Gaining Information About a Quantum Channel Via Twirling

by

Easwar Magesan

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Applied Mathematics

Waterloo, Ontario, Canada, 2008

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

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

Abstract

Finding correctable encoding that protects against a quantum process is in general a difficult task. Two main obstacles are that an exponential number of experiments are needed to gain complete information about the quantum process, and known algorithmic methods for finding correctable encodings involve operations on exponentially large matrices.

In this thesis we discuss how useful partial information of a quantum channel can be systematically extracted by averaging the channel under the action of a set of unitaries in a process known as twirling. We show that in some cases it is possible to find correctable encodings for the channel using the partial information obtained via twirling.

We investigate the particular case of twirling over the set of Pauli operators and qubit permutations, and show that the resulting quantum operation can be characterized experimentally in a scalable manner. A post-processing scheme for finding unitarily correctable codes for these twirled channels is presented which does not involve exponentially large matrices. A test for non-Markovian noise using such a twirling process is also discussed.

Acknowledgements

First and foremost I would like to thank Joseph Emerson for his guidance and assistance during my graduate studies. I would also like to thank Marcus Silva for the discussions and his willingness to listen to and answer my questions over the past two years. I am also indebted to David Kribs for his knowledge and guidance regarding my research at Waterloo. This thesis emerged from a paper with Marcus, David and Joseph, and without our frequent discussions I would not have anything to write about.

Thanks go to my readers Ray Laflamme and Achim Kempf. As well, I would like to thank Iman Marvian, Robin Blume-Kohout and anyone at IQC or PI who took the time to discuss quantum mechanics. Finally, I would be lost without the administrative support of Helen Warren, Wendy Reibel and all of the staff at applied math and IQC. Their assistance during my time at Waterloo is greatly appreciated.

Contents

List of Figures

Chapter 1

Introduction

Information in the physical world is represented by a physical system. Therefore transformations of the represented information must obey the dynamical laws that the system is governed by. Classical information theory is built upon the principles of classical physics and only recently has information been described using the laws of quantum mechanics. This new perspective on information is called quantum information theory, and is a more general theory of information than its classical counterpart. Quantum information theory allows for results that are possibly unattainable using a classical theory, as depicted by the efficient factorization of natural numbers into primes using Shor's quantum algorithm [40]. Other examples of computational speed-up using a quantum algorithm are given by Grover's search algorithm [17] and simulating the evolution of physical systems [29].

The advantages of using quantum theory for information processing can only be realized if physical implementations of a quantum information processor are possible. Such implementations have been rudimentary from the perspective of the goal of at least about 1000 qubits, and there is debate as to whether large scale quantum information processors will be a reality in the future. Two reasons for the difficulty of a large scale implementation is that quantum systems are extremely sensitive to environmental noise effects and trying to characterize the noise is exponentially hard in time.

Recent algorithms have been devised to combat noise effects once the noise model is known, however these algorithms also suffer from the drawback of being exponential in time. These methods are grouped under the title of quantum error correction, and involve finding correctable codes for the channel. The exponential nature of these algorithms is a result of requiring operations on exponentially large matrices. Thus, the exponential nature of both determining the noise model and finding methods of protecting quantum information from the noise is a central problem in the implementation of a quantum information processor.

This thesis will discuss the concept of twirling a quantum channel, which is an efficient method of partially characterizing a given channel. It will be shown that this partial characterization can lead to useful information about the channel in the form of both correctable codes and Markovicity of the channel. An algorithm for finding these correctable codes is presented that does not rely on operations on exponentially large matrices, but on known algebraic relations of Pauli operators. Finally, it is shown that the protocol is robust against experimental error.

1.1 State Space of a Quantum System

This thesis will deal only with finite dimensional quantum mechanical systems. The states of a quantum mechanical system can be represented by trace 1 positive operators acting on a complex Hilbert space $\mathcal{H}$. Elements of the Hilbert space will be labeled using Dirac notation. A vector in the Hilbert space is written in ket form $|\psi\rangle$ and the inner product of $|\psi\rangle$ with $|\phi\rangle$ is written $\langle\phi||\psi\rangle$, or more simply, $\langle\phi|\psi\rangle$. The object $\langle\phi|$ is called a bra and represents the unique linear functional f_ϕ on $\mathcal{H}$ given by

$$f_\phi(|\psi\rangle) = \langle\phi|\psi\rangle. \tag{1.1}$$

The outer product of $|\psi\rangle$ and $|\phi\rangle$ is denoted $|\phi\rangle\langle\psi|$ and by linearity of the inner product is a linear operator on $\mathcal{H}$. The set of linear operators acting on $\mathcal{H}$ is denoted by $B(\mathcal{H})$ and forms a linear space under the usual operations of addition and scalar multiplication. $B(\mathcal{H})$ can be made into a Hilbert space by defining the inner product of σ with τ to be

$$\langle\tau|\sigma\rangle = tr(\tau^\dagger\sigma) \tag{1.2}$$

where $\tau^\dagger$ is the adjoint of τ. This inner product is called the trace inner product and, unless otherwise stated, $B(\mathcal{H})$ will be assumed to have this inner product defined on it. The states ρ that satisfy

$$\rho^2 = \rho \tag{1.3}$$

are called pure states. States that are not pure are called mixed. Pure states may also be thought of as rank 1 projectors, hence the pure state $|\psi\rangle\langle\psi|$ is a representative of the equivalence class of vectors $e^{i\theta}|\psi\rangle$ in $\mathcal{H}$. Therefore the set of pure states is represented by the set of norm 1 vectors in $\mathcal{H}$, modulo phases.

1.2 Evolution

The evolution of a quantum system (assuming the system is not subject to a measurement or initially entangled with its environment) is mathematically described by a completely positive [34], trace preserving, linear map from states into states. Such a map will be called a quantum operation, or quantum channel, throughout the rest of the presentation.

A quantum channel $\Lambda : B(\mathcal{H}) \rightarrow B(\mathcal{H})$ has the following form [34]

$$\Lambda(\rho) = \sum_k A_k \rho A_k^\dagger \tag{1.4}$$

where the A_k satisfy the constraint,

$$\sum_k A_k^\dagger A_k = \mathbb{1}. \tag{1.5}$$

This representation of a quantum channel is called an operator sum decomposition. The A_k are linear operators on H called Kraus (noise) operators. The converse is also true: if a mapping is defined by a set of Kraus operators in the above manner then the mapping is completely positive, linear and trace preserving. Thus the general evolution of a quantum system under some noise model is described by an operator sum decomposition satisfying the above constraint on the Kraus operators.

There is a unitary freedom in the Kraus operators describing the quantum channel [34]. If the set of operators $\{E_i\}$ is defined by

$$E_i = \sum_k U_{ik} A_k \tag{1.6}$$

for some unitary matrix U, then $\{E_i\}$ describes the same quantum operation as the $\{A_k\}$. The converse is also true: if $\{E_i\}$ and $\{A_k\}$ define the same quantum operation then they are unitarily related as above, where the cardinality of the index sets for $\{E_i\}$ and $\{A_k\}$ can be made equal by appending the necessary number of zero operators to the smaller set. As shown below, any two sets of linearly independent Kraus operators for $\mathcal{E}$ will have index sets of equal cardinality.

Proposition 1 *For any quantum channel $\mathcal{E}$, the minimal cardinality for the index set of Kraus operators used for the operator sum decomposition of the channel is obtained by any set of linearly independent Kraus operators used to represent $\mathcal{E}$.*

Proof: Suppose $\{A_k\}_{k=1}^{n}$ is a linearly independent set of Kraus operators for the quantum operation $\mathcal{E}$ and let $\{B_j\}_{j=1}^{m}$ be another set of Kraus operators for $\mathcal{E}$ such that $m < n$. By the unitary freedom in Kraus operators we have

$$B_i = \sum_{j=1}^{n} \mu_{i,j} A_j \tag{1.7}$$

where if $i > m$ then $B_i := 0$ and $\mu_{i,j}$ is a unitary matrix. Thus, fixing $i > m$ gives $\sum_{j=1}^{n} \mu_{i,j} A_j = 0$. By the linear independence of the A_j this implies $\forall j \in \{1, ..., n\}$ $\mu_{i,j} = 0$ which contradicts the unitarity of μ. So, n is the minimal cardinality for the index set of the Kraus operators for $\mathcal{E}$.

Channels with Kraus operators satisfying the additional constraint

$$\sum_k A_k A_k^\dagger = I \tag{1.8}$$

are called unital. When a quantum system is isolated from environmental interactions the evolution is unitary. The set of Kraus operators for unitary evolution consists of a single unitary operator.

Given a linear map $\mathcal{E}$ on $B(\mathcal{H})$, the adjoint, or dual, map $\mathcal{E}^\dagger$ is uniquely defined in the usual manner by the equation

$$tr(\mathcal{E}^\dagger(\sigma)\tau) = tr(\sigma^\dagger \mathcal{E}(\tau)) \tag{1.9}$$

Proposition 2 *If $\mathcal{E}$ is completely positive with Kraus representation given by $\{A_k\}$ then the adjoint map is completely positive with Kraus operators $\{A_k^\dagger\}$.*

Proof:

We need only verify that the equation defining the adjoint map of $\mathcal{E}$ is satisfied using the set of Kraus operators $\{A_k^\dagger\}$. Indeed, by the linearity and cyclic properties of the trace we have,

$$tr(A_k^\dagger \sigma A_k \tau) = tr(A_k \tau A_k^\dagger \sigma). \tag{1.10}$$

1.3 Measurement

Measurement of a quantum system obeys different transformation rules than those described above. A measurement is described by a set of linear operators $\{M_m\}$ satisfying the completeness relation

$$\sum_m M_m^\dagger M_m = \mathbb{1}. \tag{1.11}$$

The measurement operators are indexed by the measurement outcomes m. If the state of the system is ρ, then the probability of obtaining outcome m is given by

$$p(m) = tr(M_m^\dagger M_m \rho). \tag{1.12}$$

By the completeness relation above, p(m) is a normalized probability distribution since

$$\begin{aligned} \sum_m p(m) &= \sum_m tr(M_m^\dagger M_m \rho) \\ &= tr(\sum_m M_m^\dagger M_m \rho) \\ &= tr(\rho) \\ &= 1. \end{aligned} \tag{1.13}$$

The state of the system after the measurement is

$$\frac{M_m \rho M_m^\dagger}{p(m)}. \tag{1.14}$$

The convention is to define the positive operator

$$E_m = M_m^\dagger M_m \tag{1.15}$$

and make the necessary replacement in the above expressions. The set of operators $\{E_m\}$ is called a positive operator valued measure (POVM). POVM's are especially useful when only measurement statistics are of interest.

The simplest kind of measurement is when the measurement operators consist of projection operators $P_m = |\psi_m\rangle\langle\psi_m|$ of positive rank. The P_m are uniquely defined by the fact they are Hermitian and $P_m^2 = P_m$. Such a measurement is called a projective measurement. Clearly such measurements are in a 1-1 correspondence with Hermitian operators where the eigenvalues are the measurement outcomes and projectors onto the associated eigenspace are the projectors for the measurement. The probability of outcome m is given by

$$p(m) = tr(P_m \rho) \tag{1.16}$$

and the post-measurement state is

$$\frac{P_m \rho P_m}{p(m)}. \tag{1.17}$$

1.4 Composite Quantum Systems

Let V and W be finite dimensional Hilbert spaces with bases $\{|v_i\rangle\}_{i=1}^n$ and $\{|w_j\rangle\}_{j=1}^m$ respectively. The direct sum of these two spaces has basis given by the union of the bases of its component spaces $\{|v_1\rangle, ..., |v_n\rangle, |w_1\rangle, ..., |w_m\rangle\}$. The tensor product of V and W has basis $\{|v_1\rangle \otimes |w_1\rangle, ..., |v_1\rangle \otimes |w_m\rangle, ..., |v_n\rangle \otimes |w_1\rangle, ..., |v_n\rangle \otimes |w_m\rangle\}$.

Individual state spaces of n particles combine classically through the direct sum while quantum states combine through the tensor product. Thus, the dimension of the state space of multiple classical particles grows linearly with the number of particles, since $\dim(V \times W) = \dim(V) + \dim(W)$. In the quantum case, the dimension of the composite system increases as dim(V)dim(W). The extension to multi-partite quantum systems is performed by taking the tensor product of the spaces describing each party. We will sometimes denote the t-fold tensor product of a state $|\psi\rangle$ with itself by the expression $|\psi\rangle^{\otimes t}$.

1.5 Distinguishing Quantum States

There are many methods of distinguishing quantum states [14]. Distinguishing states is often done by defining a metric, or something resembling a metric, on the space $B(\mathcal{H})$.

Definition 1 *Metric*

Let X be a set. A metric on X is a function $d : X \to \mathbb{R}$ that satisfies the following properties:

1. $\forall$ x, $y \in X$, $d(x, y) \geq 0$

2. $d(x, y) = 0$ *if and only if* $x = y$

3. $\forall$ x,$y \in X$, $d(x, y) = d(y, x)$

4. $\forall$ x,y,$z \in X$, $d(x, y) \leq d(x, z) + d(z, y)$

Functions resembling metrics on the state space give meaning to the concept of distance between states. Two important measures of distance are the trace distance, which is a metric, and the fidelity, which is not. Before defining these quantities, we look at their classical analogues.

Definition 2 *Classical Trace Distance and Fidelity*

Let $\{p_x\}$ and $\{q_x\}$ be two probability distributions over the same index set. The classical trace distance, $\mathcal{D}_C$, between the distributions is the l_1 distance between the distributions divided by 2. That is,

$$\mathcal{D}_C = \frac{1}{2} \sum_x |p_x - q_x| \tag{1.18}$$

The classical fidelity, $\mathcal{F}_C$, between the distributions is

$$\mathcal{F}_C(p_i, q_i) = \sum_i \sqrt{p_i q_i} \tag{1.19}$$

$\mathcal{F}_C$ is clearly not a metric since if $\forall i \, p_i = q_i$, $\mathcal{F}_C(p_i, q_i) = 1$. Hence, $\mathcal{F}_C(p_i, q_i)$ being near 1 indicates the probability distributions are close to each other. We now define the trace distance in the quantum case and discuss some of its properties.

Definition 3 *Trace Distance*

The trace distance $\mathcal{D}$ between two quantum states ρ and σ is one half times the metric induced by the trace inner product. Specifically,

$$\mathcal{D}(\rho, \sigma) = \frac{1}{2} tr \left|\rho - \sigma\right| \tag{1.20}$$

where as usual for $A \in B(\mathcal{H})$, $|A| = \sqrt{A^\dagger A}$.

Clearly the trace distance is unitarily invariant, ie. for all unitaries U,

$$\mathcal{D}(\rho, \sigma) = \mathcal{D}(U\rho U^\dagger, U\sigma U^\dagger). \tag{1.21}$$

The following characterizes the trace distance in terms of measurement statistics. [34].

Proposition 3 *Let ρ and σ be quantum states and $\{E_m\}$ be an arbitrary POVM. Define the probability distributions p_m and q_m by $p_m = tr(\rho E_m)$ and $q_m = tr(\rho E_m)$. Then*

$$\mathcal{D}(\rho, \sigma) = max_{\{E_m\}} \mathcal{D}_C(p_m, q_m) \tag{1.22}$$

where the maximization is over all POVMs.

Hence the trace distance is the largest possible classical trace distance between the probability distributions arising from a POVM. An important property of the trace distance is that of strong convexity [34].

Proposition 4 *Let p_m and q_m be probability distributions over the same index set. Then, for states ρ_m and σ_m defined on this index set,*

$$\mathcal{D}\left(\sum_m p_m \rho_m, \sum_m q_m \sigma_m\right) \leq \mathcal{D}_C(p_m, q_m) + \sum_m p_m \mathcal{D}(\rho_m, \sigma_m). \tag{1.23}$$

In the special case of $\forall m \, p_m = q_m$ we get

$$\mathcal{D}\left(\sum_m p_m \rho_m, \sum_m p_m \sigma_m\right) \leq \sum_m p_m \mathcal{D}(\rho_m, \sigma_m). \tag{1.24}$$

Thus the trace distance is jointly convex in its inputs. Next, we define the fidelity and discuss some analogous properties to those for the trace distance.

Definition 4 *Fidelity*

The fidelity, $\mathcal{F}$, between ρ and σ is

$$\mathcal{F}(\rho, \sigma) = tr\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}}. \tag{1.25}$$

The fidelity satisfies all of the properties of a metric except for being zero when $\rho = \sigma$. When ρ (or σ by symmetry) is a projector $|\psi\rangle\langle\psi|$ we have the following simple form for the fidelity

$$\begin{aligned} \mathcal{F}(\rho, \sigma) &= tr\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}} \\ &= tr\sqrt{\langle\psi|\rho|\psi\rangle|\psi\rangle\langle\psi|} \\ &= \sqrt{\langle\psi|\rho|\psi\rangle}. \end{aligned} \tag{1.26}$$

As with the trace distance, the fidelity is unitarily invariant. There is also an analogous characterization of the fidelity in terms of measurement statistics [34].

Proposition 5 *Let ρ and σ be quantum states and $\{E_m\}$ be an arbitrary POVM. Define the probability distributions p_m and q_m by $p_m = tr(\rho E_m)$ and $q_m = tr(\rho E_m)$. Then*

$$\mathcal{F}(\rho, \sigma) = min_{\{E_m\}}\mathcal{F}_C(p_m, q_m) \tag{1.27}$$

where the minimization is over all POVMs.

As well, analogous to the strong convexity result for the trace distance, the fidelity satisfies a strong concavity property.

Proposition 6 *Let p_m and q_m be probability distributions over the same index set. Then, for states ρ_m and σ_m defined on the same index set,*

$$\mathcal{F}\left(\sum_m p_m\rho_m, \sum_m q_m\sigma_m\right) \geq \sum_m \sqrt{p_m q_m}\mathcal{F}(\rho_m, \sigma_m). \tag{1.28}$$

In the special case of $p_m = q_m \, \forall m$ we get

$$\mathcal{F}\left(\sum_m p_m \rho_m, \sum_m p_m \sigma_m\right) \geq \sum_m p_m \mathcal{F}(\rho_m, \sigma_m). \tag{1.29}$$

Thus the fidelity is jointly concave in its inputs. We are now ready to discuss how quantum mechanics is applied to information theory

1.6 Quantum Information Theory

The fundamental unit of information in quantum information theory is a qubit, analogous to the bit in classical information theory. Physically, a qubit may by thought of as a two-dimensional quantum mechanical system. Hence it is mathematically represented by the set of trace 1 positive operators acting on a 2-dimensional complex Hilbert space. A standard physical instance of a qubit is given by photon polarization, where three sets of bases for the system are the horizontal-vertical (H/V) basis, plus-minus basis (+/-) and the right-left circular polarization (R-L) basis.

The state space for two qubits, each with basis $\{|0\rangle, |1\rangle\}$, has basis $\{|0\rangle \otimes |0\rangle, |0\rangle \otimes |1\rangle, |1\rangle \otimes |0\rangle, |1\rangle \otimes |1\rangle\}$ which will be written more compactly as $\{|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle\}$. In general an n qubit system has 2^n basis vectors. The orthonormal basis for an n qubit Hilbert space formed from tensor products of $|0\rangle$ and $|1\rangle$ is called the computational basis. More generally, we write $|x\rangle$ to mean $|b_n b_{n-1} \ldots b_0\rangle$ where b_i are the binary digits of the number x.

The state $\frac{|00\rangle+|11\rangle}{2}$ is an example of a quantum state that cannot be described in terms of the state of each of its components (qubits) separately. In other words, we cannot find a_1, a_2, b_1, b_2 such that $(a_1|0\rangle + b_1|1\rangle) \otimes (a_2|0\rangle + b_2|1\rangle) = |00\rangle + |11\rangle$ since

$$(a_1|0\rangle + b_1|1\rangle) \otimes (a_2|0\rangle + b_2|1\rangle) = a_1 a_2|00\rangle + a_1 b_2|01\rangle + b_1 a_2|10\rangle + b_1 b_2|11\rangle \tag{1.30}$$

and $a_1 b_2 = 0$ implies that either $a_1 = 0$ or $b_2 = 0$. States that cannot be decomposed in this manner are called entangled. These states are dense in the space of all quantum states [35]. This means that if an arbitrary state is chosen from the state space then any open set containing this state will also contain an entangled state. Entangled states represent situations for which we have no classical intuition.

1.6.1 Quantum Gates

As previously mentioned, the dynamics of a quantum system, when not interacting with an environment or being measured, is described by a unitary transformation. One important consequence of the fact that quantum transformations are unitary is that they are reversible.

The following are some examples of useful single-qubit quantum state transformations written in the basis $\{|0\rangle, |1\rangle\}$

$$\begin{array}{llll}
I: & |0\rangle \rightarrow |0\rangle & & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
& |1\rangle \rightarrow |1\rangle & & \\
X: & |0\rangle \rightarrow |1\rangle & & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
& |1\rangle \rightarrow |0\rangle & & \\
Y: & |0\rangle \rightarrow i|1\rangle & & \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\
& |1\rangle \rightarrow -i|0\rangle & & \\
Z: & |0\rangle \rightarrow |0\rangle & & \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \\
& |1\rangle \rightarrow -|1\rangle & &
\end{array}$$

X, Y, and Z are called the Pauli operators and the set $\{I, X, Y, Z\}$ forms an orthonormal basis for $B(\mathbb{C}^2)$ when each element is scaled by $\frac{1}{2}$. X, Y and Z are sometimes denoted as P_1, P_2, and P_3 or σ_1, σ_2 and σ_3. They satisfy the commutation and anti-commutation relations

$$[P_l, P_m] = 2i \sum_{n=1}^{3} \epsilon_{lmn} P_n \tag{1.31}$$

$$\{P_l, P_m\} = 2\delta_{l,m} \mathbb{1} \tag{1.32}$$

The controlled-NOT gate, C-NOT, operates on two qubits as follows: it flips the second qubit if the first qubit is $|1\rangle$ and leaves the second qubit unchanged when the first is $|0\rangle$. As noted, the vectors $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$ form an orthonormal basis for the set of pure states for a two-qubit system. Hence the

C-NOT transformation has representation in this basis given by

$$\begin{array}{rlcl} \text{C-NOT}: & |00\rangle & \to & |00\rangle \\ & |01\rangle & \to & |01\rangle \\ & |10\rangle & \to & |11\rangle \\ & |11\rangle & \to & |10\rangle \end{array} \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

The transformation C-NOT is unitary and cannot be decomposed into a tensor product of two single qubit transformations.

It is useful to have graphical representations of quantum state transformations, especially when several transformations are combined in sequence. This representation is given by a quantum circuit, which is read left to right in time. The number of horizontal levels in the circuit corresponds to the number of qubits involved in the computation. The following is an example of a quantum circuit.

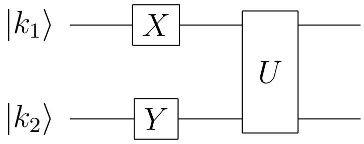

Figure 1.1: Example of a Quantum Circuit

There are two qubits, the first in the state $|k_1\rangle$ and the second in the state $|k_2\rangle$. The first qubit undergoes the unitary transformation X and the second is transformed by Y. The entire 2 qubit system then undergoes the unspecified unitary transformation U.

C-NOT is typically represented by a circuit of the form

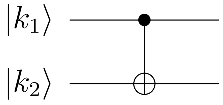

Figure 1.2: C-NOT Gate

The filled circle indicates the control qubit, and the ⊕ indicates the conditional negation of the target qubit.

The Hadamard and R_k Transformations

Another important single qubit transformation is the Hadamard Transformation defined by

$$\begin{aligned} H: \ |0\rangle &\rightarrow \tfrac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = |+\rangle \\ |1\rangle &\rightarrow \tfrac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = |-\rangle . \end{aligned}$$

The transformation H has a number of important applications. When applied to $|0\rangle$, H creates a superposition state $\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)$. Applied to n qubits individually, H generates a superposition of all 2^n possible states, which can be viewed as the binary representation of the numbers from 0 to $2^n - 1$.

$$\begin{aligned} H \otimes H \otimes \cdots \otimes H\,|00\ldots 0\rangle &= \frac{1}{\sqrt{2^n}}\left((|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle) \otimes \cdots \otimes (|0\rangle + |1\rangle)\right) \\ &= \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle . \end{aligned} \tag{1.33}$$

The R_k unitary transformation on a single qubit is given by the matrix

$$R_k = \begin{bmatrix} 1 & 0 \\ 0 & e^{-\frac{2\pi i}{2^k}} \end{bmatrix}$$

in the standard basis. We will now discuss quantum algorithms using the concepts presented in this section.

1.6.2 Quantum Algorithms

One of the main advantages of representing information through quantum systems is that certain computational problems with no known efficient classical solution are efficiently solvable using a quantum information processor. Many of the quantum algorithms that solve these problems rely on a transformation called the quantum Fourier transform [21, 34]. Due to its importance, the quantum Fourier transform and its generalization to finite abelian groups is given in the appendix.

Perhaps the most prominent computational problem is that of factoring integers. Shor's quantum algorithm for solving the factoring problem (RSA) is essentially a specific case of a more general problem called the hidden subgroup problem [22]. Due to the importance of solving this problem efficiently, an efficient quantum algorithm for solving it is given in the appendix. The algorithm relies heavily on the representation theory of finite abelian groups. Hence, basic results on group theory and representation theory of finite abelian groups are also presented in the appendix. While it is useful to know that certain computational tasks are easy on a quantum information processor, implementing such computations is extremely hard. This difficulty is due to the extreme sensitivity of quantum systems to their environment. The area of research that deals with reliably preserving information when a quantum system interacts with some environment will be examined next.

