MATHEMATICAL FOUNDATIONS OF ADAPTIVE QUANTUM PROCESSING

by

DANIEL BONIOR
B.S. Middle Tennessee State University, 2013

A dissertation submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
in the Department of Physics
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Fall Term
2018

Major Professor: Eduardo Mucciolo

ABSTRACT

Quantum information has the potential to revolutionize the way we store, process, transfer and acquire information [1,14,15,21,37]. In particular, quantum information offers exciting new approaches to secure communication, computation and sensing. However, in order to realize such technologies, we must first understand the effect that environmental noise has on a quantum system. This dissertation builds upon recent studies that have explored the underlying structure of quantum information and the effects of qubit channels in quantum communication protocols.

This work is divided into five main chapters, with Chapter 1 being a brief introduction to quantum information. We then begin Chapter 2 by defining the error function for our qubit communication protocols. From there we explore the properties of our error functions and the topological space that they form.

In Chapter 3 we consider the newly patented process Adaptive Quantum Information Processing, patent number US9838141 B2; originally outlined by Martin in [23]. We restate the adaptive scheme and exemplify its application through the Prepare and Send Protocol and Quantum Key Distribution. Applying our results from Chapter 2, we obtain an expression for the adaptability of unital channels in these two protocols and classify the channels that admit the most improvement. We dedicate Chapter 4 to the derivation of gravitational noise, and show that in certain circumstances gravity results in a channel that can be maximally improved in Adaptive QKD [3,14,16].

Lastly, we study the set of error functions through the lens of domain theory. Domain theory is a subset of mathematics that was developed in order to rigorously formalize computations. The first four chapters are all consequences of past discoveries in the mathematical structure of quantum channels. In Chapter 5 we characterize the set of error functions through domain theory, extending the mathematical foundations of quantum information. [12,18,20, 22, 23,25].

Dedicated to my Mom and Dad:

this is the culmination of all the patience, time, and effort you put into me.

ACKNOWLEDGMENTS

I can only lay claim to this work in regards to its physical construction. In terms of spirit, inspiration, and all fundamental events that permitted me to work in this area, this dissertation does not belong to me. I was given the unique and undeserved opportunity to write this work due to the overwhelming understanding, patience, and cooperation of Keye Martin and Eduardo Mucciolo. These two men not only provided me with the guidance and encouragement to take on this endeavor, but the chance to do so under the unusual means that allowed me to forfeit the clichéd lifestyle of a starving graduate student.

Throughout my graduate experience I developed as a researcher but more so as a person. For all the guidance, advice, and introductions to endless amounts of art that have not only served as a safety net but also a source of inspiration, I am forever indebted to Keye Martin. I would like to acknowledge the efforts of everyone in the informatic phenomena group at the Naval Research Laboratory. Specifically, Tanner Crowder for never allowing me to sidestep a question, constantly pushing me to analyze my own intuition, and helping me to develop my writing skills. I am indebted to Marco Lanzagorta for his helpful discussions and ideas that are not only present in this dissertation, but wholly define my undergraduate thesis. As a friend and prior teacher, I am grateful for all the efforts of Johny Feng.

Expressing my gratitude towards my family borders on a vacuous sentiment, for nothing can encapsulate twenty-eight years of financial and emotional support. To this end, all my accomplishments are due to them; and I offer my condolences for the mountain of shortcomings this includes. Both during and prior to this work, I always had comfort and encouragement. For that, I am truly thankful for all my friends, family, and ultimately Erica Ellison.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION TO QUANTUM INFORMATION

The content of this dissertation lies in the intersection of mathematics and physics, and is therefore written with two audiences in mind: mathematicians and physicists. As such, we will be providing definitions, proofs, and exposition for some material, while other topics are assumed to be known. This is because we are presuming the reader is familiar with the concepts encountered in the basic curricula of both subjects, and we will try to define anything that is sometimes missed in either field. With this in mind, we begin with a preliminary introduction to some of the basics of linear algebra and the notation that will be used throughout this work. We note that since this dissertation will include many theorems, some original and some already proven by others, we will include the name of the original author after the numbering of all previous results.

1.1 Mathematical Background

Let us first establish some essential notation. Throughout this work both real and complex numbers will appear and context will provide an unambiguous indication as to which number system is being used. That being said, if $\alpha = a + ib$ is a complex number, then we denote its conjugate by $\overline{\alpha} = a - ib$. If M is a $m \times n$ dimensional matrix with complex entries, we denote its transpose by M^t, its conjugate transpose by $M^\dagger$, and note that in the case where its entries are purely real numbers, $M^t = M^\dagger$. Further, if M is a square matrix, we denote its trace by $\text{tr}(M)$ and its determinant by $\det(M)$. Lastly, we denote the identity matrix for all dimensions by I.

1.1.1 The State Space

This dissertation will be limited to the study of two-level quantum systems, which we will often refer to as qubits. In order to discuss these dynamic systems, we must first define the space used to represent their states. We assume the reader is familiar with the definitions of an *inner product*, *norm*, and the n-dimensional complex vector space, $\mathbb{C}^n$. For the case of $\mathbb{C}^n$, we fix once and for all the following notation:

(1) The inner product $\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$, is $\langle x, y \rangle = x^\dagger y$.

(2) The norm $|| \cdot || : \mathbb{C}^n \to \mathbb{R}$, is taken to be $||x|| = \sqrt{\langle x, y \rangle}$.

(3) The distance function $d : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{R}$, is defined such that $d(x, y) = \sqrt{||x - y||}$.