1.7 Quantum Error Correction

Quantum error correction (QEC) is a subfield of quantum information theory that deals with how to preserve quantum information when it is sent through a channel. Representing information through quantum states suffers from the drawback that a quantum system is extremely sensitive to interactions with an environment. These interactions create correlations between the system of interest and the environment which results in the environment carrying away information about the system. Thus, information initially encoded in the quantum system may be lost through such interactions. Sending information from one party to another requires that the received state of the system closely resembles the initial information. Hence we must find ways to minimize the interaction of an environment with the encoded information.

There are two types of error correction, passive and active. In passive error correction once the initial state has been encoded it only interacts with the quantum channel. Thus, the main part of the error correction procedure lies in the encoding and decoding of the quantum information. Active error correction pertains to actively manipulating the state while, or after, it interacts with the channel in order to preserve the encoded information. QEC through noiseless subsystems is a type of passive error correction where certain subsystems of the state space are located as being ”unaffected” by the quantum channel.

A method for quantum error correction is given in [25] that unifies the previously known methods of error correction under one framework. This framework is called operator quantum error correction and applies to both unital and non-unital quantum channels. The previous methods of error correction are the standard model, decoherence free subspaces and noiseless subsystems. We briefly discuss these models for error correction and then present the unified method.

1.7.1 The Standard Model

The standard model may be described by a triple $(\mathcal{R}, \mathcal{E}, \mathcal{C})$, where the code $\mathcal{C}$ is a subspace of the Hilbert space $\mathcal{H}$, $\mathcal{E}$ is a quantum channel, and $\mathcal{R}$ is a recovery channel. Denote the projection onto $\mathcal{C}$ by $P_{\mathcal{C}}$. The triple must satisfy the following for all bounded linear operators $\rho = P_{\mathcal{C}} \rho P_{\mathcal{C}}$ (ie. all ρ which are reduced by $P_{\mathcal{C}}$ and whose support lies in $\mathcal{C}$),

$$(\mathcal{R} \circ \mathcal{E})(\rho) = \rho. \tag{1.34}$$

When there exists an $\mathcal{R}$ for given $\mathcal{E}$ and $\mathcal{C}$, the code $\mathcal{C}$ is said to correct $\mathcal{E}$. In the case $\mathcal{R} = \mathcal{I}$, the triple is called a decoherence free subspace. Let $\{E_a\}$ be a set of Kraus operators for $\mathcal{E}$. Then the existence of $\mathcal{R}$ for $\mathcal{E}$ and $\mathcal{C}$ is equivalent to

$$P_{\mathcal{C}} E_a^\dagger E_b P_{\mathcal{C}} = \mu_{ab} P_{\mathcal{C}} \tag{1.35}$$

for all a, b in the index set for the Kraus representation where the matrix μ_{ab} is positive semi-definite with trace equal to 1 [34]. Since different Kraus representations for a particular CP map are related by a unitary matrix, the form of the above condition is clearly independent of the Kraus representation used.

By the unitary freedom in the Kraus operators and the fact that μ_{ab} is positive semi-definite, there exists a set of Kraus operators for the channel $\mathcal{E}$ such that μ_{ab} is diagonal. For this particular set of Kraus operators, labeled $\{G_a\}$, it is clear that the code subspace $\mathcal{C}$ is mapped to orthogonal subspaces by the G_a. So, $\mathcal{C}$ is correctable for $\mathcal{E}$ if and only if there exists a Kraus representation $\{G_a\}$ such that:

1. $\forall\ |\psi\rangle \in H$ and $a \neq b$, $G_a P\ |\psi\rangle$ is orthogonal to $G_b P\ |\psi\rangle$
2. The inner product structure on $\mathcal{C}$ is preserved by the G_a.

Thus the G_a map $\mathcal{C}$ to orthogonal undeformed copies of $\mathcal{C}$ in H, which is a useful property of the $\{G_a\}$ because the recovery operation is then easily described by a measurement in a basis determined by the orthogonal copies followed by a unitary operation [34].

Note that there may exist a set of Kraus operators for $\mathcal{E}$ such that the action of at least two of the Kraus operators on $\mathcal{C}$ is the same and $\mathcal{C}$ is still a correctable subspace for $\mathcal{E}$. This phenomenon is called degeneracy and $\mathcal{C}$ in this case is called a degenerate code. Analogous types of codes can not be found in classical error correction [34].

1.7.2 Noiseless Subsystems and Decoherence Free Subspaces

Before describing the noiseless subsystem method, let us lay down some terminology. Let $\mathcal{E}$ be a quantum operation with Kraus operators $\{E_a\}$ and suppose the Hilbert space $\mathcal{H}$ factorizes as $\mathcal{H} = (\mathcal{H}_A \otimes \mathcal{H}_B) \oplus K$, with $\dim(\mathcal{H}_A) = m$, $\dim(\mathcal{H}_B) = n$, and K a subspace of arbitrary but finite dimension. Let P_{AB} be the projector onto the subspace $\mathcal{H}_A \otimes \mathcal{H}_B$, P_{kl} be projectors of the form $|\alpha_k\rangle\langle\alpha_l| \otimes \mathbb{1}_B$ for some orthonormal basis $\{|\alpha_k\rangle\} \in \mathcal{H}_A$ and the quantum operation $\mathcal{P}_{\mathcal{AB}}$ be defined by Kraus operators $\{P_{kl}\}$. The P_{kk} are called minimal reducing projections for B and $P_{AB} = \sum_{k=1}^{m} P_{kk}$ is called the minimal central projection onto $\mathcal{H}_{AB}$. Finally, let $\mathcal{S}$ be the semigroup of operators of the form $\sigma_A \otimes \sigma_B$ which are reduced by P_{AB} and have support on $P_{AB}\mathcal{H}$.

Definition 5 *Noiseless Subsystem*

B is said to be a noiseless subsystem for $\mathcal{E}$ *if* $\forall\sigma_A\,\forall\sigma_B\,\exists\tau_A$

$$\mathcal{E}(\sigma_A \otimes \sigma_B) = \tau_A \otimes \sigma_B \tag{1.36}$$

Thus, B is a noiseless subsystem for $\mathcal{E}$ if there exists a quantum operation $\mathcal{F}_{\mathcal{AA}} : B(\mathcal{H}_A) \to B(\mathcal{H}_A)$ such that $\mathcal{E}|_{B(\mathcal{H}_A)\otimes B(\mathcal{H}_B)} = \mathcal{F}_{\mathcal{AA}} \otimes \mathbb{1}$. The following proposition is proved in [25]

Proposition 7 *The following four conditions are equivalent to B being a noiseless subsystem for the quantum process* $\mathcal{E}$*:*

1. $\forall\sigma_B\,\exists\tau_A\ : \mathcal{E}(\mathbb{1}_A \otimes \sigma_B) = \tau_A \otimes \sigma_B$

2. $\forall\sigma \in \mathcal{S} : Tr_A \circ \mathcal{P}_{\mathcal{AB}} \circ \mathcal{E}(\sigma) = Tr_A(\sigma)$

3. $\forall a : E_a$ *is invariant on* $P_{AB}\mathcal{H}$ *and* $E_a|_{P_{AB}\mathcal{H}} \in B(\mathcal{H}_A) \otimes \mathbb{1}_B$

4. $\forall a,k,l : E_a P_{AB} = P_{AB} E_a P_{AB}$ *and* $P_{kk} E_a P_{ll} = \lambda_{akl} P_{kl}$ *where* λ_{akl} *is some set of scalars.*

The noiseless subsystem framework given above encompasses the notion of decoherence free subspaces in the case when $m = 1$. This is easy to see from the first condition using $\mathcal{H}_A \otimes \mathcal{H}_B \cong \mathcal{H}_B$ and trace preservation. Hence when $m = 1$, the B subsystem is actually a subspace which is undeformed by the action of $\mathcal{E}$.

1.7.3 Unified Method For Quantum Error Correction

The unified scheme that encompasses all of these models is described by a triple $(\mathcal{R}, \mathcal{E}, \mathcal{S})$ where as in the terminology for the standard model, $\mathcal{R}$ is a recovery quantum operation for the channel $\mathcal{E}$. $\mathcal{S}$ is a semigroup of operators defined as above in the noiseless subsystems section.

Definition 6 *Correctable Code*

For a triple $(\mathcal{R}, \mathcal{E}, \mathcal{S})$, the B subsystem is called correctable for $\mathcal{E}$ by the recovery operation $\mathcal{R}$ if it is noiseless for the quantum operation $\mathcal{R} \circ \mathcal{E}$. Concretely, using a definition of a noiseless subsystem given above, B is correctable of $\mathcal{E}$ by $\mathcal{R}$ if

$$\forall \sigma_B \, \exists \tau_A \; : \mathcal{R} \circ \mathcal{E}(\mathbb{1}_A \otimes \sigma_B) = \tau_A \otimes \sigma_B \tag{1.37}$$

The standard model is encompassed within this framework in the case when $\dim(\mathcal{H}_A) = m = 1$. When $\mathcal{R} = \mathcal{I}$, B is a noiseless subsystem. If both $\mathcal{R} = \mathcal{I}$ and $m = 1$ then B is a decoherence free subspace.

The case when $\mathcal{R}$ can be chosen to be a unitary operation $\mathcal{U}$ is of particular interest and in such a scenario the subsystem B is said to be a unitarily correctable subsystem (UCS). A UCS code is called a unitarily noiseless subsystem (UNS) for Λ if it is a UCS of Λ^n for all $n \geq 1$. As usual, Λ^n is the channel Λ composed with itself n times.

When $\mathcal{E}$ is a unital quantum channel we have the following result [26] that will be useful in finding correctable codes later on.

Theorem 1 *The following are equivalent:*

1. *B is a unitarily correctable subsystem for $\mathcal{E}$*
2. *B is a noiseless subsystem for $\mathcal{E}^\dagger \circ \mathcal{E}$.*

In addition to having a framework for error correction, it is necessary to be able to find correctable codes for the theory to be of any practical interest. In this thesis we are interested in finding correctable codes for unital channels because Pauli twirling a quantum channel results in a unital channel. The algorithm we will present later finds both noiseless subsystems and unitary correctable subsystems but is not exhaustive. For unital channels there is an algorithm that finds all noiseless subsystems for the channel [19]. The main drawback of the algorithm is that it is exponential in the number of qubits.

The algorithm that finds all noiseless subsystems for a unital quantum channel is presented in the appendix, along with the required background on algebra theory. The key is that for unital channels the commutant and fixed point set of the channel coincide. By the Artin-Wedderburn theorem and the fact that the commutant is a finite-dimensional C^*-algebra, the matrix algebras in the Artin-Wedderburn decomposition for the commutant are areas in the Hilbert space in which noiseless quantum information may be stored. Thus the algorithm consists of how to find this decomposition of the commutant. It was proved in [8] that every noiseless subsystem for a unital channel must reside in the commutant. Hence this algorithm finds all noiseless subsystems for a unital channel. For non-unital channels, there is an algorithm for finding noiseless subsystems [8, 23].

Unitary t-designs are introduced next, which will naturally lead into the concept of twirling a quantum channel. After specific types of twirling are discussed, the above concepts in QEC will be applied to prove various theorems regarding twirling and to give an algorithm for finding UCS codes.

Chapter 2

Unitary t-Designs and Twirling Quantum Channels

In this chapter we introduce unitary t-designs and prove some basic results about them. Unitary t-designs naturally lead to the concept of twirling a quantum channel over a subset of the unitary group U(D). We will discuss twirling over all of U(D) using the Haar measure and then look at applications that involve estimating the average gate fidelity of a quantum gate. In the next chapter we will discuss twirling over discrete subsets of the unitary group.

2.1 Unitary t-Designs

Before defining unitary t-designs we discuss the well-known concept in numerical analysis of spherical t-designs [11]. Suppose one has a function defined on the unit sphere $S^{n-1} \subseteq \mathbb{R}^n$ and wants to compute the average of the function. The maximal symmetry of the domain suggests that for "well-behaved" classes of functions, there should exist a set of fixed points on $\mathbb{S}^{n-1}$ such that for any function f in the class, the average of the values of f at these points is equal to the global average of f on $\mathbb{S}^{n-1}$.

Important functions in numerical analysis are polynomials, which are divided into a countable number of classes by their degree. A natural class of functions to analyze the existence of such a set of points is the set of polynomials of degree t. These sets are called spherical t-designs. Formally, this is stated below

Definition 7 *Spherical t-Design*

A spherical t-design is a finite set of points $\{x_1, ..., x_K\} \subseteq \mathbb{S}^{n-1}$ such that for any polynomial $p : \mathbb{S}^{n-1} \to \mathbb{R}$ of degree less than or equal to t, the average of p over $\mathbb{S}^{n-1}$ with respect to the rotationally-invariant Haar measure (see appendix) is equal to the average of the polynomial values at each x_i. The polynomials defined on $\mathbb{S}^{n-1}$ are just the set of all polynomials of degree less than or equal to t defined on $\mathbb{R}^n$, but restricted to $\mathbb{S}^{n-1}$.

In the case of $\mathbb{S}^2$ we require that for any polynomial p of degree less than or equal to t whose domain is $\mathbb{R}^3$,

$$\int_0^{2\pi} \int_0^{\pi} p(\theta, \phi) sin\theta d\theta d\phi = \frac{1}{K} \sum_{j=1}^{K} p(x_j) \tag{2.1}$$

where we utilize the usual spherical coordinate system on the sphere. A specific example is given by the 3-design for $\mathbb{S}^2$ where K = 6 and the x_j are chosen so that they form the vertices of a regular octahedron. It has been proved [39] that spherical t-designs exist of sufficiently large sizes. More precisely, there exists a number N(n,t) such that ∀N ≥ N(n,t) there exists a spherical t-design of N points on $\mathbb{S}^n$. Only estimates of the size of N(n,t) exist.

A unitary t-design is similar in principle to that of a spherical t-design. Let $\mathcal{H} = \mathbb{C}^D$ and recall that a homogeneous polynomial is one in which all the monomials making up the polynomial have the same degree. The definition of a unitary t-design is as follows [9],

Definition 8 *Unitary t-Design*

A unitary t-design is a finite set $\{U_1, ..., U_K\} \subseteq U(D)$ of unitary matrices such that for every homogeneous complex-valued polynomial p in $2D^2$ indeterminates of degree (s,s) less than or equal to (t,t),

$$\frac{1}{K} \sum_{j=1}^{K} p(U_j) = \int_{U(D)} p(U) dU \tag{2.2}$$

The integral is taken with respect to the Haar measure on U(D). p(U) is defined to be the evaluation of p at the $2D^2$ matrix entries, and their complex conjugates, of U. That is, without loss of generality, if the indeterminates are labelled $x_1, ..., x_{2D^2}$ we can relabel the them by the mapping.

$$\begin{array}{rcl} x_1 & \rightarrow & U_{1,1} \\ x_2 & \rightarrow & U_{1,2} \\ & . & \\ & . & \\ x_{D^2} & \rightarrow & U_{D,D} \\ x_{D^2+1} & \rightarrow & \overline{U_{1,1}} \\ x_{D^2+2} & \rightarrow & \overline{U_{1,2}} \\ & . & \\ & . & \\ x_{2D^2} & \rightarrow & \overline{U_{D,D}} \end{array} \tag{2.3}$$

Then the evaluation of p comes from choosing a specific U $\in$ U(D) as in the definition. Under this association, p having degree equal to (s,s) means that each monomial has 2s indeterminates where s of them are from the set $\{U_{1,1}, ..., U_{D,D}\}$ and the remaining s must come from $\{\overline{U_{1,1}}, ..., \overline{U_{D,D}}\}$. The following gives an equivalent characterization of a unitary t-design. The proof is an obvious extension of Corollary 5.2.2 in [9]. We give it here for completeness.

Proposition 8 *$\{U_1, ..., U_K\}$ is a unitary t-design if and only if $\forall s \in \{0, ..., t\}$, $\forall m, n \in \{1, ..., D\}$ and $\forall \rho \in B(\mathcal{H}^{\otimes s})$,*

$$\frac{1}{K}\sum_{j=1}^{K} P_{m,n}\left(U_j^{\otimes s}\rho U_j^{\otimes s\dagger}\right) = \int_{U(D)} P_{m,n}\left(U^{\otimes s}\rho U^{\otimes s\dagger}\right) dU. \tag{2.4}$$

Here we have denoted the s-fold tensor product of an operator A with itself by $A^{\otimes s}$, and $P_{m,n}$ corresponds to the projector onto the (m,n) entry of a matrix.

Proof:

First, suppose that $\{U_j\}_{j=1}^{K}$ form a unitary t-design. Note that the entries of $U^{\otimes s}$ are just the set of all monomials of degree s evaluated at the matrix entries of U. Similarly, the entries of $U^{\otimes s\dagger}$ are just the set of all monomials of degree s evaluated at the conjugates of the matrix entries of U. Thus, the matrix entries of $U^{\otimes s}\rho U^{\otimes s\dagger}$ are homogeneous degree (s,s) polynomials in the $2D^2$ indeterminates given by the entries of U and $U^\dagger$. This shows that for each $m, n \in \{1, ..., D\}$,

$$\frac{1}{K}\sum_{j=1}^{K} P_{m,n}\left(U_j^{\otimes t}\rho U_j^{\otimes t\dagger}\right) = \int_{U(D)} P_{m,n}\left(U^{\otimes t}\rho U^{\otimes t\dagger}\right) dU. \tag{2.5}$$

The converse is also simple. Note that every Hermitian matrix is a (real) linear combination of states. Hence, if

$$\frac{1}{K}\sum_{j=1}^{K} P_{m,n}\left(U_j^{\otimes s}\rho U_j^{\otimes s\dagger}\right) = \int_{U(D)} P_{m,n}\left(U^{\otimes s}\rho U^{\otimes s\dagger}\right) dU \tag{2.6}$$

holds for all states ρ, then it holds for every hermitian matrix. The fact that there exists a Hermitian basis for $B(\mathcal{H}^{\otimes s})$ implies that the statement holds for any linear operator A in $B(\mathcal{H}^{\otimes s})$. Finally, any monomial of degree (s,s) in $2D^2$ indeterminates can be constructed in one of the D^2 entries of $\left(U^{\otimes t}AU^{\otimes t\dagger}\right)$ by choosing A appropriately. By linearity, the definition for a unitary t-design is satisfied.

It can be seen in a manner similar to the above proof that the condition

$$\begin{aligned}&\frac{1}{K}\sum_{j=1}^{K} P_{m,n}\left(U_j^{\dagger}M_1U_jM_2...U_j^{\dagger}M_{2s-1}U_j\right)\\ &= \int_{U(D)} P_{m,n}\left(U_j^{\dagger}M_1U_jM_2...U_j^{\dagger}M_{2s-1}U_j\right) dU\end{aligned} \tag{2.7}$$

holding for all $s \in \{0,...,t\}$, all m,n $\in \{1,...,D\}$ and all linear operators $M_1,..,$ M_{2s-1} is equivalent to a t-design. Indeed, the definition of a t-design clearly implies the above and conversely any monomial of degree (s,s) can be constructed by appropriately choosing the M_1 through M_{2s-1}.

For $D = 2^n$, exact unitary t-designs have been constructed for $t = 1$ and $t = 2$ [9]. The Clifford group forms a unitary 2-design while the Pauli group forms a 1-design. It is unknown whether there exist unitary t-designs for $t \geq 3$.

From above it can be shown that $\{U_1,...,U_K\}$ satisfying the condition for a unitary 2-design is equivalent to

$$\frac{1}{K}\sum_{j=1}^{K} P_{m,n}\left(U_j\Lambda\left(U_j^{\dagger}\rho U_j\right)U_j^{\dagger}\right) = \int_{U(D)} P_{m,n}\left(U\Lambda\left(U^{\dagger}\rho U\right)U^{\dagger}\right) dU \tag{2.8}$$

being satisfied for any quantum channel Λ and any state ρ [9]. This naturally leads to the concept of twirling.

2.2 Twirling Quantum Channels

Twirling a quantum channel Λ consists of averaging Λ under the composition $\mathcal{U} \circ \Lambda \circ \mathcal{U}^\dagger$ for unitary operations $\mathcal{U}(\rho) = U\rho U^\dagger$ chosen according to some probability distribution [4, 9]. The averaged channel

$$\begin{aligned} \bar{\Lambda}(\rho) &= \int_{U(D)} d\mu(\mathcal{U})\, \mathcal{U} \circ \Lambda \circ \mathcal{U}^\dagger(\rho) \\ &= \int_{U(D)} d\mu(U)\, U\Lambda(U^\dagger \rho U)U^\dagger \end{aligned} \tag{2.9}$$

is known as the "twirled channel".

The case where the distribution over unitaries is discrete is of practical interest. In this case, the twirled channel is given by $\bar{\Lambda}(\rho) = \sum_i \mathrm{pr}(\mathcal{U}_i)\, \mathcal{U}_i \circ \Lambda \circ \mathcal{U}_i^\dagger(\rho)$, where $\{\mathrm{pr}(\mathcal{U}_i)\}$ is a probability distribution over the $\mathcal{U}_i$.

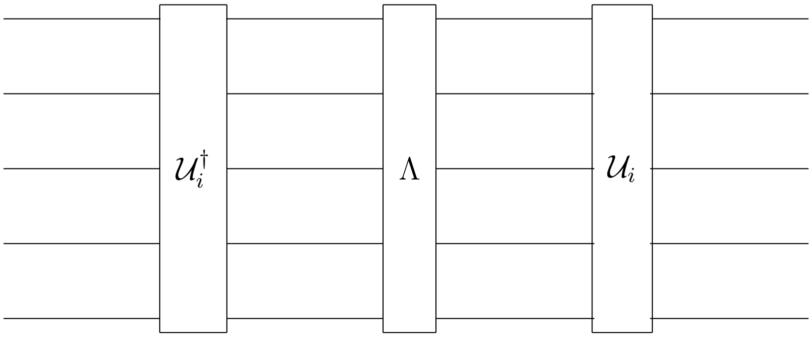

Figure 2.1: Twirling a Quantum Channel

Hence, from the previous section we have the following proposition,

Proposition 9 $\{U_1, ..., U_K\}$ *forms a unitary 2-design if and only if for any quantum channel* Λ*, the uniform twirl of* Λ *over* $\{U_1, ..., U_K\}$ *is equal to the full Haar twirl of* Λ*.*

Twirling a quantum channel over both continuous and discrete sets of unitaries has been analyzed. We briefly discuss the case of twirling over the continuous unitary group U(D) with some applications before looking at the discrete case.

2.3 Twirling Quantum Channels Over U(D)

Twirling a quantum channel Λ over U(D) is given by the expression

$$\bar{\Lambda}(\rho) = \int_{U(D)} d\mu_H(U)\, U\Lambda(U^\dagger \rho U)U^\dagger \tag{2.10}$$

where the measure μ_H on U(D) is the unique bi-invariant normalized Haar measure. Since the set of twirling elements is the full unitary group one would expect that the channel resulting from this twirl has a high amount of symmetry, or more simply, the resulting channel can be described by a small number of parameters. This is in fact the case as twirling over U(D) results in $\bar{\Lambda}$ being a depolarizing channel [12]

$$\bar{\Lambda}(\rho) = p\rho + (1-p)\frac{\mathbb{1}}{D} \tag{2.11}$$

where $p \in [0,1]$. We use this result in the next section to show a method of calculating the average fidelity between a quantum operation and a unitary operation [12].

2.3.1 Average Gate Fidelity

In many instances it is useful to estimate how "close" a quantum operation is to a given unitary operation. For instance during a quantum computation one may want to implement the unitary gate $\mathcal{U}$ and, as with any physical implementation, this will not be perfect. Thus the resulting operation will be represented by a channel $\mathcal{E}$. It would be useful to have an idea of how well $\mathcal{E}$ approximates $\mathcal{U}$. As noted previously there are many different measures of how close two quantum operations are [14], two of which, the fidelity and trace distance, were discussed in the introduction. One important measure is related to the fidelity and is called the average gate fidelity.

Definition 9 *Average Gate Fidelity*

The average gate fidelity between $\mathcal{E}$ and $\mathcal{U}$, $\overline{F_g}(\mathcal{E},\mathcal{U})$, is given by

$$\overline{F_g}(\mathcal{E},\mathcal{U}) = \int d\psi \; tr\left(\mathcal{U}\left(|\psi\rangle\langle\psi|\right)\mathcal{E}\left(|\psi\rangle\langle\psi|\right)\right). \tag{2.12}$$

The integral is over the unitarily invariant Fubini-Study measure on pure states [3]. The above can be rewritten as

$$\overline{F_g}(\mathcal{E},\mathcal{U}) = \int d\psi \; tr\left(\mathcal{U}\left(|\psi\rangle\langle\psi|\right)\Lambda \circ \mathcal{U}\left(|\psi\rangle\langle\psi|\right)\right) \tag{2.13}$$

where Λ is defined by

$$\Lambda := \mathcal{E} \circ \mathcal{U}^{-1} \tag{2.14}$$

and is called the cumulative noise operator for $\mathcal{E}$ and $\mathcal{U}$.