The function d, beyond being non-negative, is also symmetric and satisfies the triangle inequality and the identity of indiscernibles. When a function satisfies these four properties, we say it is a *metric*, and we refer to the pair $(\mathbb{C}^n, d)$ as a *metric space*. Furthermore, as d is directly defined by the norm $|| \cdot ||$, and therefore our inner product, we refer to it as the *induced metric*. We adopt the same notation when considering the vector space $\mathbb{R}^n$, where all simplifications associated with a change from complex to real numbers are applied. With a metric, we may characterize two particular classes of sequences of complex numbers $x_1, x_2, x_3, \ldots$, which we denote by (x_n).

Definition 1.1.1: Let (x_n) be a sequence in $(\mathbb{C}^n, d)$. If there exists a $x \in \mathbb{C}^n$ where for any positive real number ϵ we can find some positive integer N such that $m \geq N$ implies $d(x, x_m) < \epsilon$, then we say (x_n) *converges to* x. In such a case we write $(x_n) \to x$ where x is called the *limit* of (x_n) and is denoted by $\lim_{n \to \infty} x_n = x$.

Definition 1.1.2: For the metric space $(\mathbb{C}^n, d)$ a sequence (x_n) is *Cauchy* if for any real number $\epsilon > 0$, there is an integer $N > 0$ such that for all integers $n, m \geq N$, $d(x_m, x_n) < \epsilon$. We say

a metric is *complete* if every Cauchy sequence in the space converges. With these definitions, we can now finally define the state space of a quantum system.

Definition 1.1.3: A *Hilbert space* is a complex inner product space $\mathcal{H}$ with a complete induced metric.

The state space of a qubit, being a two-level quantum system, is characterized by a two-dimensional Hilbert space. That is, when we speak of the state space of a qubit, we are referring to the space $(\mathbb{C}^2, d)$, which we denote by $\mathcal{H}^2$. When two subsystems interact they are considered to form a larger composite system which we describe with a *tensor product*.

If S and T are vector spaces, then their *tensor product* is another vector space $S \otimes T$ with a bilinear map $\otimes : S \times T \to S \otimes T$ such that for any bilinear map $b : S \times T \to Z$ there is a unique linear map $\bar{b} : S \otimes T \to Z$ such that $\bar{b}(s \otimes t) = b(s, t)$. This condition is known as the universal property of the tensor product and is illustrated in Fig. 1.1. The tensor product is unique up to an isomorphism, meaning that any other vector space X satisfying the universal property is isomorphic to $S \otimes T$. That is, there exists a bijective linear function between X and $S \otimes T$.

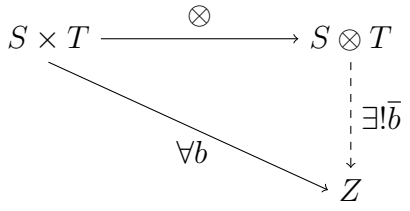

Figure 1.1: Diagram depicting the universal property of the tensor product.

With the state space defined, we introduce a few more definitions and some notation before delving into the mathematics needed to discuss operations on a quantum system.

Notation 1.1.4: Developed by physicist Paul Dirac, *Bra-ket* notation attempts to make calculations more "intuitive". The vectors of a Hilbert space are denoted by $|x\rangle$, pronounced "ket x"; while the symbol $\langle x|$, "bra x", denotes its conjugate transpose. With this notation in mind, the inner product for our state space $\mathcal{H}$ is then denoted by a "bra acting on a ket", $\langle x|x\rangle$.

The *outer product* of vectors $|x\rangle, |y\rangle \in \mathbb{C}^n$ is

$$|x\rangle\langle y| = \begin{bmatrix} x_1y_1 & x_1y_2 & \dots & x_1y_n \\ x_2y_1 & x_2y_2 & \dots & x_2y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_my_1 & x_my_2 & \dots & x_my_n \end{bmatrix}. \tag{1.1}$$

For example, in the case of a two-level quantum system, if $\{|e_1\rangle, |e_2\rangle\}$ is an orthonormal basis for the state space, then each $|x\rangle \in \mathcal{H}^2$ can be uniquely written as $x_1|e_1\rangle + x_2|e_2\rangle$ where x_1 and x_2 are complex numbers. Then

$$\Big(\sum_{i=1}^{2} |e_i\rangle\langle e_i|\Big)|x\rangle = \sum_{i=1}^{2} |e_i\rangle\Big(\langle e_i|x\rangle\Big) = \sum_{i=1}^{2} x_i|e_i\rangle = |x\rangle. \tag{1.2}$$

Therefore, $\sum_{i=1}^{2} |e_i\rangle\langle e_i| = I$; this equality is called the *completeness relation*. Lastly, we wish to establish the following definition, as it will be appear frequently.

Definition 1.1.5: Let T be a vector space. A subset $C \subseteq T$ is said to be *convex* if for all $x, y \in C$ and all $t \in (0, 1)$, then $tx + (1-t)y \in C$. Furthermore, we say that x is an *extreme point* of C if $x = py + (1-p)z$ with $y, z \in C$ and $p \in (0, 1)$ implies that $x = y = z$.

Remark: Convex sets can be generalized to more abstract spaces; however, the results of this work will only pertain to convexity for vector spaces. Therefore, Definition 1.1.5 is sufficient.

While this concludes the necessary background for the state of a quantum system, we also require an understanding of how these systems change. Therefore, we provide a brief introduction to some of the basic concepts in algebra and operator theory that are essential to describe the evolution of qubits.