Now we can again re-write the average fidelity as,

$$\begin{aligned} \overline{F_g}(\mathcal{E},\mathcal{U}) &= \int d\psi \; tr\left(U|\psi\rangle\langle\psi|U^{-1}\Lambda\left(U|\psi\rangle\langle\psi|U^{-1}\right)\right) \\ &= \int d\psi \; tr\left(|\psi\rangle\langle\psi\,|U^{-1}\left(\Lambda\left(U|\psi\rangle\langle\psi|U^{-1}\right)\right)U\right) \\ &= \int d\psi \; tr\left(|\psi\rangle\langle\psi|\,\mathcal{U}^{-1}\circ\Lambda\circ\mathcal{U}\left(|\psi\rangle\langle\psi|\right)\right). \end{aligned} \tag{2.15}$$

where U is the unitary Kraus operator for $\mathcal{U}$. Note that the last two expressions for $\overline{F_g}(\mathcal{E},\mathcal{U})$ involve a twirl of Λ. This will be used to prove the following theorem [12].

Theorem 2

$$\overline{F_g}(\mathcal{E},\mathcal{U}) = \frac{\sum_k |\mathrm{Tr}(A_k)|^2 + D}{D^2 + D}, \tag{2.16}$$

where the $\{A_k\}$ are a set of Kraus operators for Λ.

Proof:

Note that we can integrate over the Haar measure on $U(D)$,

$$
\begin{aligned}
\overline{F_g}(\mathcal{E},\mathcal{U}) &= \int_{U(D)} dU\ tr\left(\rho\, U^{-1}\left(\Lambda\left(U\rho U^{-1}\right)\right)U\right) \\
&= tr\left(\rho\int_{U(D)} dU\ \ U^{-1}\left(\Lambda\left(U\rho U^{-1}\right)\right)U\right) \qquad (2.17)
\end{aligned}
$$

for $\rho = |\psi\rangle\langle\psi\,|$ *fixed. Now,*

$$
\int_{U(D)} dU\ \ U^{-1}\left(\Lambda\left(U\rho U^{-1}\right)\right)U \qquad (2.18)
$$

is the definition of twirling Λ *over the full unitary group* $U(D)$*. As noted, twirling a quantum channel over* $U(D)$ *produces an operation* $\bar{\Lambda}$ *which is a depolarizing channel. Hence,*

$$
\overline{F_g}(\mathcal{E},\mathcal{U}) = tr\left(\rho\bar{\Lambda}(\rho)\right) \qquad (2.19)
$$

for

$$
\bar{\Lambda}\rho = p\rho + (1-p)\frac{\mathbb{1}}{D} \qquad (2.20)
$$

where from [12]

$$
p = \frac{\sum_k |tr(A_k)|^2 - 1}{D^2 - 1}. \qquad (2.21)
$$

Hence

$$
\begin{aligned}
\overline{F_g}(\mathcal{E},\mathcal{U}) &= tr\left(\rho\left(p\rho + (1-p)\frac{\mathbb{1}}{D}\right)\right) \\
&= p + \frac{1-p}{D} \\
&= \frac{\sum_k |\mathrm{Tr}(A_k)|^2 + D}{D^2 + D} \qquad (2.22)
\end{aligned}
$$

as desired.

p is called the noise strength parameter and characterizes how well $\mathcal{E}$ approximates $\mathcal{U}$. If p is close to 1 then $\mathcal{E}$ can be thought of as approximating $\mathcal{U}$ well, while if p is close to 0 then $\mathcal{E}$ does not resemble $\mathcal{U}$.

We now describe a protocol for estimating the average gate fidelity $\overline{F_g}(\mathcal{E},\mathcal{U})$. We have

$$\overline{F_g}(\mathcal{E},\mathcal{U}) = \int_{U(D)} dU \; tr \left(U\rho\, U^{-1}\left(\Lambda\left(U\rho U^{-1}\right)\right)\right) \tag{2.23}$$

which is the average over the unitary group of the function $f : U(D) \to \mathbb{R}$ defined by

$$f(U) = tr\left(\rho\, U^{-1}\left(\Lambda\left(U\rho U^{-1}\right)\right)U\right). \tag{2.24}$$

ρ may be taken as the computational basis state $|0\rangle|0\rangle...|0\rangle\langle 0|...\langle 0|\langle 0|$. Note that there can be a dependence of Λ on U. In the special case of Λ being independent of the choice of U, we can use the concentration of measure effect for large Hilbert spaces (see appendix 2). This effect essentially gives that if we choose a unitary operator at random and implement it, then f(U) will be close to the average $\overline{F_g}(\mathcal{E},\mathcal{U})$ in that

$$f(U) = p + \frac{1-p}{D} + \mathrm{O}\left(\frac{1}{\sqrt{D}}\right). \tag{2.25}$$

There are two drawbacks to the above protocol for determining the average fidelity of $\mathcal{E}$ and U. The first is that choosing a random unitary over the Haar measure is exponentially hard in the system dimension D. The second is that even if one calculates f(U) for some unitary U then they will have to either perform a measurement in the basis defined by U acting on the computational basis, or perform full process tomography on the output of $\mathcal{E}(\rho) = \Lambda\circ\mathcal{U}(\rho)$. Both of these are infeasible in general. In the case of implementing motion-reversal transformations, the second difficulty can be overcome as shown at the end of the next section.

2.3.2 Motion-Reversal Transformations

While the expression for $\overline{F_g}(\mathcal{E},\mathcal{U})$ can be thought of as the average fidelity between $\mathcal{I}$ and $\mathcal{U}^{-1}\circ\Lambda\circ\mathcal{U}$, one can't estimate $\overline{F_g}(\mathcal{E},\mathcal{U})$ by implementing $\mathcal{U}^{-1}\circ\mathcal{U}$ as there will be some noise associated with $\mathcal{U}^{-1}$. More precisely, assume that the target operation is the "motion-reversal" transformation $\mathcal{U}^{-1}\circ\mathcal{U}$. We can decompose the noise affecting the transformation in two steps as previously done. In the first step, $\mathcal{U}$ is implemented and as above,

$$\mathcal{E} = \Lambda \circ \mathcal{U}. \tag{2.26}$$

The next step is to act $\mathcal{U}^{-1}$. Suppose the actual operation performed is $\mathcal{G}$ and let $\mathcal{U} \circ \mathcal{G} = \Phi$ for some quantum operation Φ. Then,

$$\mathcal{G} = \mathcal{U}^{-1} \circ \Phi. \tag{2.27}$$

This gives that the actual operation performed is

$$\mathcal{G} \circ \mathcal{E} = \mathcal{U}^{-1} \circ \Phi \circ \Lambda \circ \mathcal{U} \tag{2.28}$$

and so the noise affecting the implementation of $\mathcal{I} = \mathcal{U}^{-1} \circ \mathcal{U}$ is

$$\tilde{\Lambda} := \Phi \circ \Lambda \tag{2.29}$$

Hence, the average fidelity between $\mathcal{G} \circ \mathcal{E}$ and $\mathcal{I} = \mathcal{U}^{-1} \circ \mathcal{U}$ is,

$$\overline{F_g}(\mathcal{G} \circ \mathcal{E}, \mathcal{U}^{-1} \circ \mathcal{U}) = \int d\psi \; tr\left(|\psi\rangle\langle\psi| \, \mathcal{U}^{-1} \circ \Phi \circ \Lambda \circ \mathcal{U}\left(|\psi\rangle\langle\psi|\right)\right) \tag{2.30}$$

which is not $\overline{F_g}(\mathcal{E}, \mathcal{U})$. Thus, implementing $\mathcal{U}^{-1} \circ \mathcal{U}$ and comparing with the identity does not give a method for estimating $\overline{F_g}(\mathcal{E}, \mathcal{U})$.

Motion-reversal transformations give useful information about the strength of a noise process [12, 24], hence estimating $\overline{F_g}(\mathcal{G} \circ \mathcal{E}, \mathcal{U}^{-1} \circ \mathcal{U})$ is of interest. Following the protocol given in the average fidelity section above we have that

$$\overline{F_g}(\mathcal{G} \circ \mathcal{E}, \mathcal{U}^{-1} \circ \mathcal{U}) = \int_{U(D)} dU \; tr\left(\rho \, U^{-1}\left(\tilde{\Lambda}\left(U \rho U^{-1}\right)\right) U\right) \tag{2.31}$$

and we define the function $f : U(D) \rightarrow \mathbb{R}$ by

$$f(U) = tr\left(\rho \, U^{-1}\left(\tilde{\Lambda}\left(U \rho U^{-1}\right)\right) U\right). \tag{2.32}$$

with $\rho = |0\rangle|0\rangle...|0\rangle\langle 0|...\langle 0|\langle 0|$. Again, if $\tilde{\Lambda}$ is independent of the unitary U chosen then we pick a unitary U at random and calculate f(U). By concentration of measure,

$$f(U) = p + \frac{1-p}{D} + \mathrm{O}\left(\frac{1}{\sqrt{D}}\right). \tag{2.33}$$

Since we are implementing a motion-reversal transformation, f(U) is calculated by taking the inner product of $\mathcal{G} \circ \mathcal{E}(\rho) = U^{-1}\left(\tilde{\Lambda}\left(U\rho U^{-1}\right)\right)U(\rho)$ with ρ, which is just the probability of obtaining ρ from a measurement in the computational basis. Hence the second difficulty described in the section on average gate fidelity can be overcome for a motion-reversal transformation.

An extension of the motion-reversal protocol is when one implements a string of motion-reversal transformations,

$$\rho(n) = \left(\mathcal{U}_n^{-1} \circ \tilde{\Lambda}_n \circ \mathcal{U}_n\right) \circ ... \circ \left(\mathcal{U}_2^{-1} \circ \tilde{\Lambda}_2 \circ \mathcal{U}_2\right) \circ \left(\mathcal{U}_1^{-1} \circ \tilde{\Lambda}_1 \circ \mathcal{U}_1\right)\rho(0). \tag{2.34}$$

We take $\rho(0) = |0\rangle|0\rangle...|0\rangle\langle 0|...\langle 0|\langle 0|$. Each $\tilde{\Lambda}_i$ is the noise arising from implementing the i'th motion reversal transformation and will be assumed to be independent of the choice of unitary chosen in the i'th time step. The above discussion has shown how to estimate the average fidelity of each motion reversal transformation. An interesting question is how the average fidelity of the composition of motion-reversal transformations with the identity behaves as n grows large.

Denote the string of motion-reversal transformations by $\mathcal{K}$. Then the fidelity of $\mathcal{K}$ with the identity is

$$\begin{aligned} F(\rho(0), \mathcal{K}\rho(0)) &= tr\left(\rho(0)\mathcal{K}\rho(0)\right) \qquad (2.35)\\ &= tr\left(\rho(0)\left(\mathcal{U}_n^{-1} \circ \tilde{\Lambda}_n \circ \mathcal{U}_n\right) \circ ... \circ \left(\mathcal{U}_1^{-1} \circ \tilde{\Lambda}_1 \circ \mathcal{U}_1\right)\rho(0)\right) \end{aligned}$$

If one integrates each motion reversal sequence over the Haar measure to obtain the average fidelity we end up with a composition of n depolarizing channels

$$\overline{F_g}(\mathcal{G}, \mathcal{I}) = tr\left(\rho(0)\overline{\tilde{\Lambda}}_1 \circ ... \circ \overline{\tilde{\Lambda}}_n\left(\rho(0)\right)\right) \tag{2.36}$$

$$\overline{\tilde{\Lambda}}_j\rho = p_j\rho + (1-p_j)\frac{\mathbb{1}}{D} \tag{2.37}$$

Hence if we make the simplifying assumption that the $\tilde{\Lambda}_j$ are all the same then we define $\forall j \;\; p_j = p$ to obtain

$$\overline{F_g}(\mathcal{G},\mathcal{I}) = p^n + \frac{1-p^n}{D}. \tag{2.38}$$

As $n \rightarrow \infty$, $\overline{F_g}(\mathcal{G},\mathcal{I}) \rightarrow \frac{1}{D}$ which is the average fidelity between randomly chosen states. Since p is a measure of how strong the cumulative noise is, the decay property of p^n is a measure of how strong the noise process is. Slow convergence of p^n to 0 implies weak noise whereas fast convergence implies strong noise. This decay property is utilized in [24]. By concentration of measure, if D is large then implementing n random motion-reversal sequences and performing a measurement in the computational basis will give a value that is close to the average fidelity in the sense shown above for a single motion-reversal transformation. Again, note that the above protocol is not strictly efficient because it is exponentially hard to generate a Haar random unitary.

In the next chapter we look at twirling quantum channels over discrete subsets of the unitary group and how to gain information about a channel via a twirling procedure. Discrete twirls are easier to implement experimentally and so are of larger practical interest than twirls over the full unitary group U(D).

Chapter 3

Gaining Information About a Quantum Channel Via Twirling

This chapter is based mainly on [41], however we give more detail and expand upon many of the results. We begin by discussing twirls over discrete subsets of U(D), in particular we will analyze twirling over both the Pauli group and random qubit permutations. We show that when a channel is twirled over any discrete subset of U(D), correctable codes for the twirled channel are correctable for the original channel, up to some unitary correction. We then prove some basic results about channels twirled over the Pauli group and random permutations that lead to an algorithm for finding certain correctable codes. This scheme is proved to be robust under experimental error. Finally, we show that Markovian quantum channels satisfy a specific composition law when twirled over the Pauli group and random permutations.

3.1 Twirling Quantum Channels Over the Pauli and Permutation Groups

Before discussing twirling over discrete subsets of U(D) let us lay down some notation. Let $\mathcal{P}_1$ be the set of single qubit Pauli operators adjoined with the identity operation, that is, $\mathcal{P}_1 = \{\mathbb{1}, X, Y, Z\}$. Define $\mathcal{P}_1^{\otimes n}$ to be all n-fold tensor products of $\mathcal{P}_1$. Up to phases of -1, i, and -i, $\mathcal{P}_1$ and $\mathcal{P}_1^{\otimes n}$ form groups under multiplication. Let $\mathcal{C}_1$ be the normalizer of the group $\mathcal{P}_1$, $\mathcal{C}_n$ be the normalizer of $\mathcal{P}_1^{\otimes n}$, and $\mathcal{C}_1^{\otimes n}$ be all n-fold tensor products of $\mathcal{C}_1$.

It was shown in [9] that twirling Λ over $\mathcal{C}_n$ results in a depolarizing channel that is characterized by a single paramenter. This depolarizing channel is the same one that is obtained when twirling over the the full unitary group. Hence from a previous characterization, $\mathcal{C}_n$ is a unitary 2-design. Recently, it has been shown that twirling Λ over both $\mathcal{C}_1^{\otimes n}$ and random qubit permutations reduces the number of parameters describing the twirled channel $\bar{\Lambda}_\Pi$ to $n+1$ [13]. We will now analyze the effect of twirling Λ over both $\mathcal{P}_1^{\otimes n}$ and random qubit permutations.

Suppose Λ is an n-qubit quantum channel described by Kraus operators $\{A_k\}$. Expanding the Kraus operators in terms of the Hermitian basis $\mathcal{P}_1^{\otimes n}$ we have

$$A_k = \sum_l \gamma_l^k P_l \tag{3.1}$$

and so,

$$\begin{aligned}
\Lambda(\rho) &= \sum_k A_k \rho A_k^\dagger \\
&= \sum_k \left(\sum_l \gamma_l^k P_l\right) \rho \left(\sum_m \gamma_m^{k\,*} P_m\right) \\
&= \sum_{l,m} \left(\sum_k \gamma_l^k \gamma_m^{k\,*}\right) P_l \rho P_m \\
&:= \sum_{l,m} \chi_{l,m} P_l \rho P_m
\end{aligned} \tag{3.2}$$

where we have defined

$$\chi_{l,m} = \sum_k \gamma_l^k \gamma_m^{k\,*}. \tag{3.3}$$

Trace preservation of Λ requires that $\sum_k A_k^\dagger A_k = \mathbb{1}$ which implies

$$\sum_{l,m} \left(\sum_k \gamma_l^k \gamma_m^{k\,*}\right) P_m P_l = \mathbb{1}. \tag{3.4}$$

Hence,

$$\sum_{l,m} \chi_{l,m} P_m P_l = \mathbb{1} \tag{3.5}$$

Taking the trace of both sides gives

$$tr(\chi) = 1 \tag{3.6}$$

and since $\chi_{m,m} = \sum_k \gamma_m^k \gamma_m^{k\,*}$, the diagonals of the chi matrix are non-negative and add up to 1. Hence the diagonal elements can be interpreted as probabilities and the set $\{\chi_{i,i} P_i\}$ define a set of Kraus operators for a quantum operation.

We can obtain the operation defined by $\{\chi_{i,i} P_i\}$, which will be denoted $\bar{\Lambda}$, from Λ by uniformly twirling over $\mathcal{P}_1^{\otimes n}$ [9],

$$\begin{aligned} \bar{\Lambda}(\rho) &= \frac{1}{4^n} \sum_i \sum_k P_i A_k P_i \rho P_i A_k^\dagger P_i \\ &= \sum_i \chi_{ii} P_i \rho P_i. \end{aligned} \tag{3.7}$$

The twirl over $\mathcal{P}_1^{\otimes n}$ strips away the off-diagonals of the χ matrix and since $\forall i$ $P_i P_i^\dagger = \mathbb{1}$, $\bar{\Lambda}$ is a unital channel. Channels that have a Kraus decomposition given by scaled Pauli operators are called Pauli channels.

The channel $\bar{\Lambda}$ still has a number of parameters that is exponential in the number of qubits comprising the system. Hence attempting to estimate these parameters is an inefficient process. However, if one considers an additional twirl over the group of permutations on n qubits, $\mathcal{S}_n$, then the number of parameters describing the resulting channel, denoted $\bar{\Lambda}_\Pi$, becomes polynomial. If an element of $\mathcal{S}_n$ is denoted S_j then since $|\mathcal{S}_n| = n!$ we have

$$\bar{\Lambda}(\rho) = \frac{1}{n!} \frac{1}{4^n} \sum_j \sum_i \sum_k S_j P_i A_k P_i S_j^\dagger \rho S_j P_i A_k^\dagger P_i S_j^\dagger. \tag{3.8}$$

Note that one need not actually implement the twirl over $\mathcal{S}_n$ as such a twirl is equivalent to neglecting any ordering of the qubits when making measurements on the system.

To see that the number of parameters describing $\bar{\Lambda}_\Pi$ becomes polynomial when twirling over $\mathcal{S}_n$, first define for an element P $\in \mathcal{P}_1^{\otimes n}$ the vector weight of P, denoted wt(P), by

$$wt(P) = (w_x, w_y, w_z). \tag{3.9}$$

w_x is the number of X operators in the decomposition of P as a tensor product of elements of $\mathcal{P}_1$, and similarly for w_y and w_z. Hence the number of identity operators in the decomposition of P is $n-(w_x+w_y+w_z)$. Any triple (w_x, w_y, w_z) satisfying w_x, w_y, $w_z \geq 0$ and $w_x + w_y + w_z \leq n$ will be denoted $\underline{w}$.

Suppose $\bar{\Lambda}$ is given and a twirl over random permutations is performed. All Kraus operators with the same $\underline{w}$ must end up with the same probability coefficient as the qubits are indistinguishable under the random permutations. This coefficient must be the average of the associated probability weights of these operators in $\bar{\Lambda}$. Hence the result of randomly permuting qubits is to create another Pauli channel where all Pauli operators with the same weight vector $\underline{w}$ are associated with a probability $p_{\underline{w}}$. These channels are called permutation invariant Pauli channels, as they are unchanged under permutations of the qubits. Hence the number of parameters describing $\bar{\Lambda}_\Pi$ is equal to the number of different possible $\underline{w}$.

If we show that the number of such $\underline{w}$ is polynomial in the number of qubits then we are done. Note that the number of possible $\underline{w}$ is equal to the number of solutions to the equation $w_x + w_y + w_z = w$ where $0 \leq w \leq n$. Clearly, the number of such solutions is

$$\sum_{w=0}^{n} \sum_{w_x=0}^{w} \sum_{w_y=0}^{w-w_x} 1. \tag{3.10}$$

From the equations

$$\sum_{w_y=0}^{w-w_x} 1 = w - w_x + 1, \tag{3.11}$$

$$\begin{aligned} \sum_{w_x=0}^{w} w - w_x + 1 &= w(w+1) - \frac{w(w+1)}{2} + w + 1 \\ &= \frac{w^2}{2} + \frac{3w}{2} + 1, \end{aligned} \tag{3.12}$$

we get,

$$\begin{aligned}\sum_{w=0}^{n}\sum_{w_x=0}^{w}\sum_{w_y=0}^{w-w_x} 1 &= \sum_{w=0}^{n} \frac{w^2}{2} + \frac{3w}{2} + 1 \\ &= \frac{1}{2}\left(\frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}\right) + \frac{3}{2}\frac{n(n+1)}{2} + n + 1 \\ &= \frac{n^3}{6} + n^2 + \frac{11n}{6} + 1. \qquad (3.13)\end{aligned}$$

Hence the number of parameters describing $\bar{\Lambda}_\Pi$ is polynomial in the number of qubits and so one can write

$$\bar{\Lambda}_\Pi(\rho) = \sum_{\underline{w}} \frac{p_{\underline{w}}}{N_{\underline{w}}} \left(\sum_{P_i \in \mathcal{P}_1^{\otimes n} : wt(P) = \underline{w}} P_i \rho P_i \right) \qquad (3.14)$$

where $N_{\underline{w}}$ is the number of Pauli's whose vector weight is $\underline{w} = (w_x, w_y, w_z)$. $N_{\underline{w}}$ can easily be shown to have the form

$$N_{\underline{w}} = \binom{n}{w_x + w_y + w_z} \binom{w_x + w_y + w_z}{w_x} \binom{w_y + w_z}{w_z}. \qquad (3.15)$$

The factor $\binom{n}{w_x+w_y+w_z}$ comes from the number of ways that $w_x + w_y + w_z$ qubits can be acted on, and the $\binom{w_x+w_y+w_z}{w_x}\binom{w_y+w_z}{w_z}$ factor comes from the different ways the single qubit operators act on a set of $w_x + w_y + w_z$ qubits.

One can rewrite the above expression for $\bar{\Lambda}_\Pi$ as

$$\bar{\Lambda}_\Pi(\rho) = p_{\underline{w}} M_{\underline{w}}^p(\rho) \qquad (3.16)$$

where $M_{\underline{w}}^p$ is the quantum channel given by

$$M_{\underline{w}}^p(\rho) = \frac{1}{N_{\underline{w}}} \sum_{P_i \in \mathcal{P}_1^{\otimes n} : wt(P) = \underline{w}} P_i \rho P_i. \qquad (3.17)$$

An interesting fact about Pauli channels is that eigenoperators of these channels are Pauli operators. To see this, note that for any two P_i, $P_j \in \mathcal{P}_1^{\otimes n}$ either $[P_i, P_j] = 0$ or $\{P_i, P_j\} = 0$. If

$$\mathcal{E}(\rho) = \sum_{i=1}^{D^2} pr(P_i)P_i\rho P_i. \tag{3.18}$$

is a Pauli channel then for $P_k \in \mathcal{P}_1^{\otimes n}$,

$$\begin{aligned}
\mathcal{E}(P_k) &= \sum_{i=1}^{D^2} pr(P_i)P_iP_kP_i \\
&= \sum_{i:[P_i,P_k]=0} pr(P_i)P_k - \sum_{i:\{P_i,P_k\}=0} pr(P_i)P_k \\
&= \left(\sum_{i:[P_i,P_k]=0} pr(P_i) - \sum_{i:\{P_i,P_k\}=0} pr(P_i) \right) P_k \\
&= C(P_k)P_k
\end{aligned} \tag{3.19}$$

where $C(P_k)$ is clearly in the interval $[-1,1]$. Thus, as $\mathcal{P}_1^{\otimes n}$ forms an orthogonal basis for $B(\mathcal{H})$, Pauli channels are Hermitian, or self-dual.

By definition, PIP channels have the property of being invariant under permutations. Hence all Pauli's with the same weight vector $\underline{w}$ have the same eigenvalue $\lambda_{\underline{w}}$ under the action of the channel. Therefore for a Pauli and permutation twirled channel $\bar{\Lambda}_\Pi$ we can write

$$\bar{\Lambda}_\Pi(\rho) = \sum_{\underline{w}} \lambda_{\underline{w}} M_{\underline{w}}^{\lambda}(\rho) \tag{3.20}$$

where the $M_{\underline{w}}^{\lambda}$ are quantum channels that are sums of projectors onto each Pauli operator with weight vector $\underline{w}$. In this form, composition of two PIP channels corresponds to multiplication of the eigenvalues $\lambda_{\underline{w}}$. It can be shown [13] that the $\lambda_{\underline{w}}$ can be estimated efficiently by experiments. More precisely, with $O\left(\frac{log(2(\frac{n^3}{6}+n^2+\frac{11n}{6}+1))}{\epsilon^2}\right)$ experiments one can estimate each of the eigenvalues to precision ϵ with constant probability.

There is a linear invertible mapping between the $\lambda_{\underline{w}}$ and $p_{\underline{w}}$, which is denoted by Ω,

$$\lambda_{\underline{w}} = \sum_{\underline{v}} \Omega_{\underline{w},\underline{v}} p_{\underline{v}}. \tag{3.21}$$

As a matrix, Ω is $\frac{n^3}{6}+n^2+\frac{11n}{6}+1$ by $\frac{n^3}{6}+n^2+\frac{11n}{6}+1$. The expression for the matrix entries of Ω and Ω^{-1} are computed to be,

$$\Omega_{\underline{w},\underline{v}} = \frac{tr\left(M_{\underline{w}}^{\lambda\dagger} M_{\underline{v}}^{p}\right)}{tr\left(M_{\underline{w}}^{\lambda\dagger} M_{\underline{w}}^{\lambda}\right)} \tag{3.22}$$

$$\Omega^{-1}_{\underline{w},\underline{v}} = \frac{tr\left(M_{\underline{w}}^{p\dagger} M_{\underline{v}}^{\lambda}\right)}{tr\left(M_{\underline{w}}^{p\dagger} M_{\underline{w}}^{p}\right)}. \tag{3.23}$$

In the next section we look at the relationship between correctable codes for the original and twirled channels.