1.1.2 Operations on Quantum Systems

Many times throughout this work our results will depend on the properties of particular collections of mappings. One such property is that of forming a *group*. A set S together with a binary operation $* : S \times S \to S$ is a *group* if the following hold for all $a, b, c \in S$:

(i) $a * (b * c) = (a * b) * c$,

(ii) $(\exists e \in S)(\forall a \in S)(e * a = a * e = a)$,

(iii) $(\forall a \in S)(\exists a^{-1} \in S)(a * a^{-1} = a^{-1} * a = e)$.

If only properties (i) and (ii) hold, we say that $(S, *)$ is a *monoid*.

Specifically, there are two groups that will be essential in our study of quantum information. The first being the *special orthogonal group*, denoted by $SO(3)$. This group consists of the set of all rotations about the origin in $\mathbb{R}^3$. In other words, $SO(3)$ is the set of 3×3 matrices r, such that $\det(r) = 1$ and $r^t = r^{-1}$. We also often make use of the *Klein* 4*-group* V, which is the *involutive* group of order four. We say that an element of a group a is *involutive* if $a * a = e$. Explicitly, $V = \{e, a, b, c\}$ where e is the identity element and $a * a = b * b = c * c = e$, while $a * b = c$, $b * c = a$, and $a * c = b$. With two examples of a group given, we now move on to define particular sets of mappings between state spaces.

Definition 1.1.6: For us *linear operators* describe the transformations of the state of a qubit. We

say $E : \mathcal{H} \to \mathcal{H}$ is a *linear operator* if for all $x, y \in \mathcal{H}$ and $\alpha, \beta \in \mathbb{C}$,

$$E(\alpha x + \beta y) = \alpha E(x) + \beta E(y). \tag{1.3}$$

Furthermore, E is *self-adjoint* if $x^\dagger(Ex) = (Ex)^\dagger x$, and *unitary* if $EE^\dagger = E^\dagger E = I$. Note, every element of $SO(3)$ is a linear unitary operator, while each involutive rotation is self-adjoint.

Definition 1.1.7: Let $E : \mathcal{H} \to \mathcal{H}$ be a linear operator. A non-zero vector $x \in \mathcal{H}$ is an *eigenvector* of E if $Ex = \lambda x$, where $\lambda \in \mathbb{C}$ is called the *eigenvalue* of E associated with x. The set of eigenvalues of E is called its *spectrum*, and is denoted by $\sigma(E)$.

Lastly, we end this section with two definitions that are essential for this work.

Definition 1.1.8: The n*-simplex* is

$$\Delta^n = \left\{ p \in [0,1]^n \;\middle|\; \sum_{i=1}^{n} p_i = 1 \right\}. \tag{1.4}$$

Two examples of such would be a single point Δ^1, and a line segment Δ^2 of unit length.

Definition 1.1.9: The *unit ball* in $\mathbb{R}^3$ is

$$\mathbb{B}^3 = \left\{ x \in \mathbb{R}^3 \;\middle|\; \|x\| \leq 1 \right\}, \tag{1.5}$$

where its boundary is given by the *two-sphere*

$$S^2 = \left\{ x \in \mathbb{R}^3 \;\middle|\; \|x\| = 1 \right\}. \tag{1.6}$$

Although we have not covered all the necessary mathematics for this work in its entirety, we have

provided enough to begin our discussion on the foundations of quantum mechanics. As such, we shift our focus to physics for the duration of the next section. We define and discuss all other mathematical concepts as they arise throughout this work.

1.2 Postulates of Quantum Mechanics

Quantum mechanics is a mathematical model used by physicists to describe and make predictions in nature. Like every theory, it is built upon a set of axiomatic statements, or postulates. We will briefly state the postulates, and then individually discuss them in the following two sections.

Postulate 1: An isolated physical system can be represented by a unit vector in some Hilbert space; such a system is said to be in a *pure state*. This vector is called the *state vector*, and the Hilbert space is called the *state space*.

Postulate 2: The closed evolution of a system is described by unitary transformations. In other words, if $|\psi_0\rangle$ is the initial state of a closed system at time t_0 and $|\psi\rangle$ denotes the state at some time t, then $|\psi\rangle = U|\psi_0\rangle$, where U is unitary and depends only on the times t_0 and t.

Postulate 3: Measurements are described by collections of operators $\{M_m\}$ acting on $\mathcal{H}$ such that $\sum_m M_m^\dagger M_m = I$ and $p(m|\psi) = \langle\psi|M_m^\dagger M_m|\psi\rangle$, where m denotes a possible outcome and $p(m|\psi)$ is the probability of the state $|\psi\rangle$ resulting in the outcome m.

Postulate 4: The state space of a composite system is given by the tensor product of the individual state spaces. If $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces for two quantum systems, then $\mathcal{H}_1 \otimes \mathcal{H}_2$ is the Hilbert space for the combined system.