3.2 Correctable Codes For Twirled Quantum Channels

From the previous section, twirling a channel over $\mathcal{P}_1^{\otimes n}$, $\mathcal{C}_1^{\otimes n}$ or $\mathcal{C}_n$ takes it to one described by a polynomial number of parameters. The twirling operation will be useful if it preserves, at least partially, properties of the original channel. Specifically, one would hope that correctable codes of the twirled channel resemble those of the original channel. As the following proposition shows, this is the case.

Theorem 3 *Let $\mathcal{E}$ be a quantum channel and $\bar{\mathcal{E}}$ be the image of $\mathcal{E}$ under a twirling operation with twirl operators given by a discrete subset of the unitary group $U(D)$. Then, a correctable triple for $\bar{\mathcal{E}}$ is a correctable triple for $\mathcal{E}$, up to a unitary correction.*

Proof: Let the Kraus operators for $\mathcal{E}$ be $\{E_a\}$ and let the twirl operators be $\{U_b\}$. Thus, a Kraus operator for $\bar{\mathcal{E}}$ is $U_bE_aU_b^\dagger$. Suppose that there exists a correctable triple $(\mathcal{R},\bar{\mathcal{E}},\mathcal{S})$ for $\bar{\mathcal{E}}$. Let $\{R_c\}$ be the Kraus operators for the recovery operation. Then from the section on correctable codes, there exists a projector P such that $\forall a,b,c$

$$PR_cU_bE_aU_b^\dagger P = R_cU_bE_aU_b^\dagger P \tag{3.24}$$

and

$$R_c U_b E_a U_b^\dagger|_P \in B(\mathcal{H}_A) \otimes \mathbb{1}_B \tag{3.25}$$

where $H_A \otimes H_B$ is such that B is the noiseless sector for $\mathcal{R} \circ \mathcal{E}$. Now, look at the channel $\mathcal{U}_i^\dagger \circ \bar{\mathcal{E}} \circ \mathcal{U}_i$ for some i in the index set of the twirling elements. WOLOG, let $i = 1$. By definition this channel is unitarily equivalent to $\bar{\mathcal{E}}$. The action of this channel is

$$\mathcal{U}_1^\dagger \circ \bar{\mathcal{E}} \circ \mathcal{U}_1(\rho) = \sum_b \sum_k U_1^\dagger U_b E_a U_b^\dagger U_1 \rho U_1^\dagger U_b E_a^\dagger U_b^\dagger U_1. \tag{3.26}$$

and so a Kraus operator for $\mathcal{U}_1^\dagger \circ \bar{\mathcal{E}} \circ \mathcal{U}_1$ is $U_1^\dagger U_b E_a U_b^\dagger U_1$. The projector $P' = U_1^\dagger P U_1$, recovery operation $\mathcal{R}'$ with Kraus operators $R'_c = U_1^\dagger R_c U_1$, and semigroup $\mathcal{S}' = U_1^\dagger \mathcal{S} U_1$ are such that

$$\begin{aligned}
P' R'_c U_1^\dagger U_b E_a U_b^\dagger U_1 P' &= U_1^\dagger P U_1 U_1^\dagger R_c U_1 U_1^\dagger U_b E_a U_b^\dagger U_1 U_1^\dagger P U_1 \\
&= U_1^\dagger P R_c U_b E_a U_b^\dagger P U_1 \\
&= U_1^\dagger R_c U_b E_a U_b^\dagger P U_1 \\
&= U_1^\dagger R_c U_1 U_1^\dagger U_b E_a U_b^\dagger U_1 U_1^\dagger P U_1 \\
&= R'_c U_1^\dagger U_b E_a U_b^\dagger U_1 P' \qquad (3.27)
\end{aligned}$$

and

$$R'_c U_1^\dagger U_b E_a U_b^\dagger U_1|_{P'} \in U_1^\dagger \left(B(\mathcal{H}_A) \otimes \mathbb{1}_B\right) U_1. \tag{3.28}$$

Therefore $(\mathcal{R}', \mathcal{U}_1^\dagger \circ \bar{\mathcal{E}} \circ \mathcal{U}_1, \mathcal{S}')$ is a correctable triple. If in the above equations we set $b = 1$ we get

$$P' R'_c E_a P' = R'_c E_a P' \tag{3.29}$$

and

$$R'_c E_a|_{P'} \in U_1^\dagger B(\mathcal{H}_A) \otimes \mathbb{1}_B U_1. \tag{3.30}$$

Thus, $(\mathcal{R}', \mathcal{E}, \mathcal{S}')$ is a correctable triple and so, up to a unitary correction, a correctable code for $\bar{\mathcal{E}}$ is correctable for $\mathcal{E}$.

As the composition of two twirls is itself a twirl, we have that the above proposition holds for a channel $\bar{\Lambda}_\Pi$ twirled over both $\mathcal{P}_1^{\otimes n}$ (or $\mathcal{C}_1^{\otimes n}$) and random qubit permuations. Moreover, as both $\mathcal{P}_1^{\otimes n}$ and $\mathcal{C}_1^{\otimes n}$ contain the identity, there is no unitary correction.

The converse is clearly not true for when the twirling elements are taken from the Clifford group, $\mathcal{C}_n$, the twirled channel is depolarizing. A more interesting counter-example is that of the collective rotation channel [20] twirled over $\mathcal{P}_1^{\otimes n}$. Noiseless susbsystems for the collective rotation channel are lost when this twirl is performed as the symmetry of the noise acting in the same direction on all qubits is broken. The following is a specific case of the above proposition, however the proof relies only on the commutation relations of the Pauli operators.

Proposition 10 *Let $\mathcal{E}$ be a quantum channel and $\bar{\mathcal{E}}$ be the $\mathcal{P}_1^{\otimes n}$ twirled channel for $\mathcal{E}$. Then a noiseless subsystem for $\bar{\mathcal{E}}$ is a noiseless subsystem for $\mathcal{E}$.*

Proof: First, note that $\bar{\mathcal{E}}$ is a unital quantum channel [9]. As well, for unital channels it is known that the commutant for the Kraus operators and fixed point sets coincide (see appendix 3). Thus, noiseless subsystems may be found from an analysis of the commutant. Moreover, any noiseless subsystem for a unital quantum channel must reside in the commutant of the channel [8]. Hence if we show that an element in the commutant of the Pauli twirled channel must be an element of the commutant of the original channel, we are done.

Indeed, let $\{E_a\}$ be the Kraus operators for $\mathcal{E}$. Then, a Kraus operator for $\bar{\mathcal{E}}$ is of the form $P_bE_aP_b$. Suppose A is in the commutant of $\bar{\mathcal{E}}$, that is, $AP_bE_aP_b = P_bE_aP_bA$ $\forall a, b$. Since $\bar{\mathcal{E}}$ is unital, A is in the fixed-point algebra of $\bar{\mathcal{E}}$, however this algebra is generated by a subset of the Pauli operators. So, WOLOG, let A be a Pauli operator and fix a, b. Then

$$\begin{aligned} AP_bE_aP_b &= P_bE_aP_bA \\ \Rightarrow P_bAP_bE_aP_b &= E_aP_bA \\ \Rightarrow P_bAP_bE_a &= E_aP_bAP_b \\ \Rightarrow AE_a &= E_aA \end{aligned} \tag{3.31}$$

where the last implication holds because of the commutation and anti-commutation relations of the Pauli operators. Thus, A is in the commutant of $\mathcal{E}$ and so the commutant of $\bar{\mathcal{E}}$ is a *sub-algebra of the commutant of $\mathcal{E}$. Therefore noiseless subsystems for $\bar{\mathcal{E}}$ are noiseless subsystems for $\mathcal{E}$.

Now that we have shown correctable codes for twirled channels are correctable for the original channel, the next question is how to find correctable codes for the twirled channel. Before discussing an algorithm for finding codes for the twirled channel some interesting properties of Pauli channels are given in the next section.

3.2.1 UCS Codes for Pauli Channels

The following is a useful result for finding UCS codes of Pauli channels.

Proposition 11 *Let Λ be a unital, diagonalizable quantum channel with eigenvalues λ_i and eigenoperators L_i. Then, the noise commutant of $\Lambda^\dagger \circ \Lambda$ is generated by the eigenoperators L_i with eigenvalues satisfying $|\lambda_i| = 1$.*

Proof: Λ is unital, thus so is $\Lambda^\dagger \circ \Lambda$. Moreover, Λ is diagonalisable, so $\Lambda^\dagger \circ \Lambda$ has eigenoperators L_i with eigenvalues $\lambda_i^* \lambda_i = |\lambda_i|^2$, where the λ_i are the eigenvalues of Λ. Since $\Lambda^\dagger \circ \Lambda$ is unital, its fixed point algebra and noise commutant coincide, and both are generated by the eigenoperators L_i with $|\lambda_i|^2 = 1$.

Pauli channels are unital channels and, since they are diagonalisable and Hermitian, these channels have a particularly simple fixed-point set structure. In particular, we immediately obtain the following result.

Corollary 1 *Let $\bar{\Lambda}$ be a Pauli channel. Then the noise commutant of $\bar{\Lambda}^\dagger \circ \bar{\Lambda}$ is the algebra generated by the eigenoperators with eigenvalues ± 1.*

Hence, from [26], we have that all UCS codes for a Pauli channel are found from the algebra generated by the Pauli operators with eigenvalues ± 1.

Another intereseting property of Pauli channels that follows from the above proposition is that every UCS is a UNS.

Corollary 2 *A UCS code of a Pauli channel $\mathcal{E}$ is also a UNS.*

Proof: Let the subsystem B of a semigroup $\mathcal{S}$ be a UCS for $\mathcal{E}$. By definition, B is a UNS for $\mathcal{E}$ if $\forall n \in \mathbb{N}$, B is a UCS for $\mathcal{E}^n$. Since $\mathcal{E}$ is unital, $\mathcal{E}^n$ is unital for every $n \in \mathbb{N}$. Hence, B is a UNS for $\mathcal{E}$ if and only if it is a noiseless subsystem of $(\mathcal{E}^n)^\dagger \circ \mathcal{E}^n$ for every $n \in \mathbb{N}$. As $\mathcal{E}$ is Hermitian, $\mathcal{E}^\dagger = \mathcal{E}$, and thus we have that $(\mathcal{E}^n)^\dagger \circ \mathcal{E}^n = (\mathcal{E}^\dagger \circ \mathcal{E})^n$. So we must show that B is a noiseless subsystem of $(\mathcal{E}^\dagger \circ \mathcal{E})^n$ for every $n \in \mathbb{N}$.

B is a UCS for $\mathcal{E}$ if and only if B is a noiseless subsystem of $\mathcal{E}^\dagger \circ \mathcal{E}$. Since the eigenvalues of $\mathcal{E}$ are real, there are no negative eigenvalues for $\mathcal{E}^\dagger \circ \mathcal{E} = \mathcal{E}^2$ and so the fixed point set of $\mathcal{E}^\dagger \circ \mathcal{E}$ is identical to the fixed point set of $(\mathcal{E}^\dagger \circ \mathcal{E})^n$. Thus, B is a noiseless subsystem of $(\mathcal{E}^\dagger \circ \mathcal{E})^n = (\mathcal{E}^n)^\dagger \circ \mathcal{E}^n$ for every $n \in \mathbb{N}$.

3.3 Algorithm For Finding UCS Codes of Twirled Quantum Channels

We now discuss an algorithm which uses the eigenvalues of $\bar{\Lambda}_\Pi$ to find UCS codes for $\bar{\Lambda}_\Pi$. From the previous section, UCS codes found in this manner are also UCS codes for Λ.

By [19, 26], there exists an algorithm for finding all UCS codes of a unital channel Λ if the linear structure of the commutant of $\Lambda^\dagger \circ \Lambda$ is known. However, this algorithm requires the manipulation of exponentially large matrices. Since the Pauli operators form an eigenbasis for PIP channels, the fixed point algebra of $\bar{\Lambda}_\Pi^\dagger \circ \bar{\Lambda}_\Pi$ is the algebra of polynomials generated by the eigenvalue ± 1 Pauli operators.

A simple way to find UCS codes is to partition the eigenvalue ± 1 Pauli operators into triplets which satisfy the single qubit Pauli commutation relations. These commutation relations can be computed without writing the observables explicitly in a particular representation. For finding noiseless subsystems, one would restrict their attention to the eigenvalue 1 Pauli operators. In this case, choosing the largest number K of mutually exclusive triplets which commute with each other, one implicitly describes how to encode a noiseless Hilbert space of dimension 2^K.

The encoding of the noiseless qubits is performed by a unitary operation that maps the first triplet to the set $\{X \otimes \mathbb{1}^{\otimes n-1}, Y \otimes \mathbb{1}^{\otimes n-1}, Z \otimes \mathbb{1}^{\otimes n-1}\}$, the second triplet to the set $\{\mathbb{1} \otimes X \otimes \mathbb{1}^{\otimes n-2}, \mathbb{1} \otimes Y \otimes \mathbb{1}^{\otimes n-2}, \mathbb{1} \otimes Z \otimes \mathbb{1}^{\otimes n-2}\}$, and so on up to the K'th triplet.

The unitary which performs the encoding of these n qubits is guaranteed to be in the Clifford group since it maps a set of Pauli operators to another set of Pauli operators with the same commutation relations. Standard techniques can be applied to determine which Clifford group operations implement a desired encoding [1].

This discussion leads to the following algorithm for finding UCS codes given the eigenvalues of the PIP channel $\bar{\Lambda}_\Pi$:

(i)Enumerate the Pauli operators with eigenvalues ±1 under the action of $\bar{\Lambda}$, Denote the set of these operators by F.

(ii) Choose a triplet of Pauli operators satisfying the single qubit Pauli commutation relations. If none can be found, the search is over.

(iii) If such a triplet can be found, remove it from F, as well as all operators that do not commute with the triplet, and go back to step (ii).

The number of mutually exclusive triplets found in this manner corresponds to an allowable number of UCS codes of dimension 2 that can be protected from the action of $\bar{\Lambda}_{\Pi}$. Finding noiseless subsystems is similarly simple. The only change to the above algorithm is in step (i) where one will enumerate all Pauli operators in the ±1 eigenspace. The question as to whether this algorithm can find all noiseless and unitarily correctable encodings for a PIP channels remains open.

3.3.1 Examples of the Algorithm

Some simple examples of how to use the algorithm are as follows:

Example 1: Consider the 2 qubit PIP channel with Kraus operators proportional to $\{\mathbb{1}\mathbb{1}, ZZ\}$. The Pauli operators with eigenvalue 1 are $\mathbb{1}\mathbb{1}, XX, YY, ZZ, XY, YX$, and $\mathbb{1}Z, Z\mathbb{1}$. Out of this set, $\{XX, XY, \mathbb{1}Z\}$ satisfy the commutation relations, and no other triplets which commute with these can be found. Hence, a single qubit can be encoded noiselessly through this channel.

Example 2: Consider the 2 qubit PIP channel with Kraus operators proportional to $\{\mathbb{1}\mathbb{1}, YX, XY\}$. The eigenoperators with eigenvalue 1 are $\mathbb{1}\mathbb{1}, XY, YX$, and ZZ. There are no triplets with the right commutation relations. The eigenoperators with eigenvalues -1 are $\mathbb{1}Z, Z\mathbb{1}, XX, YY$. If we consider the ±1 eigenspace, we obtain the same eigenoperators with eigenvalue 1 as the previous example, and thus there exists a UCS consisting of a single qubit.

In the case of the previous examples, we want to map the generating set of Pauli operators $\{\mathbb{1}X, \mathbb{1}Z\}$ to the generating set $\{XX, \mathbb{1}Z\}$, which can be done by a controlled-NOT gate. The second qubit is the control, and the first qubit is the target. The next section deals with robustness of the above protocol.

3.4 Robustness of the Method

There is always error in experimentally determining the values of λ_i. Suppose that the above algorithm is used to find a set of k qubits and the $3k$ Pauli operators in the k triplets from the algorithm are experimentally determined to have eigenvalues no smaller than $1-\epsilon$ for some $\epsilon \geq 0$. Let these Pauli operators be $K_{i,j}$ where the first index runs from 1 to 3 and determines which single qubit Pauli $K_{i,j}$ represents. If $i = 1$ then the operator represents Pauli X and so on. The second index runs from 1 to k and represents the j'th qubit. We want to use these 3k Pauli operators to transmit k qubits reliably through the twirled channel $\bar{\Lambda}_\Pi$. We first prove a result that puts a bound on the eigenvalue of any product of two operators $K_{i,j}$, $K_{l,m}$ for $\bar{\Lambda}_\Pi$.

Theorem 4 *Suppose $K_{i,j}$ and $K_{l,m}$ are two Pauli operators whose eigenvalues, $\lambda(K_{i,j})$, $\lambda(K_{l,m})$ under $\bar{\Lambda}_\Pi$ are bounded below by $1-\epsilon$. Then the eigenvalue of $K_{i,j}K_{l,m}$ is bounded below by $1-3\epsilon$.*

Proof: Define $Q = K_{i,j}$ and $R = K_{l,m}$. We know that

$$\begin{aligned} \lambda(Q) &= pr(Q) - pr(\overline{Q}) \\ &\geq 1-\epsilon \end{aligned} \tag{3.32}$$

where pr(Q) is the probability that a randomly chosen Pauli operator commutes with Q and $pr(\overline{Q})$ is the probability that a randomly chosen Pauli operator anti-commutes with Q. Hence since $pr(Q) + pr(\overline{Q}) = 1$ we get,

$$pr(Q) \geq 1 - \frac{\epsilon}{2}. \tag{3.33}$$

This implies that $pr(\overline{Q}) \leq \frac{\epsilon}{2}$. Analogous equations hold for R. For the Pauli operator QR we have,

$$\lambda(QR) = pr(QR) - pr(\overline{QR}). \tag{3.34}$$

Note that

$$pr(QR) = pr(Q \wedge R) + pr(\overline{Q} \wedge \overline{R}) \tag{3.35}$$

where $pr(Q \wedge R)$ is the probability that a randomly chosen Pauli operator commutes with both Q and R, and $pr(\bar{Q} \wedge \overline{R})$ is the probability that a randomly chosen Pauli operator anti-commutes with both Q and R. As well,

$$pr(\overline{QR}) = pr(\overline{Q} \wedge R) + pr(Q \wedge \overline{R}) \tag{3.36}$$

where $pr(\overline{Q} \wedge R)$ is the probability that a randomly chosen Pauli operator anticommutes with Q and commutes with R, and similarly for $pr(Q \wedge \overline{R})$.

Now we have the equations,

$$pr(Q \wedge R) = pr(Q) + pr(R) - pr(Q \vee R) \tag{3.37}$$

and

$$pr(\overline{Q} \wedge R) = pr(\overline{Q}) + pr(R) - pr(\overline{Q} \vee R) \tag{3.38}$$

where $pr(Q \vee R)$ is the probability that a randomly chosen Pauli operator commutes with Q or commutes with R, and similarly for $pr(\overline{Q} \vee R)$. Hence since $pr(Q \vee R) \leq 1$ we get

$$\begin{aligned} pr(Q \wedge R) &\geq 1 - \frac{\epsilon}{2} + 1 - \frac{\epsilon}{2} - 1 \\ &= 1 - \epsilon. \end{aligned} \tag{3.39}$$

Similarly, since $pr(\overline{Q} \vee R) \geq 1 - \frac{\epsilon}{2}$,

$$\begin{aligned} pr(\overline{Q} \wedge R) &\leq \frac{\epsilon}{2} + 1 - (1 - \frac{\epsilon}{2}) \\ &= \epsilon. \end{aligned} \tag{3.40}$$

By symmetry, $pr(Q \wedge \overline{R}) \leq \epsilon$. Finally, with $pr(\overline{Q} \wedge \overline{R}) \geq 0$ we get,

$$\begin{aligned} \lambda(QR) &\geq 1 - \epsilon + 0 - \epsilon - \epsilon \\ &\geq 1 - 3\epsilon \end{aligned} \tag{3.41}$$

which proves the theorem.

We immediately obtain the following corollary.

Corollary 3 *If the $K_{i,j}$ have eigenvalue bounded below by $1-\epsilon$ then any element in the algebra generated by these operators has eigenvalue bounded below by $1 - 3k^2\epsilon$ where k is the number of qubits.*

Proof: By linearity, we need only show that for a collection of Pauli operators $\{K_{i_j,j}\}_{j=1}^{k}$ *with eigenvalues bounded below by* $1-\epsilon$*, the product* $\Pi_{j=1}^{k}K_{i_j,j}$ *has eigenvalue bounded below by* $1-3k^2\epsilon$*. Here* $K_{i_j,j}$ *is essentially a single qubit Pauli operator acting on the j'th qubit. When k is even,*

$$\begin{aligned}\Pi_{j=1}^{k}K_{i_j,j} &= K_{i_1,1}K_{i_2,2}...K_{i_{k-1},k-1}K_{i_k,k} \\ &= \left(K_{i_1,1}K_{i_2,2}\right)...\left(K_{i_{k-1},k-1}K_{i_k,k}\right) \qquad (3.42)\end{aligned}$$

where from the above theorem, for each $j\in\{0,...,k-1\}$*, the operators* $\{R_{i_j,j}\}_{j=1}^{\frac{k}{2}} = K_{i_j,j}K_{i_{j+1},j+1}$ *have eigenvalue bounded below by* $1-3\epsilon$*. Iterating this process with these new operators until all of the operators have been multiplied together results in a number of steps equal to* log_2k *and we get the lower bound for* $\lambda\left(\Pi_{j=1}^{k}K_{i_j,j}\right)$ *of* $1-3^{log_2k}\epsilon$*.*

If k is odd then we get the lower bound of $1-3^{log_2(k-1)+1}\epsilon$*. In either case,*

$$\begin{aligned}\lambda\left(\Pi_{j=1}^{k}K_{i_j,j}\right) &\geq 1-3^{log_2k+1}\epsilon \\ &\geq 1-3(4^{log_2k})\epsilon \\ &= 1-3k^2\epsilon \qquad (3.43)\end{aligned}$$

which proves the corollary.

Using these results we can show that the protocol is robust. Suppose the $K_{i,j}$ are such that the square of their eigenvalues, $\lambda_{i,j}^{K}$, under $\bar{\Lambda}_\Pi$ are experimentally determined to be bounded below by $1-\epsilon$. Hence, from above the eigenvalue of any product of these operators for $\bar{\Lambda}_\Pi^\dagger\circ\bar{\Lambda}_\Pi$ is bounded below by $1-3k^2\epsilon$. When ϵ is 0 these operators form a k qubit noiseless subsystem for $\bar{\Lambda}_\Pi^\dagger\circ\bar{\Lambda}_\Pi$ and hence a k qubit UCS for $\bar{\Lambda}_\Pi$.

Suppose we want to send a k qubit state through $\bar{\Lambda}_\Pi$ using this code. We would like to determine a lower bound on the fidelity between the input and output states. Let us assume the input state is pure. As the physical state space is comprised of n qubits, the input state will be adjoined to an ancilla state in the remaining n-k qubits. Suppose that the initial state is $\rho^0\otimes\rho^{anc}$ where ρ^0 is the k qubit state to be transmitted and ρ^{anc} is an n-k qubit ancilla state. Let

$$\rho^0 = \sum_{i=0}^{2^{2k}-1} \rho_i^0 P_i \tag{3.44}$$

be a pure state where $P_i \in \mathcal{P}_1^{\otimes k}$ and

$$\rho^{anc} = \sum_{j=0}^{2^{2(n-k)}-1} \rho_j^{anc} P_j \tag{3.45}$$

with $P_j \in \mathcal{P}_1^{\otimes(n-k)}$. Hence,

$$\begin{aligned} \rho^0 \otimes \rho^{anc} &= \sum_{i=0}^{2^{2k}-1} \rho_i^0 P_i \otimes \sum_{j=0}^{2^{2(n-k)}-1} \rho_j^{anc} P_j \\ &= \frac{1}{2^{n-k}} \sum_{i=0}^{2^{2k}-1} \rho_i^0 P_i \otimes \mathbb{1}_{2^{n-k}} + \sum_{i=0}^{2^{2k}-1} \sum_{j=1}^{2^{2(n-k)}-1} \rho_i^0 \rho_j^{anc} P_i \otimes P_j \end{aligned} \tag{3.46}$$

From above there is a Clifford operation that maps $P_i \otimes \mathbb{1}_{2^{n-1}}$ to $K_{i,1}$ and so on for each of the k qubits. Call this Clifford element the encoding operation and label it as ENC. Then,

$$ENC\left(\rho^0 \otimes \rho^{anc}\right) = \frac{1}{2^{n-k}} \sum_{i=0}^{2^{2k}-1} \rho_i^0 M_i + \sum_{i=0}^{2^{2k}-1} \sum_{j=1}^{2^{2(n-k)}-1} \rho_i^0 \rho_j^{anc} R_{i,j} \tag{3.47}$$

where the first sum represents the encoded state, with $\{M_i\}_{i=1}^{2^{2k}-1}$ being all products of the $K_{i,j}$. The second sum is over Pauli operators $R_{i,j}$ such that no $R_{i,j}$ is the identity. We send this state through $\bar{\Lambda}_\Pi$ to obtain

$$\bar{\Lambda}_\Pi\left(ENC\left(\rho^0 \otimes \rho^{anc}\right)\right) = \frac{1}{2^{n-k}} \sum_{i=0}^{2^{2k}-1} \lambda_i^M \rho_i^0 M_i + \sum_{i=0}^{2^{2k}-1} \sum_{j=1}^{2^{2(n-k)}-1} \lambda_{i,j}^R \rho_i^0 \rho_j^{anc} R_{i,j}.$$

Applying the decoding procedure and tracing out the n-k ancilla qubits eliminates the second sum because it contains only non-identity Pauli operators. Hence the output state ρ^F is

$$\rho^F = \sum_{i=0}^{2^{2k}-1} \lambda_i \rho_i^0 P_i \tag{3.48}$$

Since ρ^0 was assumed to be a pure state, the fidelity of ρ^0 and ρ^F is given by $tr(\rho^0 \rho^F)$. However,

$$\begin{aligned} tr(\rho^0\rho^F) &= tr\left(\sum_{i=0}^{2^{2k}-1} \rho_i^0 P_i \sum_{j=0}^{2^{2k}-1} \lambda_j \rho_j^0 P_j\right) \\ &= 2^k \sum_{i=0}^{2^{2k}-1} \left(\rho_i^0\right)^2 \lambda_i \\ &\geq \min_i \lambda_i \left[2^k \sum_{i=0}^{2^{2k}-1} \left(\rho_i^0\right)^2\right] \\ &= \min_i \lambda_i \left[tr(\left(\rho^0\right)^2)\right] \\ &\geq 1 - 3k^2\epsilon \end{aligned} \tag{3.49}$$

where we have used the fact that $tr((\rho^0)^2) = 1$ since the trace of the square of a pure state is 1. By concavity of fidelity, this bound holds for any input state and so the procedure is robust. The next section shows that other useful information about a quantum process can be obtained by twirling. Specifically, one can gain information regarding the Markovicity of the channel.

3.5 A Composition Law for Twirled Markovian Quantum Channels

The quantum operation formalism presented to this point describes the evolution of a quantum system from one specific time instance to another. A more general method of description is to model the evolution of the system in a continuous manner.

Definition 10 *One Parameter Family of Quantum Channels*

Let $\Lambda(t)$ denote a mapping from the non-negative real numbers to the space of quantum channels acting on $B(\mathcal{H})$. Such a mapping is called a one parameter family of quantum channels.

A one parameter family that does not have any memory effects is called Markovian.

Definition 11 *Markovian Quantum Process*

$\Lambda(t)$ is a Markovian quantum process [5] if

1. $\Lambda(0) = \mathbb{1}$.

2. For any $t_1, t_2 \geq 0$, $\Lambda(t_2 + t_1) = \Lambda(t_2) \circ \Lambda(t_1)$.

For t_1, $t_2 \geq 0$, let $\Lambda_{t_1 \to t_2}$ denote the quantum channel given by $\Lambda(t)$ acting from t_1 to t_2.

Theorem 5 *Let $\Lambda(t)$ be a Markovian quantum process and $\bar{\Lambda}_{0\to\delta t}$, $\bar{\Lambda}_{\delta t\to 2\delta t}$ be $\mathcal{P}_1^{\otimes n}$ $(\mathcal{C}_1^{\otimes n})$ twirls of $\Lambda(\delta t) = \Lambda_{0\to\delta t}$ and $\Lambda_{\delta t\to 2\delta t}$ respectively. Then the eigenvalues $\tilde{c_w}$ associated with the channel $\bar{\Lambda}_{\delta t\to 2\delta t} \circ \bar{\Lambda}_{0\to\delta t}$ are the squares of the eigenvalues c_w associated with $\bar{\Lambda}_{0\to\delta t}$.*

Proof: Since $\Lambda(t)$ is Markovian,

$$\Lambda_{0\to 2\delta t} = \Lambda_{0\to\delta t} \circ \Lambda_{0\to\delta t}. \tag{3.50}$$

Moreover by definition,

$$\Lambda_{0\to 2\delta t} = \Lambda_{\delta t\to 2\delta t} \circ \Lambda_{0\to\delta t}. \tag{3.51}$$

This implies that

$$\bar{\Lambda}_{\delta t\to 2\delta t} \circ \bar{\Lambda}_{0\to\delta t} = \bar{\Lambda}_{0\to\delta t} \circ \bar{\Lambda}_{0\to\delta t}. \tag{3.52}$$

$\bar{\Lambda}_{0\to\delta t}$ is diagonalizable with orthogonal eigenbasis given by $\mathcal{P}_1^{\otimes n}$ and the c_w are the eigenvalues of $\bar{\Lambda}_{0\to\delta t}$ implies that c_w^2 are the eigenvalues of $\bar{\Lambda}_{0\to\delta t} \circ \bar{\Lambda}_{0\to\delta t}$. Hence since $\tilde{c_w}$ are the eigenvalues of $\bar{\Lambda}_{\delta t\to 2\delta t} \circ \bar{\Lambda}_{0\to\delta t}$, the $\tilde{c_w}$ are the squares of the c_w.

The above theorem shows that twirling a quantum process can lead to extraction of useful information regarding Markovicity of the original channel. Indeed, if the eigenvalues of $\bar{\Lambda}_{\delta t\to 2\delta t} \circ \bar{\Lambda}_{0\to\delta t}$ are not equal to the squares of the eigenvalues of $\bar{\Lambda}_{0\to\delta t}$, then $\Lambda(t)$ can not have been a Markovian process.

3.5.1 Non-Markovian Channels Don't Obey A Composition Law

A natural question to ask is whether non-Markovian processes obey the above constraint on the eigenvalues when the channel is twirled. It is clear that, in general, a non-Markovian process will not satisfy this condition. A counterexample is the non-Markovian process on a system S given by the circuit below.

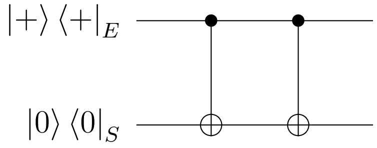

Figure 3.1: C-NOT Gates in Series as a Non-Markovian Process

This operation is clearly equal to the identity. The system of interest, S, is a single qubit that is coupled to an environment E, also comprised of a single qubit. The initial state $\rho_{ES} = |+\rangle\langle+|_E \otimes |0\rangle\langle0|_S$ is a product state, and hence by Stinespring's dilation theorem the entire dynamics for S is described by a CP map. However, one can see that the input to the second C-NOT gate is the maximally entangled state given by

$$\begin{aligned} CNOT(\rho_{ES}) = \frac{1}{2}(&|0\rangle\langle0| \otimes |0\rangle\langle0| + |0\rangle\langle1| \otimes |0\rangle\langle1| \\ &+|1\rangle\langle0| \otimes |1\rangle\langle0| + |1\rangle\langle1| \otimes |1\rangle\langle1|) \end{aligned} \tag{3.53}$$

$$\mapsto \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}. \tag{3.54}$$

Hence the dynamics for S after the initial CNOT gate is non-CP, and so the overall dynamics for S is non-Markovian.

Let us assume that each of the CNOT gates is executed in time δt. We can test the eigenvalue condition by Clifford twirling S on each gate, as shown in the circuit diagram below.

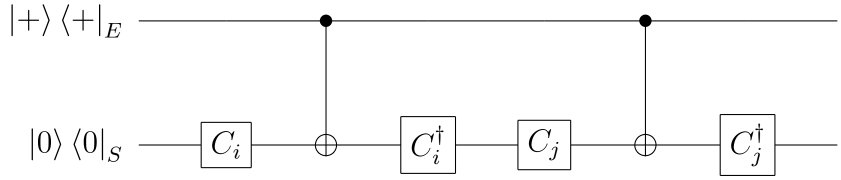

Figure 3.2: Clifford Twirling a Non-Markovian Process

Recall that the Clifford group on n qubits is a unitary 2-design, that is, for all quantum channels Λ and states ρ,

$$\frac{1}{|\mathcal{C}_1^{\otimes n}|}\sum_{U_k\in\mathcal{C}_1^{\otimes n}} U_k^\dagger\Lambda(U_k\rho U_k^\dagger)U_k = \int_{U(D)} U^\dagger\Lambda(U\rho U^\dagger)U \tag{3.55}$$

where the integral is with respect to the Haar measure on $U(D)$. In [16], a condition is given for a finite set of unitary operators to constitute a unitary 2-design. If $G = \{U_1, ..., U_K\}$ is a set of unitary operators, then G constitutes a 2-design if and only if

$$\sum_{U_k,U_{k'}\in G}\frac{|tr(U_k^\dagger U_k)|^4}{K^2} = 2 \tag{3.56}$$

Thus, one can use any set of operators that constitute a 2-design to perform a Clifford twirl, since the action of either will be the same. The following set of unitary matrices constitutes a 2-design on a single qubit, as can easily be verified.

$$\begin{aligned}
U_1 &= \mathbb{1} \\
U_2 &= e^{\frac{i\pi X}{2}} \\
U_3 &= e^{\frac{i\pi Y}{2}} \\
U_4 &= e^{\frac{i\pi Z}{2}} \\
U_5 &= e^{\frac{i\pi X}{4}} e^{\frac{i\pi Y}{4}} \\
U_6 &= e^{\frac{-i\pi X}{4}} e^{\frac{i\pi Y}{4}} \\
U_7 &= e^{\frac{i\pi X}{4}} e^{\frac{-i\pi Y}{4}} \\
U_8 &= e^{\frac{-i\pi X}{4}} e^{\frac{-i\pi Y}{4}} \\
U_9 &= e^{\frac{i\pi Y}{4}} e^{\frac{i\pi X}{4}} \\
U_{10} &= e^{\frac{-i\pi Y}{4}} e^{\frac{i\pi X}{4}} \\
U_{11} &= e^{\frac{i\pi Y}{4}} e^{\frac{-i\pi X}{4}} \\
U_{12} &= e^{\frac{-i\pi Y}{4}} e^{\frac{-i\pi X}{4}}
\end{aligned} \tag{3.57}$$

The eigenvalue condition is tested by calculating the c_w values directly after the first Clifford twirl, and the $\tilde{c_w}$ values after the second Clifford twirl. Since S is a single qubit, there are two possibilities, 0 and 1, for w. As the CNOT gates can not be decomposed into a gate acting on S alone, one can not explicitly calculate the c_w or $(\tilde{c_w})$ values as eigenvalues of some twirled channel. Instead, we must revert to the original definition of these quantities as described in [13]. We have for c_w (and similarly for $\tilde{c_w}$),

$$c_w = 2q_w - 1 \tag{3.58}$$

where q_w is the probability that, upon measurement in the computational basis, a random subset of w qubits has even parity. Hence, q_0 and c_0 are always 1.

We calculate the density operator of the composite system both directly after the first twirl, and after the second twirl. Labeling these as ρ and $\tilde{\rho}$ respectively, we have,

$$\rho = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{6} & 0 \\ 0 & 0 & 0 & \frac{1}{3} \end{bmatrix}, \tag{3.59}$$

$$\tilde{\rho} = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{5}{18} & 0 \\ 0 & 0 & 0 & \frac{2}{9} \end{bmatrix}. \tag{3.60}$$

This gives reduced density operators ρ_s and $\tilde{\rho_s}$

$$\rho_s = \begin{bmatrix} \frac{2}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}, \tag{3.61}$$

$$\tilde{\rho_s} = \begin{bmatrix} \frac{7}{9} & 0 \\ 0 & \frac{2}{9} \end{bmatrix}. \tag{3.62}$$

Clearly, q_1 is just the probability of obtaining 0 from a measurement in the computational basis, and so $q_1 = \frac{2}{3}$. Similarly, $\tilde{q_1} = \frac{7}{9}$. Thus, $c_1 = \frac{1}{3}$ and $\tilde{c_1} = \frac{5}{9}$, which shows that $\tilde{c_1} \neq c_1^2$. Hence the eigenvalue condition is not satisfied for this non-Markovian process on S.

3.6 Conclusion

We have discussed twirling a quantum channel Λ over a subset of U(D) and in particular over the Pauli and permutation groups. This specific twirl results in a channel $\bar{\Lambda}_\Pi$ that is described by a number of parameters that is polynomial in the number of qubits comprising the system. These parameters can be estimated efficiently in an experimental setting and so the twirled channel can be estimated without resorting to quantum process tomography.

We have shown that useful partial information for Λ can be gained via twirling. Correctable codes for the channel resulting from twirling Λ over any discrete subset of U(D) are also correctable for Λ up to some unitary operation. We proved various results regarding UCS codes for Pauli channels that naturally lead to an algorithm for finding these types of codes. The algorithm is exponential however it does not involve operations on exponentially large matrices. Some simple examples of the algorithm were given and the protocol was shown to be robust against experimental error.

When Λ is Markovian and the twirl is over the Pauli and permutation groups, the twirled channel obeys a specific composition law. Hence if this law is not satisfied, the original channel can't be Markovian. It was shown that a general non-Markovian channel will not satisfy such a composition law.

3.7 Further Work

There are some open questions remaining from this work. While we have shown that correctable codes for Λ are correctable for $\bar{\Lambda}_{\Pi}$ up to a unitary correction, it is not clear how to quantify the loss of possible codes from a twirling procedure, specifically when twirling over the Pauli group and random permutations. Finding the amount of codes that are lost will give an idea as to how useful the protocol presented here is. The algorithm presented for finding UCS codes for the twirled channel still has a large overhead. It would be useful to find a way of using the algebraic relations of the Pauli operators to obtain a polynomial time algorithm for finding a UCS code for $\bar{\Lambda}_{\Pi}$. In addition, finding the largest possible UCS code for $\bar{\Lambda}_{\Pi}$ via this algorithm would be of interest.

A more general question is the existence of unitary t-designs for $t \geq 2$. As we have shown, the Clifford group forms a unitary 2-design and twirling a channel over the Clifford group gives direct information regarding the average fidelity of the channel. It can be seen that if one wants information about the variance of the fidelity as a distribution, a unitary 4-design is required. Similarly, information about higher order moments of the fidelity distribution require higher order unitary t-designs.

Appendices

Appendix A

Group Theory, The Quantum Fourier Transform and Applications to Quantum Algorithms

A.1 The Cyclic Quantum Fourier Transform

Before defining the quantum Fourier transform, we recall the classical discrete Fourier transform (CDFT).

Definition 12 *Classical Discrete Fourier Transform*

The CDFT is a mapping $\mathcal{F}_{\mathcal{C}} : \mathbb{C}^N \rightarrow \mathbb{C}^N$ defined by $(x_0, ..., x_{N-1}) \rightarrow (y_0, ..., y_{N-1})$ where

$$y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{\frac{2\pi ijk}{N}}. \tag{A.1}$$

The cyclic QFT is defined similarly in the following manner.

Definition 13 *Cyclic Quantum Fourier Transform*

Let the Hilbert space $\mathcal{H}$ have dimension N and suppose $\mathcal{H}$ has orthonormal basis $\mathcal{B} = \{|0\rangle, ..., |N-1\rangle\}$. The cyclic QFT is a mapping $\mathcal{F}_{\mathcal{N}} : \mathcal{H} \rightarrow \mathcal{H}$ defined by $\{|0\rangle, ..., |N-1\rangle\} \rightarrow \{|y_0\rangle, ..., |y_{N-1}\rangle\}$ where

$$|y_k\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} e^{\frac{2\pi ijk}{N}} |j\rangle. \tag{A.2}$$

If one writes out a general state in $\mathcal{B}$ as

$$|\psi\rangle = x_0|0\rangle + ... + x_{N_1}|N-1\rangle \to (x_0, ..., x_{N-1}) \tag{A.3}$$

then

$$\begin{aligned}
\mathcal{F_N}|\psi\rangle &= x_0\mathcal{F_N}|0\rangle + ... + x_{N_1}\mathcal{F_N}|N-1\rangle \\
&= x_0\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1} e^{\frac{2\pi ij0}{N}}|j\rangle + ... + x_{N-1}\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1} e^{\frac{2\pi ij(N-1)}{N}}|j\rangle \\
&= \frac{1}{\sqrt{N}}\left(x_0 e^{\frac{2\pi i00}{N}} + ... + x_{N-1}e^{\frac{2\pi i0(N-1)}{N}}\right)|0\rangle + ... \\
&\quad + \frac{1}{\sqrt{N}}\left(x_0 e^{\frac{2\pi i(N-1)0}{N}} + ... + x_{N-1}e^{\frac{2\pi i(N-1)(N-1)}{N}}\right)|N-1\rangle \\
&= y_0|0\rangle + ... + y_{N-1}|N-1\rangle \to (y_0, ..., y_{N-1}) \qquad (A.4)
\end{aligned}$$

where the y_i are as defined in the CDFT. Hence the cyclic QFT is just the CDFT when one represents a state by its coordinates in the computational basis. It is easy to verify that the cyclic QFT is a unitary transformation.

Clearly, from the definition of the cyclic QFT, the representation of the cyclic QFT in the basis $\mathcal{B}$ is

$$\mathcal{F_N} \mapsto \frac{1}{\sqrt{N}} \begin{bmatrix}
1 & 1 & 1. & . \; . \; . & 1 \\
1 & \omega_N & \omega_N^2 & . \; . \; . & \omega_N^{N-1} \\
1 & \omega_N^2 & \omega_N^4 & . \; . \; . & \omega_N^{2(N-1)} \\
. & . & . & . \; . \; . & . \\
. & . & . & . \; . \; . & . \\
. & . & . & . \; . \; . & . \\
1 & \omega_N^{N-1} & \omega_N^{2(N-1)} & . \; . \; . & \omega_N^{(N-1)^2}
\end{bmatrix}$$

where ω_N is the primitive N-th root of unity $e^{\frac{2\pi i}{N}}$. This can also be written using the outer product formalism

$$\mathcal{F} = \sum_{g,h \in \mathbb{Z}_N} \omega_N^{gh} |g\rangle\langle h| \tag{A.5}$$

As we will see, the cyclic QFT is a specific case of a more general definition of a quantum Fourier transform over finite abelian groups. As suggested by its name, the cyclic QFT results from the underlying group being cyclic.

Next, we give a simple formula for the output of the cyclic QFT for an n-fold tensor product of $\mathbb{C}^2$, ie. for an n-qubit Hilbert space. This will give insight into what a quantum circuit decompostion for the cyclic QFT will look like. Let $N = 2^n$ where $n \in \mathbb{P}$. Write the basis $\{|0\rangle, ..., |2^n - 1\rangle\}$ in binary representation, ie. $|j\rangle \rightarrow |j_1 j_2 ... j_n\rangle$ where $j = j_1 2^{n-1} + ... + j_n 2^0$. Then by definition,

$$\begin{aligned}
|k\rangle \quad &\rightarrow \quad \frac{1}{2^{\frac{n}{2}}} \sum_{j_1=0}^{1} ... \sum_{j_n=0}^{1} e^{2\pi i k\left(\frac{j_1 2^{n-1} + ... + j_n 2^0}{2^n}\right)} |j_1 ... j_n\rangle \\
&= \quad \frac{1}{2^{\frac{n}{2}}} \left(\sum_{j_1=0}^{1} e^{2\pi i k j_1 2^{-1}} |j_1\rangle \right) \otimes ... \otimes \left(\sum_{j_n=0}^{1} e^{2\pi i k j_n 2^{-n}} |j_n\rangle \right) \\
&= \quad \frac{1}{2^{\frac{n}{2}}} \left(|0\rangle + e^{2\pi i \left(k_1 2^{n-1} + ... + k_n 2^0\right) 2^{-1}} |1\rangle \right) \otimes ... \\
&\quad \otimes \left(|0\rangle + e^{2\pi i \left(k_1 2^{n-1} + ... + k_n 2^0\right) 2^{-n}} |1\rangle \right) \\
&= \quad \frac{1}{2^{\frac{n}{2}}} \left(|0\rangle + e^{2\pi i 0.k_n} |1\rangle \right) \otimes ... \otimes \left(|0\rangle + e^{2\pi i 0.k_1 ... k_n} |1\rangle \right)
\end{aligned} \tag{A.6}$$

where we have used $e^{2\pi i k_m 2^{n-r}} = 1$ for $k_m \in \{0, 1\}$ and $r \leq n$, and the binary fraction representation

$$0.k_l k_{l+1} ... k_d = \frac{k_l}{2} + \frac{k_{l+1}}{2^2} + ... + \frac{k_d}{2^{d-l+1}} \tag{A.7}$$

As noted, this formula allows one to construct a simple quantum circuit implementing the cyclic QFT over n qubits, which is given below. The circuit gives the above state (in reverse order) for input $|k\rangle = |k_1 ... k_n\rangle$. To get the correct order, one needs to use $\frac{n}{2}$ swap gates. Hence, by inspection, the circuit uses $O(n^2) = O(log^2(N))$ elementary gates or operations. Therefore the circuit is polynomial in the number of qubits n and poly-logarithmic in N.

It can be shown that there exist circuit decompositions that approximate the cyclic QFT for any $N \in \mathbb{N}$ which uses a number of gates poly-logarithmic in N

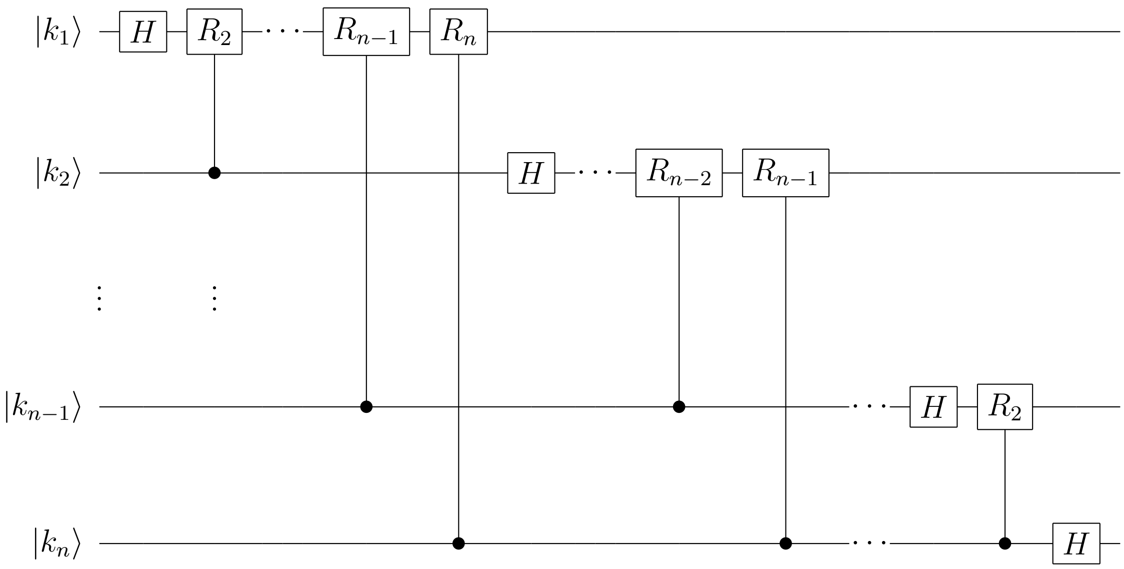

Figure A.1: Circuit Decomposition for Cyclic QFT

[30]. The circuit decompositions for arbitrary $N \in \mathbb{N}$ are more complicated than for $N = 2^n$.

A.2 Group Representations and Characters

Definition 14 *Group*

A group is a set G together with a binary operation denoted by $+$ in additive notation (or $\cdot$ in multiplicative notation) that satisfies the following properties:

1. $\forall\, a, b, c \in G,\ a + (b + c) = (a + b) + c$ (Associativity)

2. $\exists 0 \in G\, \forall a \in G,\ a + 0 = 0 + a$ (Identity)

3. $\forall a \in G\, \exists b \in G,\ a + b = 0 = b + a$ (Inverses)

The element b in 3 above is called the inverse of a, and is denoted by $-a$ in additive notation. G will be used to denote both the set on which the group operation is defined and the group itself.

When the binary operation is also commutative G is called abelian. An example of a finite abelian group is given by the integers modulo n. A useful group in mathematics is that of the invertible n by n matrices over a field F. This group is called the general linear group of dimension n and denoted $GL(n, F)$.

Definition 15 *Homomorphism*

A homomorphism from a group G into a group H is a function f that satisfies

$$f(a+b) = f(a) + f(b) \tag{A.8}$$

$\forall\, a, b \in G$*. If in addition f is one to one and onto, f is called an isomorphism.*

Definition 16 *Group Representation*

A representation of a group G on $\mathbb{C}^n$ is a group homomorphism from G into $GL(n, \mathbb{C})$.

A representation of a group G on $\mathbb{C}^n$ will just be called a representation. Unless otherwise stated, the group operation on the generic group G will be taken to be additive.

Definition 17 *Group Character*

Let G be a group and $\sigma : G \to GL(n, \mathbb{C})$ be a representation of G. If n=1 we call σ a character of G. The character $\chi : G \to \mathbb{C}/\{0\}$ defined by $\chi(g) = tr(\sigma(g))$ is called the character of G afforded by σ.

The set of characters of G will be denoted char(G).