There is a rather bothersome implication that follows from the first postulate in conjunction with the third. Every unit vector of the form $e^{i\theta}|\psi\rangle$ is equivalent in the sense that they all make the same

predictions. That is, $|\psi\rangle$ and $e^{i\theta}|\psi\rangle$ result in the same probabilities in the event of a measurement. This is commonly stated as "states are equal up to a global phase". Explicitly,

$$p(m|e^{i\theta}\psi) = e^{i\theta}e^{-i\theta}\langle\psi|M_m^\dagger M_m|\psi\rangle = \langle\psi|M_m^\dagger M_m|\psi\rangle = p(m|\psi). \tag{1.7}$$

However, it would be mathematically inconsistent to say things like "$|\psi\rangle = -|\psi\rangle$", since this equality is false for all unit vectors. One way to resolve this issue is to associate with each unit vector $|\psi\rangle$, a linear operator $\rho : \mathcal{H} \to \mathcal{H}$, where

$$\rho = |\psi\rangle\langle\psi| = e^{i\theta}e^{-i\theta}|\psi\rangle\langle\psi|. \tag{1.8}$$

Then all pure states of the form $e^{i\theta}|\psi\rangle$ are instead represented by the operator ρ, and we now have a mathematical equivalence that is consistent with the third postulate.

Definition 1.2.1: A *density operator* $\rho : \mathcal{H} \to \mathcal{H}$ is a trace one, self-adjoint, positive semi-definite (i.e. all eigenvalues are non-negative) linear operator.

The density operator also provides an intuitive means of constructing a quantum state distribution. In classical theory, each distribution $x \in \Delta^2$ may be written as

$$x = p_1 \cdot e_1 + p_2 \cdot e_2, \tag{1.9}$$

where $p_1 + p_2 = 1$. The physical meaning of Eq. (1.9) is that the classical system is either in state e_1 or e_2 with respective probabilities p_1 and p_2. Similarly, if a quantum system is in state $\rho_i = |\psi_i\rangle\langle\psi_i|$ with probability p_i, the natural way to represent this distribution is with the operator

$$\rho = \sum_{i=1}^{n} p_i|\psi_i\rangle\langle\psi_i|, \tag{1.10}$$

where $p \in \Delta^n$. Note, if a density operator has the form of Eq. (1.8), then it is an *idempotent* and conversely. That is, ρ is a pure state if and only if $\rho^2 = \rho$. With our new representation of a state, we revisit and revise our postulates.

Postulate 1*: The state of an isolated system is $\rho : \mathcal{H} \to \mathcal{H}$. In particular if the system is in the state $|\psi_i\rangle$ with probability p_i, then $\rho = \sum_{i=1}^{n} p_i |\psi_i\rangle\langle\psi_i|$, where $\mathcal{H}$ is the state space of the system, $|\psi_i\rangle \in \mathcal{H}$, and $p \in \Delta^n$. We denote the set of n-dimensional quantum states by Ω^n.

Postulate 2*: The evolution of a closed system through time is described by a unitary transformation. If ρ_0 is the initial state of a closed system at time t_0, and ρ denotes the state at some time t, then $\rho = U\rho_0 U^\dagger$, where U is unitary and depends only on the times t_0 and t.

Postulate 3*: Measurements are described by collections of operators $\{M_m\}$ acting on $\mathcal{H}$ such that $\sum_m M_m^\dagger M_m = I$ and $p(m|\rho) = \text{tr}\big(M_m^\dagger M_m \rho\big)$, where m denotes a possible outcome and $p(m|\psi)$ is the probability of ρ resulting in the outcome m when the system is in the state $|\psi\rangle$.

Postulate 4*: The state space of a composite system is given by the tensor product of the individual state spaces. Letting $\mathcal{H}_1$ and $\mathcal{H}_2$ be the Hilbert spaces for two quantum systems, then $\mathcal{H}_1 \otimes \mathcal{H}_2$ is the Hilbert space for the combined system.

Remark: The density operator formalism makes the same predictions as the state vector formalism. It is upon these revised postulates that our analysis of quantum information will be founded.

1.3 The Quantum Channel

A *quantum channel* describes the evolution of a system when it is not isolated. While we have used the term channel to describe the abstract evolution of a quantum state, in information theory it has strong implications for the storage and transfer of information. It is the purpose of this

section to describe the former, while the relationship between the evolution of a quantum system and information theory will be discussed in detail in the following section.

Postulate 2* tells us that when an isolated system evolves, its evolution is described by the conjugation by some unitary operator. However, when sending information between parties, such as in qubit communication, we can not usually assume that the system is closed. In other words, it is often the case that the system interacts with some environment. Systems like these are called *open*, and any change in their state is referred to as *open evolution*.

In order to discuss the evolution of an open quantum system, we must begin with the composite space formed by the system and its environment. The fourth postulate of quantum mechanics tells us that if the state spaces for the system and its environment are $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, then the composite system's state space is $\mathcal{H}_1 \otimes \mathcal{H}_2$. Any state from this isolated system then evolves unitarily; after which one may "trace out" the environment. That is, the evolution of the state $\rho \in \mathcal{H}_1$, when allowed to interact with its environment $\xi \in \mathcal{H}_2$, is characterized by the mapping

$$\epsilon(\rho) = \mathrm{tr}_E\big(\mathbf{U}^t \rho \otimes \xi \mathbf{U}\big), \tag{1.11}$$

where $\mathbf{U}$ is a unitary operator acting on the composite state space $\mathcal{H}_1 \otimes \mathcal{H}_2$ and tr_E is the partial trace over the state space of the environment.

While this narrative might appear straight-forward, we have made a subtle assumption that is worth mentioning. We assumed that the system and environment are in a state described by a pure tensor product (or a product state), which is not true in general [30]. However, we will not delve into the justifications for our definition of a quantum channel, but instead simply provide a characterization for the case when a system and environment are in a product state. For a detailed discussion on the assumptions behind our definition, see [9].