Proposition 12 *If χ is a character of a finite group G, then each function value $\chi(g)$ is a root of unity.*

Proof: Let $g \in G$. Since G is finite, $|g|$ is finite and so let $|g| = n$. If e denotes the identity in G, then since χ is a homomorphism,

$$1 = \chi\,(e) = \chi\,(ng) = \chi\,(g)^n \tag{A.9}$$

Hence $\chi(g)$ is a root of unity

Since char(G) is just the set of homomorphisms from G into $\mathbb{C}/\{0\}$, char(G) is an abelian group, called the character group, under the multiplicative operation defined by $fg(a) = f(a)g(a)$. The identity of the group is called the principal character and maps every element of G to 1. The rest of the discussion will involve only finite abelian groups.

The following is a structure theorem for finite abelian groups. A proof can be found in [38].

Theorem 6 *Let G be a finite abelian group such that $|G| > 1$. Let the unique factorization of $|G|$ into primes be*

$$|G| = p_1^{\alpha_1} ... p_n^{\alpha_n} \tag{A.10}$$

where $p_1 > ... > p_n$. Then,

1. *$G \cong A_1 \oplus ... \oplus A_n$ where $|A_i| = p_i^{\alpha_i}$*
2. *$\forall A_i \in \{A_1, ..., A_n\}$*

$$A_i \cong \mathbb{Z}_{p_i^{\beta_1}} \oplus ... \oplus \mathbb{Z}_{p_i^{\beta_t}} \tag{A.11}$$

where $\beta_1 \geq ... \geq \beta_t \geq 1$, $\beta_1 + ... + \beta_t = \alpha_i$, and t and each β_j depend on i.

By the above structure theorem, we will WOLOG suppose throughout the rest of the presentation that the generic group G has a decomposition of the form

$$G \cong \mathbb{Z}_{N_1} \oplus ... \oplus \mathbb{Z}_{N_k} \tag{A.12}$$

where the N_i are prime powers. Hence for any $g \in G$ we can make the association

$$g \mapsto (g_1, ..., g_k) \tag{A.13}$$

where $g_i \in \mathbb{Z}_{N_i}$. Denote the identity of G by $e \mapsto (0, ..., 0)$. Let $\beta_1 \mapsto (1, 0, ..., 0)$, $\beta_2 \mapsto (0, 1, 0, ..., 0)$,..., $\beta_k \mapsto (0, 0, ..., 1)$ and note that any element $g = (g_1, ..., g_k)$ can be written as

$$g \mapsto g_1\beta_1 + ... + g_k\beta_k = \sum_{i=1}^{k} g_i\beta_i. \tag{A.14}$$

Theorem 7 *A finite abelian group G is isomorphic to char(G)*

Proof: Let $\chi : G \to \mathbb{C}$ be a character of G. Then we have

$$\chi(g) = \chi\left(\sum_{j=1}^{k} g_j\beta_j\right) = \prod_{j=1}^{k} \chi\left(\beta_j\right)^{g_j} \tag{A.15}$$

which implies that χ is completely determined by its action on the β_j. Moreover, since the β_j have order N_j, $\chi(\beta_j)$ must have order dividing N_j. Therefore $\chi(\beta_j) =$

$\omega_{N_j}^{h_j}$ *where* ω_{N_j} *is the primitive* N_j*th root of unity* $e^{\frac{2\pi i}{N_j}}$ *and* h_j *is an integer in* $\{0, 1, ..., N_j - 1\}$.

This gives that any given character $\chi : G \to \mathbb{C}$ *is determined by a k-tuple* $(h_1, ..., h_k)$ *and clearly each such k-tuple is associated with an element in G in a one to one manner. Therefore the characters may be labeled uniquely by elements of G and a character* χ_g *is given by*

$$\chi_g(h) = \prod_{j=1}^{k} \omega_{N_j}^{g_j h_j} \tag{A.16}$$

for $h \in G$. *Note from the above that* $\forall g, h \in G$,

$$\chi_g(h) = \chi_h(g) \tag{A.17}$$

and

$$\chi_g(-h) = \frac{1}{\chi_g(h)}. \tag{A.18}$$

Multiplication in the character group is given by $\chi_g \chi_h = \chi_{g+h}$, *where* χ_e *is the identity. Therefore the mapping* $g \mapsto \chi_g$ *is a homomorphism and so an isomorphism.*

Next we introduce the notion of orthogonal elements relative to a set.

Definition 18 *Orthogonal Element*

Let G be finite and abelian and $X \subset G$ *with* $h \in G$. *We say that h is orthogonal to X if* $\forall x \in X$, $\chi_h(x) = 1$.

Proposition 13 *If* $X \subseteq G$ *then the set of all elements that are orthogonal to X, denoted* $X^\perp$, *is a group*

Proof: $\chi_e \in X^\perp$. *If a, b* $\in X^\perp$ *then for all* $x \in X$, $\chi_{a-b}(x) = \chi_x(a-b) = \frac{\chi_x(a)}{\chi_x(b)} = 1$. *Thus,* $a - b \in X^\perp$ *and so* $X^\perp$ *is a subgroup of G.*

Definition 19 *Orthogonal Subgroup*

If $H \leq G$ *then* $H^\perp$ *is called the orthogonal subgroup for H.*

Theorem 8 *Let $\chi_h \in$ char(G). Then*

1. $h = e \Rightarrow \sum_{g\in G} \chi_h(g) = |G|$

2. $h \neq e \Rightarrow \sum_{g\in G} \chi_h(g) = 0$

Proof:

$$\begin{aligned}\sum_{g\in G} \chi_h(g) &= \sum_{g_1\in\mathbb{Z}_{N_1}} \ldots \sum_{g_k\in\mathbb{Z}_{N_k}} \prod_{j=1}^{k} \omega_{N_j}^{h_j g_j} &\text{(A.19)}\\ &= \left(\sum_{g_1\in\mathbb{Z}_{N_1}} \omega_{N_1}^{h_1 g_1}\right) \ldots \left(\sum_{g_k\in\mathbb{Z}_{N_k}} \omega_{N_k}^{h_k g_k}\right). &\text{(A.20)}\end{aligned}$$

If for some j, $h_j \neq 0$ then $\omega_{N_j}^{h_j} \neq 1$ and so the geometric sum formula gives

$$\sum_{g_j\in\mathbb{Z}_{N_j}} \left(\omega_{N_j}^{h_j}\right)^{g_j} = \frac{1-\left(\omega_{N_j}^{h_j}\right)^{N_j+1}}{1-\omega_{N_j}^{h_j}} = 0 \qquad \text{(A.21)}$$

Now, $h_j \neq 0$ for some j if and only if $h \neq e$. If $h = e$ then clearly the sum is $|G|$.

Corollary 4 *Label the elements of the finite abelian group G as $\{g_1 = e, ..., g_{|G|}\}$. Then each character χ_h may be represented by a vector*

$$\frac{1}{\sqrt{|G|}} \begin{bmatrix} \chi_h(g_1) \\ . \\ . \\ . \\ \chi_h(g_{|G|}) \end{bmatrix}$$

and the set of these vectors form an orthonormal basis for $\mathbb{C}^{|G|}$.

Proof: Let h, $h' \in G$. Then

$$\begin{aligned}\frac{1}{|G|}\sum_{g\in G} \chi_h(g)\chi_{h'}(g)^* &= \frac{1}{|G|}\sum_{g_1\in\mathbb{Z}_{N_1}} \ldots \sum_{g_k\in\mathbb{Z}_{N_k}} \prod_{j=1}^{k} \omega_{N_j}^{h_j g_j} \prod_{l=1}^{k} \omega_{N_l}^{-h_l' g_l} &\text{(A.22)}\\ &= \frac{1}{|G|}\sum_{g_1\in\mathbb{Z}_{N_1}} \ldots \sum_{g_k\in\mathbb{Z}_{N_k}} \prod_{j=1}^{k} \omega_{N_j}^{(h_j - h_j')g_j} &\text{(A.23)}\end{aligned}$$

From the above theorem this is 0 *if and only if* $h \neq h'$. *Otherwise, if* $h = h'$ *it is* 1. *Hence the set of vectors forms an orthonormal basis in* $\mathbb{C}^{|G|}$.

The following is proved in [30]

Theorem 9 *The following relations hold:*

1.

$$G/H \cong H^{\perp} \tag{A.24}$$

2.

$$H^{\perp\perp} = H. \tag{A.25}$$

Using the representation and character theory introduced in this section we can now define the quantum Fourier transform over more general groups than the cyclic group.

A.3 The Finite Abelian Group Quantum Fourier Transform

Suppose we associate the elements of G to an orthonormal basis of a finite dimensional Hilbert space of dimension $|G|$ by $g \mapsto |g\rangle$. By the decomposition theorem for finite abelian groups, the Hilbert space will be assumed to have the tensor product structure defined by $g = (g_1, ..., g_k) \mapsto |g_1\rangle ... |g_k\rangle = |g\rangle$.

Definition 20 *Quantum Fourier Transform, Translation Operator and Phase Change Operator*

We define the following three operators

1. The finite abelian group quantum Fourier transform

$$\mathcal{F}_G = \frac{1}{\sqrt{|G|}} \sum_{g,h \in G} \chi_g(h) |g\rangle\langle h| \tag{A.26}$$

2. The translation operator

Let $t \in G$.

$$\tau_t = \sum_{g \in G} |t + g\rangle\langle g| \tag{A.27}$$

3. The phase change operator

Let $h \in G$,

$$\phi_h = \sum_{g \in G} \chi_g(h)|g\rangle\langle g| \tag{A.28}$$

A direct result of the fact that $char(G)$ forms an orthonormal basis in $\mathbb{C}^{|G|}$ is the Fourier transform defined above is a unitary operator. The next proposition shows that $\mathcal{F}_G$ can be decomposed as a tensor product of cyclic QFT's using the structure theorem for finite abelian groups.

Proposition 14 $\mathcal{F}_G = \otimes_{j=1}^{k} \mathcal{F}_{N_j}$

Proof: We have that

$$\begin{aligned}
F_G &= \frac{1}{\sqrt{|G|}} \sum_{g,h \in G} \chi_g(h)|g\rangle\langle h| && \text{(A.29)}\\
&= \frac{1}{\sqrt{|G|}} \sum_{g,h \in G} \left(\prod_{j=1}^{k} \omega_{N_j}^{g_j h_j} \right) |g_1\rangle ... |g_k\rangle\langle ... h_1| ... \langle h_k| && \text{(A.30)}\\
&= \frac{1}{\sqrt{|G|}} \sum_{g,h \in G} \left(\omega_{N_1}^{g_1 h_1} |g_1\rangle\langle h_1| \otimes ... \otimes \omega_{N_k}^{g_k h_k} |g_k\rangle\langle h_k| \right) && \text{(A.31)}\\
&= \frac{1}{\sqrt{N_1}} \left(\sum_{g_1, h_1 \in \mathbb{Z}_{N_1}} \omega_{N_1}^{g_1 h_1} |g_1\rangle\langle h_1| \right) \otimes ... && \text{(A.32)}\\
&\quad \otimes \frac{1}{\sqrt{N_k}} \left(\sum_{g_k, h_k \in \mathbb{Z}_{N_k}} \omega_{N_k}^{g_k h_k} |g_k\rangle\langle h_k| \right) && \text{(A.33)}\\
&= \otimes_{j=1}^{k} F_{N_j} && \text{(A.34)}
\end{aligned}$$

The following is a simple result [30].

Theorem 10 *For $H \leq G$, let $|H\rangle = \frac{1}{\sqrt{|H|}} \sum_{h \in H} |h\rangle$. Then*

1.

$$F_G|H\rangle = |H^{\perp}\rangle \tag{A.35}$$

2.

$$F_G\tau_t = \phi_t F_G \tag{A.36}$$

Next, we state and prove the hidden subgroup problem using the results of this section.

A.4 The Finite Abelian Hidden Subgroup Problem

First we give the statement of the hidden subgroup problem for finite abelian groups.

Definition 21 *The Finite Abelian Hidden Subgroup Problem (HSP):*

Let G be a finite abelian group, X a finite set, and $f : G \to X$ a function such that there is $H \leq G$ which satisfies the following property

$$\forall a, b \in G, f(a) = f(b) \textit{ if and only if} a - b \in H. \tag{A.37}$$

Then, with probability at least $1 - \frac{1}{|G|}$ determine a generating set for H in $O(polylog|G|)$ operations.

There always exists the $O(|G|)$ time algorithm for determining H that involves evaluating the image of each element under f. The question becomes whether there exists an algorithm that reduces the naive $O(|G|)$ time algorithm to one that is $O(polylog|G|)$ time. In the quantum case, the answer is yes, while in the classical case there is no known $O(polylog|G|)$ time algorithm. Before describing the quantum algorithm we briefly discuss the assumptions.

Assumptions:

1. The finite set X will without loss of generality be thought of as the integers modulo $|X|$ under addition.

2. The decomposition of G into a direct product of cyclic groups is known. Given generators of G, finding such a decomposition is possible via a quantum algorithm that runs in $O(polylog|G|)$ time [7].

3. There is an efficient encoding of G and X to basis states of the quantum computer being used. Thus, we assume that the first register of our quantum computer is associated with G and the second register is associated with X.

4. There is a method of evolving the state of the system under a unitary operation that is equivalent to calling the function f. This is done by assuming there is a quantum "black-box" that performs the unitary transformation $U_f|g\rangle|x\rangle = |g\rangle|x \oplus f(g)\rangle$ where $g \in G$, $x \in X$. To see that U_f is unitary, note that $U_f|g_1\rangle|x_1\rangle = U_f|g_2\rangle|x_2\rangle$ if and only if $|g_1\rangle|x_1 \oplus f(g_1)\rangle = |g_2\rangle|x_2 \oplus f(g_2)\rangle$. Hence, $g_1 = g_2$ and $x_1 \oplus f(g_1) = x_2 \oplus f(g_2)$. Therefore $g_1 = g_2$ and $x_1 = x_2$.

We can now present the polynomial time algorithm for solving the hidden subgroup problem.

Algorithm:

1. Prepare the initial state $|0\rangle|0\rangle$ and apply F_G to the first register to obtain the state $\frac{1}{\sqrt{|G|}}\sum_{g\in G}|g\rangle|0\rangle$

2. Apply U_f to obtain the state $\frac{1}{\sqrt{|G|}}\sum_{g\in G}|g\rangle|f(g)\rangle$. Since f is promised to be constant and distinct on cosets, if we let $T = \{t_1, ..., t_m\}$ be a set of representatives for the different cosets then

$$\frac{1}{\sqrt{|G|}}\sum_{g\in G}|g\rangle|f(g)\rangle = \frac{1}{\sqrt{|T|}}\sum_{t\in T}\left(\frac{1}{\sqrt{|H|}}\sum_{h\in H}|t+h\rangle\right)|f(t)\rangle \quad \text{(A.38)}$$

$$= \frac{1}{\sqrt{|T|}}\sum_{t\in T}\left(\tau_t|H\rangle\right)|f(t)\rangle \quad \text{(A.39)}$$

3. Apply F_G to the first register to get $\frac{1}{\sqrt{T}}\sum_{t\in T}(F_G\tau_t|H\rangle)|f(t)\rangle$. Then, by the above theorems,

$$\frac{1}{\sqrt{|T|}}\sum_{t\in T}(F_G\tau_t|H\rangle)|f(t)\rangle = \frac{1}{\sqrt{T}}\sum_{t\in T}(\phi_t F_G|H\rangle)|f(t)\rangle \tag{A.40}$$

$$= \frac{1}{\sqrt{|T|}}\sum_{t\in T}(\phi_t|H^\perp\rangle)|f(t)\rangle \tag{A.41}$$

$$\tag{A.42}$$

Now, the action of ϕ_t on a state $|g\rangle$, $g \in G$, serves only to introduce a phase. Hence if one performs a measurement on the first register in the basis $\{|g\rangle : g \in G\}$, then in terms of measurement statistics, the above state is equivalent to

$$\left(\frac{1}{\sqrt{|H^\perp|}}\sum_{h'\in H^\perp}|h'\rangle\right)\otimes\left(\frac{1}{\sqrt{|T|}}\sum_{t\in T}|f(t)\rangle\right) \tag{A.43}$$

Therefore performing a measurement on the first register is equivalent to uniformly sampling from $H^\perp$.

This concludes the quantum mechanical aspects of the algorithm. Let us determine the time cost of a single run of the above algorithm. The only resources required are those needed for performing $\mathcal{F}_G$. By a result proved above,

$$\mathcal{F}_G = \otimes_{j=1}^{k}\mathcal{F}_{N_j} \tag{A.44}$$

where the N_j are the prime powers arising from the decomposition $G \cong \mathbb{Z}_{N_1} \oplus ... \oplus \mathbb{Z}_{N_k}$. Now, since the N_j are powers of the primes that $|G|$ factorizes into, there are at most $log|G|$ of them, ie. k $\leq log|G|$. We also know that each cyclic QFT $\mathcal{F}_{N_j}$ has a circuit decomposition requiring $log(N_j)^2$ gates. Hence, by the decomposition for $\mathcal{F}_G$ given above we have that the number of gates required to implement $\mathcal{F}_G$ is at most $log|G|max_j\{log(N_j)^2\}$ which is bounded by $log^3|G|$. Hence there exist circuit decompositions for $\mathcal{F}_G$ that are poly-logarithmic in $|G|$.

What remains is to find a generating set for $H^\perp$ and, from this set, determine a generating set for H. It can be shown [30] that choosing $t + log|G|$ uniformly random elements of any finite group G will generate G with probability at least $1 - \frac{1}{2^t}$. Hence, running the above algorithm $t_1 + log|H^\perp| \leq t_1 + log|G|$ times gives a generating set for $H^\perp$ with probability at least $1 - \frac{1}{2^{t_1}}$. It can also be shown [30] that if one has a set of t_1 generators for $H^\perp$ then finding a uniformly random element of H is equivalent to randomly solving a set of t_1 linear equations.

Randomly solving this set of linear equations $t_2 + log|H| \leq t_2 + log|G|$ times generates H with probability at least $1 - \frac{1}{2^{t_2}}$.

Solving the set of equations once can be done in $O(log|G|loglog|G|)$ steps. Therefore if one runs the algorithm $t_1 + log|G|$ times to get t_1 generators of $H^{\perp}$ with probability at least $1 - \frac{1}{2^{t_1}}$, and then randomly solves the set of t_1 linear equations $t_2 + log|G|$ times, they will have a generating set of H with probability at least $(1 - \frac{1}{2^{t_1}})(1 - \frac{1}{2^{t_2}})$ in $(t_1 + log|G|)log^3|G| + (t_2 + log|G|)O(log|G|loglog|G|)$ operations. If we choose $t_1 = t_2 = log|G| + 1$ then the probability of success is at least $1 - \frac{1}{|G|}$ with $O(polylog|G|)$ number of operations.

Appendix B

Concentration of Measure

B.1 Topology and Measure Theory

B.1.1 Topology

The following is a basic introduction to topology. A reference for further reading is [32]

Definition 22 *Topology*

A topology τ on a set X is a subset of the power set, $\mathcal{P}(X)$, of X that satisfies

1. *If A_i, $i \in I$, are in τ then $\cup_{i\in I} A_i \in \tau$.*
2. *If A and B are in τ then $A \cap B$ is in τ.*
3. *$\emptyset$, $X \in \tau$.*

Elements of τ are called open sets. A set X with a topology τ defined on it is called a topological space and denoted (X, τ).

Definition 23 *Open Cover, Compact Topological Space*

If X is a topological space then an open cover of X is a subset A_i of τ such that $\cup_{i\in I} A_i = X$. A compact topological space is one for which every open cover contains a finite sub-collection that also covers X.

If X and Y are topological spaces with topologies τ_1 and τ_2, there is a natural topology, τ_p, one can put on $X \times Y$ called the product topology. Elements of τ_p are arbitrary unions of sets of the form $U \times V$ where $U \in \tau_1$ and $V \in \tau_2$.

Definition 24 *Continuous Function*

A function, f, from a topological space (X, τ_1) *to a topological space* (Y, τ_2) *is called continuous if for every open set* V *in* τ_2, $f^{-1}(V)$ *is open in* τ_1.

Definition 25 *Topological Group*

A topological group is both a group and a topological space (X, τ) *such that*

1. *The group operation is continuous from* $(X \times X, \tau_p)$ *to* (X, τ).
2. *The mapping defined by* $g \rightarrow g^{-1}\, \forall g \in X$ *is continuous from* (X, τ) *to* (X, τ).

An example of a compact topological group is the unitary group U(D) under the usual operation of multiplication. A compact Lie group [28] is, loosely speaking, a differentiable manifold with a group operation that is smooth with respect to the defined manifold. U(D) is a compact Lie group as a submanifold of $\mathbb{C}^{D^2}$.

B.1.2 Measure Theory

The following is a basic introduction to measure theory. A reference for further reading is [6].

Definition 26 *Algebra, σ-Algebra*

Let X be a set and $\mathcal{M}$ *be a subset of* $\mathcal{P}(X)$. $\mathcal{M}$ *is called an algebra of sets if*

1. $\emptyset \in \mathcal{M}$.
2. $A,B \in \mathcal{M} \Rightarrow A \cup B \in \mathcal{M}$.
3. $A \in \mathcal{M} \Rightarrow X \setminus A \in \mathcal{M}$.

If the second property is extended to countable unions, $\mathcal{M}$ *is called a* σ*-algebra of sets.*

Definition 27 *Measure, Measure Space, Measurable Sets and Probability Measure*

Let $\mathcal{M}$ *be a* σ*-algebra of subsets of some set X. A function* $\mu : \mathcal{M} \rightarrow \mathbb{R} \cup \{\infty\}$ *is called a measure if*

1. $\mu(\emptyset) = 0$.

2. For any countable disjoint collection of sets X_i*,* $\mu(\cup_i X_i) = \sum_i \mu(X_i)$*.*

3. $\mu(A) \geq 0 \, \forall A \in \mathcal{M}$*.*

The triple $(X, \mathcal{M}, \mu)$ *is called a measure space and elements of* $\mathcal{M}$ *are called measurable sets. If in addition to the above conditions* $\mu(X) = 1$ *then* μ *is called a probability measure.*

An example of a probability measure on a finite set X is the counting measure μ_C defined by

$$\mu_C\,(A) = \frac{|A|}{|X|} \tag{B.1}$$

for any subset A of X.

For any set X, $\mathcal{P}$(X) is a σ-algebra of subsets of X. Hence, if $\mathcal{S}$ is a subset of $\mathcal{P}(X)$ one can define the Borel algebra of $\mathcal{S}$, $B(\mathcal{S})$, as the smallest σ-algebra containing $\mathcal{S}$. In the case of a topological space (X, τ), the Borel algebra on (X, τ), $B\,(X, \tau)$ is the smallest σ-algebra containing all of the open sets of τ. A measure defined on $B\,(X, \tau)$ is called a Borel measure on (X, τ). The following is an important result for Borel measures on compact groups.

Theorem 11 *If* (X, τ) *is a compact topological group then there exists, up to a constant, a unique Borel measure* μ_H *on* (X, τ)*, called the bi-invariant Haar measure, satisfying the following conditions*

1. $\mu_H(xE) = \mu_H(E) = \mu_H(Ex) \, \forall x \in X \, \forall E \in B\,(X, \tau)$*.*

2. $\mu_H(U) > 0$ *for every non-empty open set* $U \in \tau$*.*

3. $\mu_H(K) < \infty$ *for every compact set* K*.*

Since the bi-invariant Haar measure is unique up to a constant and the third property implies $\mu_H(X) < \infty$ there exists a unique bi-invariant Haar probability measure on a compact group.

B.2 Concentration of Measure

Concentration of measure [31, 15, 37, 27] is a phenomenon that can be empirically understood by considering an unbiased coin-tossing experiment consisting of N trials where N is large. Let X be the state space composed of sequences of single trial outcomes. Define the function f from X to $\mathbb{N}$ by $f(x)$ = the number of heads observed. Empirically, f(x) is concentrated around the median value $\frac{N}{2}$ for f. Thus, under the counting measure μ_C on X, $\mu_C(f^{-1}(N-n, N+n))$ is close to 1 even for small n $\in \mathbb{N}$ as N grows large. The following discussion attempts to make these ideas rigorous.

Suppose (X, d, μ) is a metric space with Borel probability measure μ.

Definition 28 *ϵ-Neighbourhood*

For $S \subseteq X$ we define

$$N_\epsilon(S) = \{x \in X : \exists y \in S \text{ with } d(x,y) < \epsilon\} \tag{B.2}$$

and call it the ϵ-neighbourhood of S.

Definition 29 *Median*

Let $f : X \to \mathbb{R}$ be a continuous function. A median of f, denoted M(f), is defined by the inequalities

$$\frac{1}{2} = \mu\left(\{x \in X : f(x) \leq M(f)\}\right) = \mu\left(\{x \in X : f(x) \geq M(f)\}\right). \tag{B.3}$$

Definition 30 *Modulus of Continuity*

Let $f : X \to \mathbb{R}$ be continuous and $\epsilon > 0$. The modulus of continuity for f and epsilon, denoted $\omega_f(\epsilon)$, is

$$\omega_f(\epsilon) = sup\{|f(x) - f(y)| : d(x,y) \leq \epsilon\}. \tag{B.4}$$

Definition 31 *η-Lipschitz Functions*

A function $f : X \to \mathbb{R}$ is called η-Lipschitz if $\forall x, y \in X$,

$$|f(x) - f(y)| \leq \eta\, d(x,y) \tag{B.5}$$

Definition 32 *Concentration Function*

$\forall \epsilon > 0$ the concentration function of X with respect to ϵ is

$$\alpha_X(\epsilon) = 1 - inf\Big\{\mu\left(N_\epsilon(S)\right) : S \subseteq X \textit{ is Borel measurable and} \, \mu(S) \geq \frac{1}{2}\Big\}. \quad \text{(B.6)}$$

Two equivalent expressions for the concentration function are

$$\alpha_X(\epsilon) = sup\Big\{\mu\left(X - N_\epsilon(S)\right) : \text{ S is Borel measurable and} \mu(S) = \frac{1}{2}\Big\} \quad \text{(B.7)}$$

and

$$\alpha_X(\epsilon) = sup\{\mu\left(\{x : f(x) \geq Med(f) + \epsilon\}\right) : \text{f is 1-Lipschitz}\}. \quad \text{(B.8)}$$

Note that the concentration functions decrease as ϵ grows.

Definition 33 *Levy Family*

$\forall i \in \mathbb{N}$ let $\mathcal{F} = \{((X_i, d_i), \mu_i)\}$ be a family of metric spaces with Borel probability measures. $\mathcal{F}$ is called Levy if for any sequence of Borel sets $S_i \subseteq X_i$ such that lim inf $\mu_i(S_i) > 0$, and every $\epsilon > 0$,

$$lim_{i\to\infty}\mu_i(N_\epsilon(S_i)) = 1 \quad \text{(B.9)}$$

Equivalently, $\mathcal{F}$ is a Levy family if $\forall\epsilon > 0$, $\alpha_{X_i}(\epsilon) \to 0$ as $i \to \infty$.

Definition 34 *Normal Levy Family*

$\mathcal{F}$ is called a normal Levy family if there exist constants A, B such that $\forall i$ and $\epsilon > 0$

$$\alpha_{X_i}(\epsilon) \leq Ae^{-B\epsilon^2 i} \quad \text{(B.10)}$$

The following lemma, which is proved from the definitions of the concentration function and a normal Levy family, is required for the main result.

Lemma 1 *Let $\epsilon > 0$ and f be a continuous function on (X, d, μ) with modulus of continuity $\omega_f(\epsilon)$. Then,*

$$\mu\left(f^{-1}\left(M(f) - \omega_f\left(\epsilon\right), M(f) + \omega_f\left(\epsilon\right)\right)\right) \geq 1 - 2\alpha_X(\epsilon). \quad \text{(B.11)}$$

This gives the desired result [31].

Theorem 12 *Let $\mathcal{F}$ be a normal Levy family. For each i, let f_i be a continuous function on (X_i, d_i, μ_i) with median $M(f_i)$ and modulus of continuity $\omega_{f_i}(\epsilon)$ for each $\epsilon > 0$. Then $\forall \epsilon > 0$,*

$$\mu\left(f_i^{-1}\left(M(f_i) - \omega_{f_i}(\epsilon), M(f_i) + \omega_{f_i}(\epsilon)\right)\right) \geq 1 - 2\alpha_{X_i}(\epsilon) \geq 1 - 2Ae^{-B\epsilon^2 i} \quad \text{(B.12)}$$

The phenomenon of concentration of measure for high dimensional structures is depicted in the above inequality which says that the functions f_i become concentrated near their medians in terms of the measure on X_i. Next we look at various examples of normal Levy families.

B.3 Examples of Concentration of Measure

Example 1 *The standard example of a normal Levy family is that of unit spheres in $(\mathbb{R}^n, ||\,||_2)$ with the geodesic metric. The 2-norm, $||\,||_2$, on $\mathbb{R}^n$, is defined through the Euclidean inner product on $\mathbb{R}^n$. Taking $x = (x_1, ..., x_n) \in \mathbb{R}^n$,*

$$||x||_2 := \sqrt{< x, x >} = \sqrt{\sum_{i=1}^{n} x_i^2}. \quad \text{(B.13)}$$

As usual,

$$\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n : ||x|| = 1\}. \quad \text{(B.14)}$$

Let d be the geodesic (Riemannian) metric on $\mathbb{S}^{n-1}$,

$$d(x, y) = arccos < x, y > \quad \text{(B.15)}$$

which is the angle between x and y in $\mathbb{R}^n$. As well, let μ be the unique Borel (Haar) measure on $\mathbb{S}^{n-1}$ generated by the topology induced by d. Suppose f: $\mathbb{S}^{n-1} \to \mathbb{R}$ is continuous and let M(f) be the median of f. Then,

$$\mu\left(\{x \in \mathbb{S}^{n-1} : |f(x) - M(f)| \leq \epsilon\}\right) \geq 1 - \sqrt{\frac{\pi}{2}} e^{-\epsilon^2 \frac{(n-2)}{2}}. \quad \text{(B.16)}$$

Thus the set $\{(\mathbb{S}^{n+1} \subset \mathbb{R}^{n+2}, d_{n+1}, \mu_{n+1})\}$ indexed by $n \in \mathbb{N}$ is a Levy family with $A = \sqrt{\frac{\pi}{8}}$ and $B = \frac{1}{2}$.

Example 2 *Next, we look at the unitary group U(n). $\forall n \in \mathbb{N}$, $U(n)$ is a Lie group. In addition, $U(n)$ is compact and so can be equipped with a unique bi-invariant (ie. left and right invariant) metric. Let d_n be this bi-invariant metric on $U(n)$ which is actually induced by the trace inner product*

$$d(U,V) = \sqrt{tr((U-V)^\dagger(U-V))} \tag{B.17}$$

Denote the Haar measure on $U(n)$ by μ_n. The following theorem is proved in [31].

Theorem 13 *The family $\{(U(n), d_n), \mu_n\}$ is a normal Levy family with constants $A = \sqrt{\frac{\pi}{8}}$ and $B = \frac{1}{8}$.*

Example 3 *This example also deals with $\mathbb{S}^n$ defined above, however the metric defined on it is the Euclidean metric $\| \|_2$. In this case, by Appendix V of [31], we have for an η-Lipschitz function $f : \mathbb{S}^n \to \mathbb{R}$ (with respect to $\| \|_2$) and the Haar measure of example 1,*

$$\mu\left(f^{-1}\left(-\infty, \int f d\mu - \epsilon\right)\right) \leq 2e^{\frac{-C\epsilon^2(n+1)}{\eta^2}} \tag{B.18}$$

and

$$\mu\left(f^{-1}\left(\int f d\mu + \epsilon, +\infty\right)\right) \leq 2e^{\frac{-C\epsilon^2(n+1)}{\eta^2}} \tag{B.19}$$

where $C = \frac{1}{9\pi^3 ln2}$ and $\int f d\mu$ is the integral of f with respect to the Haar measure.This implies,

$$\mu\left(f^{-1}\left(\int f d\mu - \epsilon, \int f d\mu + \epsilon\right)\right) \geq 1 - 4e^{\frac{-C\epsilon^2(n+1)}{\eta^2}}. \tag{B.20}$$

Additionally, a relationship between the measure of $f^{-1}(\int f d\mu - \epsilon, \int f d\mu + \epsilon)$ and $f^{-1}(M(f) - \epsilon, M(f) + \epsilon)$ is given which results in analogous inequalities,

$$\mu\left(f^{-1}\left(-\infty, M(f)-\epsilon\right)\right) \leq e^{\frac{-D\epsilon^2(n-1)}{\eta^2}}, \tag{B.21}$$

$$\mu\left(f^{-1}\left(M(f)+\epsilon, +\infty\right)\right) \leq e^{\frac{-D\epsilon^2(n-1)}{\eta^2}}, \tag{B.22}$$

and

$$\mu\left(f^{-1}\left(M(f)-\epsilon, M(f)+\epsilon\right)\right) \geq 1-2e^{\frac{-D\epsilon^2(n-1)}{\eta^2}} \tag{B.23}$$

where $D = \frac{1}{2\pi^2 ln2}$.

Other examples of Normal Levy families are the permutation groups and Hamming cubes $\{0,1\}^n$ of all binary strings of length n. Both are equipped with the normalized Hamming distance and the normalized counting measure. Further examples can be found in in [37, 31, 27].

Appendix C

Finding Noiseless Subsystems for Unital Channels

C.1 Basic Algebra Theory

C.1.1 C^*-Algebras

The theory of algebras plays a central role in the theory of noiseless subsystems, thus it is imperative to present basic definitions and theorems regarding these objects [2, 18, 33].

Definition 35 *Ring*

Let $\mathcal{R}$ be a set such that two binary operations, $+$ and $\cdot$, are defined on $\mathcal{R}$ with the following properties:

1. $\mathcal{R}$ is an abelian group over $+$

2. $\forall\, a, b, c \in \mathcal{R}$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (Associativity)

3. $\exists 1 \in \mathcal{R}$ such that $\forall a \in \mathcal{R}$, $a \cdot 1 = 1 \cdot a$ (Identity)

4. $\forall\, a, b, c \in \mathcal{R}$, $a \cdot (b + c) = a \cdot b + a \cdot c$ (Distributivity)

5. $\forall\, a, b, c \in \mathcal{R}$, $(a + b) \cdot c = a \cdot c + b \cdot c$.

Then $\mathcal{R}$ is called a ring.

Definition 36 *Algebra*

Let $\mathcal{A}$ be a set satisfying the following properties

1. $\mathcal{A}$ is a ring

2. The abelian group $(\mathcal{A}, +)$ is a linear space over some field F

3. If x, y are in $\mathcal{A}$ and $\alpha \in F$ then $\alpha(xy) = (\alpha x)y = x(\alpha y)$

Then $\mathcal{A}$ is called an algebra over F.

For the rest of the definitions up to and including C^*-algebras, F will be $\mathbb{C}$.

Definition 37 *Normed Algebra*

Suppose $\mathcal{A}$ is an algebra and that there is a norm, $\| \ \|$, defined on $\mathcal{A}$ such that

4. If x, y are in $\mathcal{A}$ then $\|xy\| \leq \|x\|\|y\|$.

Then $\mathcal{A}$ is called a normed algebra.

Definition 38 *Banach Algebra*

Suppose the normed algebra $\mathcal{A}$ is a Banach space with respect to $\| \ \|$, then it is called a Banach algebra.

The following defines a *-algebra in terms of an involution mapping.

Definition 39 *Involution and *-Algebra*

Let * be a mapping from a set D into itself satisfying the following properties

$$(x^*)^* = x \tag{C.1}$$

$$(x + y)^* = x^* + y^* \tag{C.2}$$

$$(xy)^* = y^* x^* \tag{C.3}$$

$$(\alpha x)^* = \bar{\alpha} x^* \tag{C.4}$$

Then * is called an involution on D. An algebra with an involution * is called a *-algebra.

Finally, we can characterize a C^*-algebra.

Definition 40 *C^*-Algebra*

Suppose that $\mathcal{A}$ is a Banach *-algebra satisfying the following $\forall x \in \mathcal{A}$

$$\|x^*x\| = \|x\|^2. \tag{C.5}$$

Then $\mathcal{A}$ is called a C^*-algebra.

From the definitions above, it is clear that the complex field is a C^*-algebra where it is treated as a vector space over itself. The norm is the usual Euclidean norm and the involution is just the mapping taking each element to its complex conjugate. References for further clarification of the terminology used in the upcoming theorems regarding C^*-algebras can be found in [2, 10].

Next, we will show that the set of bounded linear operators on a Hilbert space H forms a C^*-algebra. For a Hilbert space H, denote the set of all bounded linear operators on H by $B(\mathcal{H})$.

Theorem 14 *$B(\mathcal{H})$ is a C^*-algebra under the operator norm induced by the norm defined on H.*

Proof: The set of all bounded linear operators defined on any inner product space forms a ring. The addition operation is defined pointwise and multiplication of operators is given by composition. The multiplication operation is associative, the addition of operators forms an abelian group and the required distributive laws hold. Since $B(\mathcal{H})$ on an inner product space forms a vector space under addition and scalar multiplication of operators, and property 3 *in Definition* 1 *clearly holds, $B(\mathcal{H})$ forms an algebra. The norm defined on $B(\mathcal{H})$ is the operator norm induced by the inner product on H. This norm satisfies*

$$F, G \in B(\mathcal{H}) \Rightarrow \|FG\| \leq \|F\|\|G\| \tag{C.6}$$

since if $x \in H$,

$$\|FG(x)\| = \|F[G(x)]\| \leq \|F\|\|G(x)\| \leq \|F\|\|G\|\|x\|. \tag{C.7}$$

This implies $\forall x \neq 0$

$$\frac{\|FG(x)\|}{\|x\|} \leq \|F\|\|G\|. \tag{C.8}$$

Therefore $\|FG\| \leq \|F\|\|G\|$, and $B(\mathcal{H})$ is a normed algebra. Since H is a Banach space, $B(\mathcal{H})$ is a Banach space, and so $B(\mathcal{H})$ is a Banach algebra. The involution * defined on $B(\mathcal{H})$ is the obvious one taking each element to its adjoint relative to the inner product on H. That is, $\forall A \in B(\mathcal{H})$,

$$^*(A) = A^*. \tag{C.9}$$

The adjoint operation * on $B(\mathcal{H})$ satisfies all of the properties required for * to be an involution. Finally, it needs to be shown that $\forall L \in B(\mathcal{H})$

$$\|L^*L\| = \|L\|^2. \tag{C.10}$$

From above and the fact that $\|L^*\| = \|L\|$ we have,

$$\|L^*L\| \leq \|L^*\|\|L\| = \|L\|^2 \tag{C.11}$$

So if we can show $\|L^*L\| \geq \|L\|^2$ then we are done. However,

$$\|Lx\|^2 = \langle Lx|Lx\rangle = \langle L^*Lx|x\rangle \tag{C.12}$$

and by the Cauchy-Schwartz inequaltiy,

$$\langle L^*Lx|x\rangle \leq \|L^*Lx\|\|x\| \leq \|L^*L\|\|x\|^2. \tag{C.13}$$

So for all x

$$\|Lx\|^2 \leq \|L^*L\|\|x\|^2 \tag{C.14}$$

which gives $\|L^*L\| \geq \|L\|^2$. Thus $B(\mathcal{H})$ is a C^*-algebra.

Next, we give conditions for when a subset of a C^*-algebra is itself a C^*-algebra.

Theorem 15 *A subset K of a C^*-algebra $\mathcal{A}$ is a sub C^*-algebra if and only if the following are satisfied*

1. The vector space and multiplication operations on K are closed

2. K is closed under the involution operation

3. K is closed in the norm-induced metric topology

Proof: The first condition of operations being closed is an obvious one. These conditions give that K is itself a ring and a vector space. This implies K is a normed linear associative algebra since all of the other required properties are inherited from the structure of $\mathcal{A}$. The second property defines an involution on K in the same manner as was defined on $\mathcal{A}$. The requirement of being closed under the norm topology just states that K is complete and hence a Banach space, since every closed subset of a Banach space is complete. Finally, the requirement $\forall L \in K$, $\|L^*L\| = \|L\|^2$ is satisfied since the involution defined on K is exactly the same as the one defined on A.

C.1.2 Wedderburn Structure Theorems

Before presenting the Wedderburn structure theorems, a few more definitions must be made.

Definition 41 *Ideal*

Let R be a ring and I a non-empty subset of R. Then I is called a (two-sided) ideal if the following are satisfied

1. I is an additive subgroup of R

2. For every $r \in R$, $rI \subseteq I$ and $Ir \subseteq I$, where $rI = \{ra|a \in I\}$ and Ir is defined similarly.

If only $rI \subseteq I$ is satisfied for all $r \in R$, then I is called a left ideal of R. Right ideals are defined similarly.

Definition 42 *Division Ring*

Let R be a ring and suppose that for all a and b in R, $ab = 0 \Rightarrow a = 0$ or $b = 0$. Then R is called a division ring.

The above definition for a division ring may be equivalently stated as for all $a \in R$, there exists $b \in R$ such that $ab = ba = 1$. That is every element has an inverse. An algebra whose ring structure is a division ring is called a division algebra.

Definition 43 *(Proper) Nilpotent Elements and the Radical*

Let $\mathcal{A}$ be an algebra and $x \neq 0$ be in $\mathcal{A}$. Then x is said to be nilpotent if for some $n \in \mathbb{N}$, $x^n = 0$. x is said to be properly nilpotent if $\forall a \in \mathcal{A}$, ax and xa are either nilpotent or zero. The set of all properly nilpotent elements of an algebra is called the radical of the algebra.

It can easily be shown that the radical of an algebra is an ideal.

Definition 44 *Simple Algebra*

Let $\mathcal{A}$ be a non-zero algebra. If the set of all ideals of $\mathcal{A}$ consists only of the zero ideal and $\mathcal{A}$ itself then $\mathcal{A}$ is called a simple algebra.

Definition 45 *Semi-Simple Algebra*

If $\mathcal{A}$ is an algebra such that the radical of $\mathcal{A}$ is the zero ideal then $\mathcal{A}$ is called semi-simple.

Note that the set of all linear operators on a finite dimensional vector space is simple, and since it contains an identity element, a semi-simple algebra. We now present the Wedderburn structure theorems [33, 38].

Theorem 16 *First Wedderburn Structure Theorem*

$\mathcal{A}$ is a finite dimensional semi-simple algebra if and only if $\mathcal{A}$ is isomorphic to a direct sum of simple algebras. Moreover, the direct sum decomposition of $\mathcal{A}$ is unique up to ordering.

Theorem 17 *Second Wedderburn Structure Theorem*

Let $\mathcal{A}$ be a finite dimensional simple algebra. Then $\mathcal{A}$ is isomorphic to $\mathcal{A} = M \otimes D$ where M is a full matrix algebra and D is a division algebra

These two theorems taken together imply that every finite dimensional semi-simple algebra is isomorphic to the direct sum of factors of matrix and division algebras. These theorems are the foundation for representations of algebras. For our presentation we will be concerned with finite-dimensional C^*-algebras over $\mathbb{C}$. In this case we have the following stronger representation theorem. [2]

Theorem 18 *Unitary Equivalence (UE) Theorem*

Let $\mathcal{A}$ be a finite dimensional C^*-algebra over $\mathbb{C}$. Then $\mathcal{A}$ is isometrically *-isomorphic to an orthogonal direct sum of the form

$$\mathcal{A} \cong (M_{n_1} \otimes I_{m_1}) \oplus ... \oplus (M_{n_d} \otimes I_{m_d}) \tag{C.15}$$

where the above representation is unique up to ordering. M_{n_i} is a full matrix algebra on an n_i dimensional space and I_{m_i} is the identity on an m_i dimensional space. All spaces are over $\mathbb{C}$.

C.2 Unital Channels

The UE theorem will give us our starting point for the discussion of unital channels, however before proceeding let us discuss a few results given by the theorem. First, note that the UE theorem splits the Hilbert space on which the representation of $\mathcal{A}$ acts into orthogonal subspaces, each of which is a direct product of subspaces. Moreover, the the algebra acts invariantly on one of these factor spaces.

If we have a set E of linear operators acting on a finite dimensional Hilbert space H, and this set of operators is closed under ring and vector space operations, and is closed under adjoints (†-closed) then E is a sub C^*-algebra of $B(H)$. The condition of being closed in the metric topology is trivially satisfied since E is a finite dimensional subspace of $B(H)$. Thus, E has a decomposition of the form given in the UE theorem.

The above remarks give some intuition as to how to proceed. If a unital channel can be described by a finite dimensional C^*-algebra of operators acting on a finite dimensional Hilbert space then the UE theorem tells us there exist subsystems in the Hilbert space unaffected by the quantum channel. The problem will then be to find these subsystems. Before explicitly giving an algorithm for finding these subsystems we state a few definitions and prove a structure theorem for unital channels.

Definition 46 *Reduction and Invariance of an Operator*

A projection P on a Hilbert space reduces an element T of $B(H)$ if T and P commute, that is $T(PH) \in PH$ and $T(P^{\perp}H) \in P^{\perp}H$. As well, $T \in B(H)$ is said to be invariant on PH if $TP = PTP$, that is only $T(PH) \in PH$ is satisfied.

We now state without proof [19] an important lemma for unital channels.

Lemma 2 *Suppose Λ is a unital quantum channel and P is a projection on H. Then the following are equivalent.*

A) $\Lambda(P) = P$

B) P reduces any set of Kraus operators describing Λ

C) PH is invariant for any set of Kraus operators describing Λ

D) PH is invariant for $\{A_i^\dagger\}$ where the A_i are any set of Kraus operators describing Λ

In this lemma it is clear that if P is a rank one projection $|\psi\rangle\langle\psi|$ then $|\psi\rangle$ is a joint eigenstate of the A_k. Lemma 1 will be of importance throughout the rest of the paper.

Definition 47 *Interaction Algebra*

Let Λ be a quantum channel and $A_1, ..., A_n$ be Kraus operators for the operator sum decomposition. The interaction algebra generated by $A_1, ..., A_n$, denoted by $\mathcal{A} = Alg\{Ai\}$, is the algebra of polynomials generated by $\{A_i\}$

The Cayley-Hamilton theorem states that every operator in $B(H)$ is the root of some polynomial of degree at most the dimension of the space the operator acts on [18]. As a result, the degree of $Alg\{Ai\}$ is bounded uniformly by the same positive integer.

Definition 48 *Noise Commutant*

Let $\mathcal{A}$ be the interaction algebra generated by $A_1, ..., A_n$. The noise commutant of $\mathcal{A}$, denoted by $\mathcal{A}'$, is the set of all linear operators that commute with every element in $\mathcal{A}$.

Note that since the A_i generate $\mathcal{A}$, it is sufficient to have $\mathcal{A}'$ be the set of all linear operators that commute with each of the A_i. As well, clearly $\mathcal{A}'$ is an algebra.

Definition 49 *Fixed Point Set*

The fixed point set of a quantum channel Λ is given by the set of all elements in $B(H)$ that are fixed points of Λ. The fixed point set is denoted by $Fix(\Lambda)$.

Note that $Fix(\Lambda)$ is †-closed by the representation of Λ through an operator sum decomposition. Now, we state a well known theorem due to Von Neumann [36] which will be of use later on.

Theorem 19 *Von Neumann Double Commutant Theorem (VNDC)*

Let $\mathcal{A}$ be an algebra of bounded operators on a Hilbert space H, containing the identity operator and closed under taking adjoints, that is, $\mathcal{A}$ is a unital †-closed subalgebra of $B(H)$. Then the closures of $\mathcal{A}$ in the weak operator topology and the strong operator topology [36] are equal, and are in turn equal to the double commutant $(\mathcal{A}')'$ of $\mathcal{A}$. This algebra is the von Neumann algebra generated by $\mathcal{A}$.

If the interaction algebra, $Alg\{Ai\}$, is †-closed, that is $Alg\{Ai\}$ is a C^*-algebra, then $Alg\{Ai\} = (\mathcal{A}')'$. This is clear since by definition, $Alg\{Ai\}$ is an algebra in $B(H)$. Also, $Alg\{Ai\}$ contains the identity by observing either the unitality or trace preserving conditions. Moreover, since the algebra of operators is finite dimensional, the closures of $Alg\{Ai\}$ in the weak operator and strong operator topology are just $Alg\{Ai\}$. Thus by VNDC if $Alg\{Ai\}$, is †-closed $Alg\{Ai\} = (\mathcal{A}')'$. The following theorem shows that in fact $Alg\{Ai\}$ is †-closed and gives some of the main results needed for the rest of the paper.

C.2.1 Structure Theorem For Unital Channels

Theorem 20 *Structure Theorem*

Let Λ:$B(H) \to B(H)$ be a unital quantum channel represented by the Kraus operators $A_1, ..., A_n$. Then the following are true

A) The interaction algebra $\mathcal{A} = Alg\{Ai\}$ is †-closed and depends only on Λ

B) $\mathcal{A}'$ is a †-closed algebra

C) $Fix(\Lambda) = \mathcal{A}'$

Proof: The theorem will be proved if C is shown. This is because if C is true then since $\mathcal{A}'$ is an algebra and $Fix(\Lambda)$ is †-closed B is true. Since $\mathcal{A} = (\mathcal{A}')'$ by VNDC and $Fix(\Lambda) = \mathcal{A}'$ we have $\mathcal{A} = Fix(\Lambda)'$. This implies that $\mathcal{A}$ is a function of Λ only, and hence $\mathcal{A}$ is independent of the choice of Kraus operators for its representation. $\mathcal{A}$ will †-closed since if $B \in \mathcal{A}$ and $C \in Fix(\Lambda) = \mathcal{A}'$,

$$[B, C] = 0 \Rightarrow BC - CB = 0 \Rightarrow [B^\dagger, C^\dagger] = 0 \Rightarrow B^\dagger \in \mathcal{A} \tag{C.16}$$

where the last implication holds since $Fix(\Lambda)$ is assumed to be $\dagger$-closed and so $B^\dagger$ commutes with every element in $\mathcal{A}'$. So by VNDC, $B^\dagger \in \mathcal{A}$. Thus, we only prove C.

First, let $T \in \mathcal{A}'$. Then,

$$\Lambda(T) = \sum_k A_k T A_k^\dagger = T \sum_k A_k A_k^\dagger = TI = T \tag{C.17}$$

and $T \in Fix(\Lambda)$. So, $\mathcal{A}' \subseteq Fix(\Lambda)$.