We would like to further mention that in certain circumstances an additional assumption may be made beyond those that define a quantum channel. When sending the completely mixed state through a channel, it is reasonable in some cases to assume that the output would not be "less" mixed than the input. With this in mind, we sometimes require that a quantum channel leaves the completely mixed state unchanged, i.e. fixes it. As shown in [21], this assumption is equivalent to considering the channels in which entropy increases.

Definition 1.3.1: A *quantum channel* $\epsilon : \Omega^n \to \Omega^n$ is a completely positive, trace preserving, convex-linear map. We say a channel is *unital* if it fixes the completely mixed state $\frac{I}{n}$.

Remark: Unitality is not included in the definition of a quantum channel as there exist natural examples of non-unital channels. For instance, when preparing a quantum state, we interact with a system such that independent of the initial state we obtain some predetermined output. Since there is no reason this interaction could not exist, we are justified in excluding unitality as a defining characteristic of a quantum channel.

Here is an alternative characterization of quantum channels from [9,11]:

Theorem 1.3.2 (Chuang and Nielsen): The map $\epsilon : \Omega^n \to \Omega^n$ is completely positive, trace preserving, and convex-linear if and only if it has the form

$$\epsilon(\rho) = \sum_{i=1} A_i \rho A_i^\dagger, \tag{1.12}$$

where each A_i is a linear operator on the state space of the system and $\sum_i A_i^\dagger A_i = I$. Since the Hilbert space of our system is $\mathbb{C}^n$, each A_i is a matrix and $\epsilon(\rho)$ is self-adjoint for all ρ . Further, by Choi's theorem [8], a linear opeator $\epsilon : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ is completely positive if and only if it

has the form

$$\epsilon(\rho) = \sum_{i=1} E_i \rho E_i^\dagger, \tag{1.13}$$

where each E_i is a complex matrix and, in general, can be taken to be linearly independent. The E_i's are called *Kraus operators* and Eq.(1.13) is referred to as the *Kraus representation* of the quantum channel ϵ. Simply put, Theorem 1.3.2 tells us that every quantum channel is the restriction of a completely positive, trace preserving, linear map on the Hermitian matrices, while conversely, Choi's Theorem says that every quantum channel can be uniquely extended to a completely positive, trace preserving, linear map of Hermitian matrices. It is also shown in [9,11], as well as many other works, that the set of completely positive trace preserving unital maps on $M_n(\mathbb{C})$ is a compact, convex set that is closed under composition. Throughout this work, we will restrict ourselves to the set of unital quantum channels on two-level systems, which we will denote by $\mathcal{U}$.

1.4 Information Theory

In this dissertation we consider two types of communication: classical and quantum. Let us explain each of these in turn.

A common case of classical communication is the transfer of a binary string (a list of 0's and 1's) between two parties that is encoded in some number of classical systems. We refer to each individually encoded system corresponding to an entry of the string as a *bit*. If party one (Alice) prepares and sends a series of bits to party two (Bob), then assuming interaction with the environment, there exists the potential of any given bit changing its state. In other words, a 0 turning to a 1 or vice versa. Therefore, in order to successfully communicate, Alice and Bob require a means of studying possible errors. Such can be accomplished by modeling this interaction with a *classical binary channel*. That is, a channel with two possible inputs and two possible outputs (e.g. 0 and 1). A

classical binary channel is represented by a 2×2 stochastic matrix; we denote the set of classical binary channels by $(2, 2)$. In other words, a member of $(2, 2)$ is a matrix of the following form:

$$\begin{pmatrix} x & 1-x \\ y & 1-y \end{pmatrix}, \tag{1.14}$$

where $x = p(0|0)$ is the probability that a 0 is received when a 0 is sent, and $y = p(0|1)$ is the probability that a 0 is received when a 1 is sent.

A central concern of information theory is, "How much information can I send?". The answer to this question depends on many variables; one of which is the environment we send our information through. Classically, the largest amount of information one may transmit through an environment is called the *capacity* of the channel. As shown in [19], the capacity for a classical binary channel is

$$C(x, y) = \log_2 \left(2^{\frac{\overline{x}H(y) - \overline{y}H(x)}{x-y}} + 2^{\frac{yH(x) - xH(y)}{x-y}} \right), \tag{1.15}$$

where $H(x) = -x \log_2(x) - (1-x) \log_2(1-x)$ is the base two entropy, $\overline{x} = 1 - x$ and $\overline{y} = 1 - y$. For a detailed discussion on capacity and Shannon entropy, see [10,32].

At no point have we utilized any quantum effects in our example of classical communication. That is, the postulates of quantum mechanics are completely irrelevant in our description. And while the principles are the same in quantum communication (information is encoded in a system, which is then sent between parties through some environment in which it may evolve), the particulars are very different. To illustrate this difference we now consider the case where information is encoded in a quantum system.

Alice and Bob first choose a basis of the state space $\{|\psi\rangle, |\phi\rangle\}$, after which they let $|\psi\rangle$ respresent 0 and $|\phi\rangle$ represent 1, or vice versa. Preparing a quantum system in one of the two basis elements,

Alice encodes information as $|*\rangle\langle *|$, which we then refer to as a quantum bit or a qubit. Sending her qubit to Bob, the system evolves according to some quantum channel $\epsilon\big(|*\rangle\langle *|\big)$. Afterwards, Bob receives the qubit, performs a measurement on the system with respect to the $\{|\psi\rangle, |\phi\rangle\}$ basis, and obtains a state that corresponds to either a 0 or a 1. This process defines a classical binary channel, which in turn has a capacity.