To prove the reverse inclusion we first note that if the set of fixed points only consists of scalar multiples of I, then since all scalar multiples of I are in $\mathcal{A}'$, $\mathcal{A}'$ $= Fix(\Lambda)$. So, suppose there exists a non-scalar operator S in $Fix(\Lambda)$. Every operator S can be broken up into its real and imaginary parts in the following manner

$$S = Re(S) + iIm(S) = \frac{1}{2}(S + S^\dagger) + i(\frac{1}{2i}(S - S^\dagger)) \tag{C.18}$$

Now, since $Fix(\Lambda)$ is $\dagger$-closed and each of $Re(S)$ and $Im(S)$ are self-adjoint, $Re(S)$ and $Im(S)$ are self-adjoint operators in $Fix(\Lambda)$. Moreover, since S is non-scalar, at least one of $Re(S)$ and $Im(S)$ is non-scalar. Thus without loss of generality, if $Fix(\Lambda)$ contains a non-scalar operator, we may choose the non-scalar operator to be self-adjoint. In fact, we can choose the non-scalar operator to be positive by the fact that if S is self-adjoint, $S + \|S\|$ is clearly positive and in $Fix(\Lambda)$. Thus, let S be a positive non-scalar operator in $Fix(\Lambda)$. Order the eigenvalues of S in an increasing sequence $\{\lambda_i\} \geq 0$ with

$$S = \lambda_1 P_1 + ... + \lambda_r P_r \tag{C.19}$$

Thus, $0 \leq S \leq \|S\| I = \lambda_r I$ where $\leq$ is the partial order defined on $B(H)$ by $C \leq D$ if $\forall\, |\in\rangle\, H$, $\langle C|\,|\psi\rangle \leq \langle\psi|D|\psi\rangle$. If $|\psi\rangle$ is in the eigenspace associated with λ_r so that $S|\psi\rangle = \lambda_r|\psi\rangle$ then since λ_r is the extremal eigenvalue for S,

$$\lambda_r\langle\psi||\psi\rangle = \sum_k \langle\psi|A_k S A_k^\dagger|\psi\rangle = \sum_k \langle A_k^\dagger\psi|S|A_k^\dagger\psi\rangle \leq \lambda_r \sum_k \langle A_k^\dagger\psi||A_k^\dagger\psi\rangle \tag{C.20}$$

Now, unitality and linearity of the channel gives,

$$\lambda_r \sum_k \langle A_k^\dagger \psi || A_k^\dagger \psi \rangle = \lambda_r \langle \psi || \psi \rangle \tag{C.21}$$

This implies that in fact $\langle A_k^\dagger \psi | S | A_k^\dagger \psi \rangle = \lambda_r \langle A_k^\dagger \psi | A_k^\dagger \psi \rangle$ and so $A_k^\dagger |\psi\rangle$ is in the eigenspace associated with λ_r, call it H_r for all k. Thus H_r is an invariant subspace for $A_k^\dagger$ and by Lemma 1, H_r is a reducing subspace for the algebra $\mathcal{A}$ and P_r is in $\mathcal{A}'$. So, $S - \lambda_r P_r = 0P_r + \lambda_1 P_1 + ... + \lambda_{r-1} P_{r-1} \in \mathcal{A}'$. This argument can clearly be iterated to obtain that each $P_i \in \mathcal{A}'$. Therefore $S \in \mathcal{A}'$. Now, for an arbitrary operator $T \in Fix(\Lambda)$, we can write T in terms of its real and imaginary parts, each of which is self-adjoint and also in $Fix(\Lambda)$. As well, each self-adjoint operator R produces a positive operator $R + \|R\|$ which is in $Fix(\Lambda)$. Therefore $R \in \mathcal{A}'$, which implies T is in $\mathcal{A}'$. So, $Fix(\Lambda) \subseteq \mathcal{A}'$ and we have $Fix(\Lambda) = \mathcal{A}'$.

C.3 The Structure of the Commutant: An Algorithmic Approach

As stated previously, the structure theorem along with the UE theorem form the basis for the algorithmic approach to finding noiseless subsystems. From the UE and structure theorems, the interaction algebra $\mathcal{A}$ for a unital quantum channel is a finite dimensional unital C^*-algebra and so there exists a basis for the Hilbert space in which the representation of $\mathcal{A}$ takes the form $(M_{n_1} \otimes I_{m_1}) \oplus ... \oplus (M_{n_d} \otimes I_{m_d})$. Moreover, the direct sum $(M_{n_1} \otimes I_{m_1}) \oplus ... \oplus (M_{n_d} \otimes I_{m_d})$ is unique up to ordering.

Definition 50 *Minimal Central Projections for $\mathcal{A}$*

A projection in $\mathcal{A}$ which corresponds to the identity on $(M_{n_k} \otimes I_{m_k})$ is called a minimal central projection for $\mathcal{A}$.

By the representation of $\mathcal{A}$ as above, the Hilbert space H decomposes as $(H_{n_1} \otimes H_{m_1}) \oplus ... \oplus (H_{n_d} \otimes H_{m_d})$. Choose one term in the direct sum, say $H_{n_k} \otimes H_{m_k}$. The space $H_{n_k} \otimes H_{m_k}$ can equivalently be thought of as housing (being a direct sum of) m_k orthogonal copies of H_{n_k}. These copies are such that when an operator in $\mathcal{A}$ acts on H, and an orthonormal basis is fixed for H_{n_k}, the same matrix in M_{n_k} acts on each of the copies of H_{n_k}. Moreover the particular H_{n_k} subspaces for which this is true are unique. Thus, these particular subspaces are linked by the fact that there is a redundancy in the action of M_{n_k} on these subspaces. Clearly, the projections onto one of these subspaces reduces $\mathcal{A}$, but will not be a minimal central projection for $\mathcal{A}$.

The projections onto one of the above mentioned subspaces are called minimal $\mathcal{A}$-reducing projections, and since the representation of $\mathcal{A}$ is unique up to ordering, the maximal family of minimal $\mathcal{A}$-reducing projections is unique. Since these projections reduce $\mathcal{A}$, they are in $\mathcal{A}'$, however they may not be in $\mathcal{A}$. The reason they may not be in $\mathcal{A}$ is that such a minimal $\mathcal{A}$-reducing projection onto one of the subspaces does not act redundantly on the m_k copies. Thus, such a projection will not be in $\mathcal{A}$, unless there is only one such copy, that is, $m_k = 1$ in the decomposition. As well, one can easily see that since $H_{n_k} \otimes H_{m_k}$ is comprised of a unique orthogonal sum of these minimal $\mathcal{A}$ reducing subspaces, the sum of the projections onto these subspaces is a minimal central projection for $\mathcal{A}$ corresponding to the projection onto $H_{n_k} \otimes H_{m_k}$. This discussion motivates the following definition.

Definition 51 *Linked Minimal $\mathcal{A}$-Reducing Projections*

Suppose $\{P_j\}$, $j \in S$, is the unique maximal family of minimal $\mathcal{A}$-reducing projections discussed above. A subset $\{Q_i\}$, with $i \in S_Q$ of this family is linked if the following are true

A) $Q = \sum_{i \in S_Q} Q_i$ is in $\mathcal{A}$

B) If $S_0 \subsetneq S_Q$ then $\sum_{i \in S_0} Q_i$ is not in $\mathcal{A}$

Thus, the sum of linked projections correspond exactly to minimal central projections for $\mathcal{A}$ and the cardinality of each S_k is m_k. Once the representation of $\mathcal{A}$ has been identified, the structure of $\mathcal{A}'$ is easily found by the fact that

$$\mathcal{A}' = (\oplus_k M_{n_k} \otimes I_{m_k})' = \oplus_k (M_{n_k} \otimes I_{m_k})' = \oplus_k (I_{n_k} \otimes M_{m_k}) \qquad \text{(C.22)}$$

Thus, noiseless subsystems will be associated with the M_{m_k} and these subsystems represent the possibility of encoding quantum information. We are now ready to construct the algorithm for finding noiseless subsystems. The only assumption that is required is that the linear structure of $\mathcal{A}'$ is already known. There are two broad parts to the algorithm.

Part 1: Obtain the unique maximal family of minimal A-reducing projections in $\mathcal{A}'$

Part 2: Obtain the minimal central projections for $\mathcal{A}$ by computing all linked projections in the unique maximal family from Part 1.

C.3.1 Part 1

If $\mathcal{A}'$ contains only scalar multiples of the identity, then the identity operator is the only minimal central projection for $\mathcal{A}$. That is, $\mathcal{A}$ is unitarily equivalent to a full matrix algebra on the entire Hilbert space. This is easily seen to be equivalent to the fact that there are no non-trivial reducing subspaces of $\mathcal{A}$. So, suppose $\mathcal{A}'$ contains non-trivial operators. As shown in the proof of the structure theorem, this implies that there exists a self-adjoint operator, T, in $\mathcal{A}'$. We will use the spectral projections for T as the first step in Part 1. Note that if an eigenvalue λ of T has degeneracy larger than 1, then the spectral projection associated to λ corresponds to the projection onto the entire eigenspace for λ.

Lemma 3 *For self-adjoint $T \in B(H)$, TFAE*

A) $\Lambda(T) = T$

B) If P is a spectral projection of T then $\Lambda(P) = P$

C) If P is a spectral projection of T then $P \in \mathcal{A}'$

Proof: The proof is straightforward. The equivalence of B and C follows directly from the structure theorem. $B \Rightarrow A$ follows from linearity of the unital channel. $A \Rightarrow B$ will follow by supposing R commutes with T and showing R commutes with P. Indeed, let $\{\lambda_i | i = 1, ..., m\}$ consist of the distinct eigenvalues of T. Then $TR = RT$ if and only if

$$R(\lambda_1 P_1 + ... + \lambda_m P_m) = (\lambda_1 P_1 + ... + \lambda_m P_m)R \tag{C.23}$$

or

$$\lambda_1 [R, P_1] + ... + \lambda_m [R, P_m] = 0 \tag{C.24}$$

Since the λ_i were assumed to be distinct, they are linearly independent variables. Thus, each of the $[R, P_i]$ must be zero.

If P is a spectral projection for T and $P_0 < P$ then the proof of A $\Rightarrow$ B above shows it may not be the case that P_0 reduces $\mathcal{A}$.

By the above lemma, if $\{P_j\}$ are the spectral projections of T then for each j we can define a unital channel Λ_j from P_jH into P_jH of the form

$$\Lambda_j(T) = \sum_k A_{k,j} T A_{k,j}^{\dagger} \tag{C.25}$$

with $A_{k,j} = A_k|_{P_jH}$. These channels are clearly unital, with the identity given by $I|_{P_jH}$, and they are trace preserving. Thus Λ_j is a unital quantum channel with underlying Hilbert space corresponding to P_jH. The following proposition will give the result required to finish Part 1 of the algorithm

Proposition 15 *Suppose $\Lambda : B(H) \to B(H)$ is a unital channel with interaction algebra $\mathcal{A}$ determined by Kraus operators A_i. Let P be a projection which reduces $\mathcal{A}$. Then the mapping $\Lambda_P : B(H) \to B(H)$ given by $\Lambda_P(T) = \sum_k A_{k,P} T A_{k,P}^{\dagger}$ is a unital quantum channel such that*

$$Fix(\Lambda_P) = (\mathcal{A}|_{PH})' = P\mathcal{A}'|_{PH} \tag{C.26}$$

and P is a minimal $\mathcal{A}$-reducing projection if and only if $Fix(\Lambda_P) = P\mathcal{A}'|_{PH} = \mathbb{C}I|_{PH}$.

Proof: By the discussion above Λ_P is a unital channel. We introduce the notation $P\mathcal{A}'|_{PH} = P\mathcal{A}'P$ and $(\mathcal{A}|_{PH})' = (\mathcal{A}P)'$. This is done because $P\mathcal{A}'|_{PH}$ is just a set of operators from PH into PH and each function in this set may be extended in its domain to all of H by defining it to be zero on all subspaces orthogonal to PH. That is, $P\mathcal{A}'|_{PH}$ is associated to

$$P\mathcal{A}'P = \begin{bmatrix} P\mathcal{A}'|_{PH} & 0 \\ 0 & 0 \end{bmatrix}, \tag{C.27}$$

Similarly, $\mathcal{A}|_{PH}$ is associated with

$$\mathcal{A}P = \begin{bmatrix} A|_{PH} & 0 \\ 0 & 0 \end{bmatrix}, \tag{C.28}$$

which gives the correspondence $(\mathcal{A}|_{PH})' = (\mathcal{A}P)'$.

Now, suppose $T \in \mathcal{A}'$ so that $PT|_{PH} \in P\mathcal{A}'|_{PH} = P\mathcal{A}'P$. Then,

$$PT|_{PH}A_i|_{PH} = PA_i|_{PH}T|_{PH} = A_i|_{PH}PT|_{PH} \tag{C.29}$$

Thus, $P\mathcal{A}'P \subseteq (\mathcal{A}P)'$. Note that $PT|_{PH}A_i|_{PH} = A_i|_{PH}PT|_{PH}$ if and only if $(PTP)(A_iP) = (A_iP)(PTP)$, which is clear from the discussion following our definition of notation. Now, suppose $T \in (\mathcal{A}P)'$. Then, T may be associated with an operator over the entire Hilbert space H of the form $T = PTP$. That is, PTP has the form of T from PH to PH and maps all subspaces orthogonal to PH to 0. This gives,

$$TA_i = PTPA_i = TPA_i = T(A_iP) = (A_iP)T = A_iT \tag{C.30}$$

Thus, $T \in P\mathcal{A}'P$ and so $(\mathcal{A}P)' \subseteq P\mathcal{A}'P$. Therefore $(\mathcal{A}P)' = P\mathcal{A}'P$ and $Fix(\Lambda_P) = P\mathcal{A}'P$. Finally, P is a minimal $\mathcal{A}$-reducing projection if and only if there are no projections which reduce the restricted noise algebra $A_{i,P}$. However, this is equivalent to $Fix(\Lambda_P) = \mathbb{C}I|_{PH}$, or by the above, $P\mathcal{A}'P = \mathbb{C}P$.

If we use the spectral projections P_j of self-adjoint T in $\mathcal{A}'$, then the above proposition gives that P_j is a minimal $\mathcal{A}$ reducing projection if and only if $P_j\mathcal{A}'|_{P_jH} = \mathbb{C}I|_{P_jH}$. Or, equivalently, $P_j\mathcal{A}'P_j = \mathbb{C}P_j$. So, since we have assumed we know the linear structure of $\mathcal{A}'$, we take a basis B_k of this linear space and compute $P_jB_kP_j$. If for each k, this is proportional to P_j then P_j is in the maximal family of minimal $\mathcal{A}$ reducing projections we seek. If there exists a B_m such that this is not true, then there are smaller projections that reduce the original interaction algebra. They can be found by iterating the above process. This is done by noting that Λ_j is a unital channel with non-scalar fixed point $P_jB_mP_j$. This operator will give us, through its' real and imaginary parts, a self-adjoint operator whose spectral projections define unital channels on subspaces of P_jH. Since H is finite-dimensional, this process will terminate at a minimal $\mathcal{A}$ reducing projection. This concludes Part 1.

C.3.2 Part 2

Suppose the maximal family of minimal $\mathcal{A}$ reducing projections, $\{P_j\}$, has been found from Part 1. It is clear that a sub-family of $\{P_j\}$ is linked only if the projections in the sub-family are of the same rank. Thus for all positive natural numbers k let $\{P_{j,k}\}$ be the subfamily of $\{P_j\}$ whose elements are of rank k. Let S_k be the index set for $\{P_{j,k}\}$. Now, by Von Neumann's Double Commutant Theorem $A =(\mathcal{A}')'$ and so if $\{B_i\}$ is a basis for $\mathcal{A}'$, we have that an element of A must commute with each of the B_i. This gives a systematic method to find the linked projections in $\{P_{j,k}\}$. The method is to let $S \subseteq S_k$ and take P_S to be the sum of all of the projectors indexed by elements in S. Then, P_S is a minimal central projection for $\mathcal{A}$ if and only if $[P_S, B_i] = 0 \forall i$ and there is no proper subset of S for which this is true. Again, this searching type algorithm will terminate in finite time and is not in NP.

The rank one projections are easier to test for links than higher rank projections by the following theorem which we present without proof [19].

Theorem 21 *Let $\{P_{j,1}\} := \{R_k\}$ be the set of rank one projectors in the maximal family obtained from Part 1 and suppose $R_k = |\psi_k\rangle\langle\psi_k|$. Suppose the interaction algebra for the unital quantum channel Λ is generated by $\{A_i\}$, $i \in \{1, ..., n\}$ and define the n-tuple $(\lambda_{1k}, ..., \lambda_{nk})$ by*

$$A_i|\psi_k\rangle = \lambda_{ik}|\psi_k\rangle \tag{C.31}$$

This defines a mapping from $\{R_k\}$ into $\mathbb{C}^n$ given by

$$f(R_k) = (\lambda_{1k}, ..., \lambda_{nk}) \tag{C.32}$$

Suppose $\lambda \in \mathbb{C}^n$ is such that $f^{-1}(\lambda) \neq \emptyset$. Then $f^{-1}(\lambda)$ forms a linked set of projections.

What this means is that we can group together all of the rank one projections with the same eigenvalues under the A_i and the resulting partition will correspond to minimal central projections for $\mathcal{A}$. Before we summarize the algorithm just constructed, it is interesting to note that for the case of an abelian algebra, the maximal family obtained from part 1 will only consist of rank one projectors [19].

C.3.3 Summary of the Algorithm

Assumption: The linear structure of $\mathcal{A}'$ is known

- If $\mathcal{A}' = \mathbb{C}I$ then we are done and the identity is the only minimal central projection in $\mathcal{A}$. If $\mathcal{A}' \neq \mathbb{C}I$ then choose a non-scalar, self-adjoint $T \in \mathcal{A}'$, and compute the corresponding spectral projections
- For each P_j, compute $P_j B_i P_j$ for a basis $\{B_i\}$ of $\mathcal{A}'$. If for every i, the result is proportional to P_j, then P_j is in the desired maximal family. If this is not the case then the channel Λ_j has fixed points that are non-scalar. This brings us back to the start of the algorithm, and we repeat this process until we find the desired maximal family

- Assuming the maximal family has been found, group together projectors of the same rank. A subset of one of the groups is linked if the sum of all projectors in the subset commutes with each B_i, and this is not true for a smaller sum of these projectors

- The projectors of rank one may be subdivided by their eigenvalues under the A_i. Each class in this subdivision corresponds to a linked set of projectors

- Once the links have been found, the minimal central projections will be determined. Thus, the spatial location of each block $M_{n_k} \otimes I_{m_k}$ in the representation of $\mathcal{A}$ will be apparent and will correspond to the subalgebra AP where P is a minimal central projection. The factor of m_k will correspond to the number of linked (redundant) minimal $\mathcal{A}$-reducing projections which form P. Moreover, each minimal $\mathcal{A}$-reducing projection will correspond to a projection onto a matrix block M_{n_k}. The structure of the commutant is now easily determined and has blocks $I_{n_k} \otimes M_{m_k}$ and will give the possible noiseless subsystems we are in search of

The algorithm presented here involves parts where searches must be made which will not run in polynomial time. Therefore, the algorithm will become virtually intractable for very large Hilbert spaces.

Appendix D

List of Abbreviations

Table D.1: List of Abbreviations Used.

Abbreviation	Full Name
POVM	Positive Operator Valued Measure
QEC	Quantum Error Correction
PIP Channel	Permutation Invariant Pauli Channel
UCS	Unitarily Correctable Subsystem
UNS	Unitarily Noiseless Subsystem
CDFT	Classical Discrete Fourier Transform
QFT	Quantum Fourier Transform
VNDC	Von Neumann Double Commutant Theorem

References

[1] S. Aaronson and D. Gottesman. Improved simulation of stabilizer circuits. *Physical Review A*, 70(052328), 2004. 41

[2] W. Arveson. *An Invitation to C^*-Algebra.* Springer-Verlag, 1976. 77, 79, 82

[3] Ingemar Bengtsson and Karol Zyczkowski. *Geometry of Quantum States.* Cambridge University Press, Cambridge, UK, 2006. 25

[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), 1996. 23

[5] H.P. Breuer and F. Petruccione. *The Theory of Open Quantum Systems.* Oxford University Press, 2002. 48

[6] Andrew M. Bruckner, Judith B. Bruckner, and Brian S. Thomson. *Real Analysis.* Prentice-Hall, 1997. 70

[7] K. H. Cheung and M. Mosca. Decomposing finite abelian groups. *Quantum Information and Computation*, 1(2), 2001. 66

[8] M.D. Choi and D.W. Kribs. A method to find quantum noiseless subsystems. *Physical Review Letters*, 96(050501), 2006. 18, 39

[9] Christoph Dankert, Richard Cleve, Joseph Emerson, and Etera Livine. Exact and approximate unitary 2-designs: Constructions and applications. arXiv:quant-ph/0606161, June 2006. 20, 21, 22, 23, 31, 33, 39

[10] Kenneth R. Davidson. *C^* - Algebras by Example.* Fields Institute Monographs 6, 1996. AMS. 79

[11] P. Delsarte, J.M. Goethals, and J.J. Seidel. Spherical codes and designs. *Geometriae Dedicata: Constructions and Applications*, 6, 1997. 19

[12] Joseph Emerson, Robert Alicki, and Karol Zyczkowski. Scalable noise estimation with random unitary operators. *Journal of Optics B: Quantum and Semiclassical Optics*, 7(10), 2005. 24, 25, 26, 28

[13] Joseph Emerson, Marcus Silva, Osama Moussa, Colm Ryan, Martin Laforest, Jonathan Baugh, David G. Cory, and Raymond Laflamme. Symmetrized characterization of noisy quantum processes. *Science*, 317(5846), 2007. 32, 36, 51

[14] Christopher A. Fuchs. *Distinguishability and Accessible Information in Quantum Theory.* PhD thesis, University of New Mexico, Albuquerque, New Mexico, 2006. Preprint quant-ph 9601020. 7, 24

[15] M. Gromov and V.D. Milman. A topological application of the isoperimetric inequality. *American Journal of Mathematics*, 105(4), 1983. 71

[16] D. Gross, D. Audenaert, and J. Eisert. Evenly distributed unitaries: On the structure of unitary designs. *J Math Phys*, 48, 2007. 50

[17] L.K. Grover. A fast quantum mechanical algorithm for database search. In *28th Annual ACM Symposium on the Theory of Computing*, 1996. 1

[18] I.N. Herstein. *Topics In Algebra.* John Wiley and Sons, Inc., 2004. 77, 84

[19] J.A. Holbrook, D.W. Kribs, and R. Laflamme. Noiseless subsystems and the structure of the commutant in quantum error correction. *Quantum Information Processing*, 2, 2004. 18, 41, 84, 92

[20] J.A. Holbrook, D.W. Kribs, R. Laflamme, and D. Poulin. Noiseless subsystems for collective rotation channels in quantum information theory. *Integral Equations and Operator Theory*, 15, 2005. 39

[21] Richard Jozsa. Quantum algorithms and the fourier transform. arXiv:quant-ph/9707033, 1997. 13

[22] Richard Jozsa. Quantum factoring, discrete logarithms, and the hidden subgroup problem. *Computing in Science and Engineering*, 03(2), 2001. 13

[23] E. Knill. Protected realizations of quantum information. *Physical Review A*, 74(042301), 2006. 18

[24] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland. Randomized benchmarking of quantum gates. *Physical Review A*, 77(012307), 2008. 28, 30

[25] D.W. Kribs, R. Laflamme, and Poulin. A unified and generalized approach to quantum error correction. *Physical Review Letters*, 94(180501), 2005. 14, 16

[26] D.W. Kribs and R.W. Spekkens. Quantum error correcting subsystems are unitarily recoverable subsystems. *Physical Review A*, 74(042329), 2006. 17, 40, 41

[27] Michel Ledoux. *The Concentration of Measure Phenomenon.* American Mathematical Society, 2001. 71, 76

[28] John M. Lee. *Introduction to Smooth Manifolds.* Springer, 2002. 70

[29] S. Lloyd. Universal quantum simulators. *Science*, 273(5278), 1996. 1

[30] C. Lomont. The hidden subgroup problem - review and open problems. arXiv:quant-ph/0411037, 2006. 58, 63, 64, 67

[31] Vitali D. Milman and Gideon Schechtman. *Asymptotic Theory of Finite Dimensional Normed Spaces.* Springer-Verlag, 1980. Lecture Notes in Mathematics-1200. 71, 74, 75, 76

[32] James R. Munkres. *Topology.* Prentice Hall Inc., second edition edition, 2000. 69

[33] W.K. Nicholson. *Introduction to Abstract Algebra.* John Wiley and Sons, Inc., 1999. 77, 82

[34] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Information.* Cambridge University Press, Cambridge, UK, 2000. 3, 8, 9, 13, 15, 16

[35] Roman Orus and Rolf Tarrach. Weakly entangled states are dense and robust. *Physical Review A*, 70(050101), 2004. 10

[36] G.K. Pedersen. *Analysis Now.* Springer-Verlag, 1989. 85

[37] Vladimir Pestov. On the geometry of similarity search: Dimensionality curse and concentration of measure. *Information Processing Letters*, 73, 2000. 71, 76

[38] J.J. Rotman. *Advanced Modern Algebra.* Pearson Education Inc., 2006. 59, 82

[39] P.D. Seymour and T. Zaslavsky. Averaging sets: A generalization of mean values and spherical designs. *Advances in Mathematics*, 52, 1984. 20

[40] Peter W. Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In *Proceedings of the 35'th Annual Symposium on Foundations of Computer Science (FOCS)*, Los Alamitos, CA, 1994. IEEE Press. 1

[41] M. Silva, E. Magesan, D.W. Kribs, and J. Emerson. Scalable protocol for identification of correctable codes. *Accepted for publication in Physical Review A*. arXiv:quant-ph/0710.1900. 31