While this discussion appears eerily similar to classical communication, there is a subtlety to this procedure that results in a huge difference between sharing classical information with a qubit versus a bit. At the beginning, Alice and Bob fix a basis of the state space; however, in principle, many unit vectors in the state space are possible states of the system. That is, there could be literally an infinite number of different bases with which Alice and Bob can choose to communicate. In turn, this means that associated with every quantum channel there are an infinite number of classical channels; each induced by a particular basis, and each having their own capacity. Therefore, varying over each basis of the state space, we obtain a range of capacities which we call the *scope* of the channel. For a detailed discussion on scope, see [21].

We then see that one fundamental difference between classical and quantum communication is the dependence on how we represent our information. In the classical case, our representation of a bit is not a variable when calculating how much information can be sent; while how we choose to prepare our system is at the very heart of this question in quantum communication. This provides the potential of an ideal representation for a given quantum channel. Meaning, it can be the case that encoding qubits with one basis will allow for more information to be exchanged than if a different basis were used. A detailed example of such can be found in [21], for which we provide a quick overview that will also serve to illustrate the complexity of calculating the scope.

We consider the bit-flip channel $\epsilon(\rho) = (1-p)\rho + pU\rho U^\dagger$, where $p \in [0,1]$ and

$$U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{1.16}$$

This is arguably the simplest non-trivial example of a channel where the state is flipped with probability p and unchanged otherwise. Suppose that Alice and Bob choose to encode their information such that $|\psi\rangle$ represents 0, while the state $|\phi\rangle$ represents 1. Letting $|e_1\rangle = [1\ 0]^t$ and $|e_2\rangle = [0\ 1]^t$, then there exists complex numbers a, b, c, and d such that $|\psi\rangle = a|e_1\rangle + b|e_2\rangle$ and $|\phi\rangle = c|e_1\rangle + c|e_2\rangle$. By assumption, after the qubit $|*\rangle\langle *|$ is sent through the environment, Bob receives the state

$$(1-p)|*\rangle\langle *| + pU|*\rangle\langle *|U^\dagger. \tag{1.17}$$

In order to calculate the capacity, we must first compute the probability that Bob receives a 0 when Alice sends 0:

$$\begin{aligned} x = p(0|0) &= \text{tr}\Big(|\psi\rangle\langle\psi|\big((1-p)|\psi\rangle\langle\psi| + pU|\psi\rangle\langle\psi|U^\dagger\big)\Big) \\ &= (1-p) + p\Big((a\bar{b})^2 + (\overline{a}b)^2 + 2|a|^2|b|^2\Big), \end{aligned} \tag{1.18}$$

and the probability that Bob receives a 0 when Alice sends 1

$$\begin{aligned} y = p(0|1) &= \text{tr}\Big(|\phi\rangle\langle\phi|\big((1-p)|\phi\rangle\langle\phi| + pU|\phi\rangle\langle\phi|U^\dagger\big)\Big) \\ &= p\Big(|a|^2|d|^2 + |b|^2|c|^2 + a\bar{b}c\overline{d} + \overline{a}b\overline{c}d\Big). \end{aligned} \tag{1.19}$$

Upon plugging these probabilities into Eq. (1.15), we then obtain the capacity of the channel when information is represented in the basis $\{|\psi\rangle, |\phi\rangle\}$. As you can see, calculating the capacity for a single basis is tedious, and obtaining the range for an infinite number of bases would prove beyond daunting. The question then remains, "How do we find the basis that allows for the largest amount

of information to be sent?". In this particular example a choice of basis where $a = b = c = -d = \frac{1}{\sqrt{2}}$ gives us $x = 1$ and $y = 0$ regardless of p. In turn, this results in $C(x, y) = 1$, which corresponds to perfect transmission.

While we have provided the ideal basis for the case of the bit flip channel, it is not obvious as to how we found it in the first place. Forfeiting the density operator representation and instead using the *Bloch representation*, this task is significantly simplified. Specifically, in [21] Martin shows how to calculate the scope of a unital qubit channel from the eigenvalues of a systematically constructed 3×3 real symmetric matrix. With this in mind, the next section will be devoted to the introduction of the Bloch representation in quantum information.

1.5 The Bloch Representation

Considering only qubits, we know that every density operator is described by a 2×2 self-adjoint matrix. Upon standard computations, it can be shown that the identity and spin operators, also known as the Pauli matrices,

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{1.20}$$

form a linearly independent set of self-adjoint matrices, and thus they form a basis for the set of self-adjoint 2×2 matrices. Therefore, each density operator can be uniquely written in the following form:

$$\rho = x_0 I + x_1 \sigma_1 + x_2 \sigma_2 + x_3 \sigma_3, \tag{1.21}$$

where $x_0,\ x_1,\ x_2$, and x_3 are real numbers. Furthermore, since each ρ has trace one and the spin operators are traceless, we must have that $x_0 = \frac{1}{2}$. For brevity, we factor $\frac{1}{2}$ out of each x_i, rename

the result x_i, and let $\sigma = [\sigma_1\ \sigma_2\ \sigma_3]^t$ be a vector of matrices. Then every state can be uniquely written as

$$\rho = \tfrac{1}{2}\big(I + \langle x, \sigma\rangle\big), \tag{1.22}$$

where $x \in \mathbb{R}^3$ is called the *Bloch vector*. It can further be shown that ρ is positive-semidefinite if and only if $x \in \mathbb{B}^3$; and $||x|| = 1$ is equivalent to $\rho^2 = \rho$. That is, the state of a qubit is fully described by a vector in $\mathbb{B}^3$, where the points in S^2 correspond to pure states [9].

Moreover, each pair of antipodal points on the two-sphere corresponds to a set of orthogonal unit vectors in $\mathcal{H}^2$. In other words, every pair of unit vectors x and $-x$ represents a means of encoding information in a quantum system. Therefore, we refer to the set $\{x, -x\}$ as a communication basis. Note, each representation, being a set of antipodal points, can be fully characterized by only one of its vectors. Consequently, we will henceforth refer to the set $\{x, -x\}$ as the x communication basis. To put it simply, the two-sphere S^2 can be used to denote the set of all representations, where the point x and $-x$ correspond to the same communication basis. With this in mind, extending postulate 3^* to the Bloch representation, the probability of obtaining the state x when performing a measurement on the state y is

$$p(x|y) = \frac{1 + \langle x, y\rangle}{2}. \tag{1.23}$$

Now turning our attention to the evolution of a Bloch vector, it can be shown that every qubit channel ϵ induces a map on the set of Bloch vectors $f_\epsilon : \mathbb{B}^3 \to \mathbb{B}^3$. Furthermore, it is shown in [9,11] that every f_ϵ is an affine map on the Bloch sphere and is linear if and only if ϵ is unital. Moreover, the set of unitary channels is given by $SO(3)$, and its convex closure is the set of unitals; which as we previously mentioned, will be the focus of this work.

The Bloch representation of a quantum channel is a non-expansive, linear operator on the unit ball that satisfies particular properties. With that said, we would like to highlight some of these properties with the following two propositions which are proven in [21]:

Proposition 1.5.1 (Martin): Let f_ϵ be the Bloch representation of a qubit channel ϵ.

(i) The function f_ϵ is convex linear.

(ii) The Bloch representation of a convex sum is the convex sum of the Bloch representations.

(iii) The composition of qubit channels corresponds to the composition of Bloch representations.

Proposition 1.5.2 (Martin): The Bloch representation of a unital channel f_ϵ is symmetric if and only if it is a convex sum of involutive rotations that form a copy of the Klein four group. Explicitly,

$$f_\epsilon = \sum_{i=0}^{4} p_i r^t s_i r, \tag{1.24}$$

where $p \in \Delta^4$, r is a member of $SO(3)$, and the s_i's are the Bloch representations of the identity and spin operators given by:

$$s_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad s_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad s_2 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad s_3 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{1.25}$$

Not only does the Bloch representation "look nicer", but it offers major simplifications when performing information theoretic calculations. For instance, Martin shows in [21] that if f is the Bloch representation of a unital qubit channel, then its scope is

$$s(f) = \left[\frac{1 + \operatorname{sgn}(\lambda_1 \lambda_3)}{2} \left(1 - H\left(\tfrac{1+\min(|\lambda_i|)}{2} \right),\ 1 - H\left(\tfrac{1+\max(|\lambda_i|)}{2} \right) \right) \right], \tag{1.26}$$

where $\lambda_3 \leq \lambda_2 \leq \lambda_1$ are the eigenvalues of $\varphi(f) = \frac{1}{2}(f + f^t)$ and H is the binary entropy function. Therefore, in the Bloch representation the task of calculating the scope of a unital qubit channel f is

reduced to obtaining the eigenvalues of the symmetric matrix $\varphi(f)$! While this result alone shows the significance of the Bloch representation, this dissertation will further expand upon its benefits. With this in mind, from now on when we speak of a unital channel, it should be understood that we are referring to its Bloch representation, unless otherwise stated.

1.6 Summary

In this chapter we introduced a way to quantify the amount of information one may send through a quantum channel, called scope. We also formulated a representation of qubits and channels where states are points on the unit ball and channels are nonexpansive, linear operators on $\mathbb{B}^3$. With this groundwork established, we begin the next chapter with our analysis on qubit communication protocols and their associated error rates.

CHAPTER 2: ERROR RATES

Before defining an error function for an arbitrary protocol, we need to introduce some basic concepts from Topology. Topology is a branch of mathematics concerned with properties that are preserved under continuous functions; or simply put, an important part of the foundation for real analysis.

2.1 Preliminaries: Point Set Topology

This is in no way meant to be a representation of topology as a whole, but rather a short introduction to the subject that will enable anyone unfamiliar with the topics to understand the results in this dissertation. With this being the purpose, we will introduce several definitions and examples but omit proofs of any theorems that are stated. For those new to this branch of mathematics, we recommend [27] by James Munkres. This section is based on Munkres.

Definition 2.1.1: Let X be a set. A collection of subsets of X, denoted by τ, is a *topology* if the following properties are satisfied:

(i) X and the empty set $\varnothing$ are members of τ.

(ii) The union of the elements of any arbitrary subcollection of τ is contained in τ.

(iii) The intersection of the elements of any finite subcollection of τ is contained in τ.

The pair (X, τ) is called a topological space, and every $U \in \tau$ is called an *open* set while its complement $X \backslash U$ is referred to as *closed*. Thus, for every topology both $\varnothing$ and X are open and closed. Furthermore, every subset $Y \subseteq X$ with the collection $\tau_Y = \{Y \cap U \mid U \in \tau\}$ satisfies

properties (i)-(iii) and is referred to as a topological *subspace* of (X, τ).

Definition 2.1.2: Let (X, τ) and (Y, σ) be topological spaces. The function $f : X \to Y$ is *continuous* if for every σ-open set $V \subseteq Y$, the set $f^{-1}(V)$ is τ-open in X. One may show that this is equivalent to the well-known ϵ-δ definition found in real analysis where both topological spaces are $(\mathbb{R}, \tau_{\mathcal{U}})$, which we will define briefly. A continuous bijection whose inverse is also continuous is called a *homeomorphism*.

In this dissertation we will be dealing with two specific topologies defined on a particular set of continuous functions. When it is the case that the elements of our underlying set are functions themselves, we refer to the space as a *topological function space*. In the remainder of this section, we will introduce several more definitions, including the topologies on the very function spaces we plan to study. With this in mind, let us denote the set of all continuous functions $f : X \to Y$ by $\mathcal{C}(X, Y)$. We begin with the definition of a very important subset of a topology.

Definition 2.1.3: A *basis* $\mathcal{B}$ is a collection of subsets of X that satisfies the following properties:

(i) For each $x \in X$, there is at least one basis element B containing x.

(ii) If x is contained in the intersection of two basis elements B_1 and B_2, then there exists a basis element B_3 such that $x \in B_3 \subseteq B_1 \cap B_2$.

Remark: With a basis $\mathcal{B}$ for the set X, we can then define a topology τ as follows: The set U is open in X if for each $x \in U$, there exists a $B \in \mathcal{B}$ such that $x \in B \subseteq U$.

Throughout this chapter we will make use of two particular topologies on $\mathcal{C}(X, Y)$. We characterize these topologies by constructing their respective bases; one of which will necessitate the following metric:

Let $f, g : X \to Y$ be continuous functions, where Y is a metric space with the metric d. The *uniform metric* on the set $\mathcal{C}(X, Y)$

$$\rho(f, g) = \sup_{x \in X} \Big\{ \bar{d}\big(f(x), g(x)\big) \;\Big|\; x \in X \Big\} \tag{2.1}$$

makes it a topological space, where

$$\bar{d}\big(f(x), f(y)\big) = \min \Big\{ d\big(f(x), f(y)\big), 1 \Big\}. \tag{2.2}$$

Remark: If X is compact and each function in $\mathcal{C}(X, Y)$ is bounded, then the uniform metric can be defined simply by d and the induced metric $\bar{d}$ is not necessary. This is because under such conditions the value of $d\big(f(x), g(x)\big)$ is finite and the supremum exists when varying over X.

Note, just as we can in a metric space consisting of points rather than functions, with the use of the metric we can then define the convergence of a sequence. That is, if $f_n : X \to Y$ is a sequence in $\mathcal{C}(X, Y)$, we say that f_n converges to f if $\rho(f_n, f) \to 0$ as $n \to \infty$. Furthermore, if the space $\mathcal{C}(X, Y)$ has a metric ρ, then a basis for it is given by:

If $f \in \mathcal{C}(X, Y)$, where Y is a metric space and ϵ is a positive real number, then the collection of all sets of the form

$$B_\rho(f, \epsilon) = \Big\{ g \in \mathcal{C}(X, Y) \;\Big|\; \rho(f, g) < \epsilon \Big\}, \tag{2.3}$$

is a basis for the *uniform topology* on $\mathcal{C}(X, Y)$.

In order to construct a basis for our second topology, we need to define the following property.

Definition 2.1.4: Let (X, τ) be a topological space. The collection $C = \big\{ U_\alpha \mid \alpha \in \Lambda \big\}$ is an *open cover* of $K \subseteq X$ if $K \subseteq \bigcup_{\alpha \in \Lambda} U_\alpha$, where each U_α is open. A set K is called *compact* if every open

cover of K contains a finite subcover.

With this property in mind, we can then construct the basis for the *compact-open topology* on $\mathcal{C}(X, Y)$ with the collection of sets

$$B(K, U) = \Big\{ f \;\Big|\; f(K) \subseteq U \Big\}, \tag{2.4}$$

where K is compact in X and U is open in Y.

The results of this work will be primarily dependent on the uniform and compact-open topologies. That is, when we set forth to prove topological statements, most results will pertain to a function space whose basis is characterized by either Eq. (2.3) or (2.4). The results that do not utilize these bases will instead make use of the topological space that is the foundation of real analysis. Namely, the set of real numbers $\mathbb{R}$ with the *usual topology* $\tau_{\mathcal{U}}$, which is generated by the basis with elements of the form:

$$B_\epsilon(x) = \Big\{ y \in \mathbb{R} \;\Big|\; ||x - y|| < \epsilon \Big\}, \tag{2.5}$$

where ϵ is some positive real number. Having defined the topological spaces we plan to use, we would like to introduce a couple more properties that they may posses.

Definition 2.1.5: A topological space (X, τ) is said to be *disconnected* if it is the union of two disjoint nonempty open sets. Otherwise it is said to be *connected*.

Definition 2.1.6: A space (X, τ) is called ***Hausdorff*** if for every pair of distinct elements $x, y \in X$, there exist disjoint open sets U and V such that $x \in U$ and $y \in V$.

The topological space $(\mathbb{R}, \tau_{\mathcal{U}})$ is a connected, Hausdorff space in which every closed interval $[a, b]$ is a compact subset. This result will be essential when studying the images of our error functions; however, we omit its proof as it is already well-established in the literature and would only serve as

a mathematical tangent. While this concludes all the definitions we need from topology, we would like to end this section with a list of results that will be used throughout this chapter. Again, we will not be including proofs as they are not necessary to follow the content of this dissertation. All of the following theorems can be found in Munkres, [27].

Theorem 2.1.7: A *topological invariant* is a property of a topological space that is invariant under homeomorphisms. Compactness, connectedness, and Hausdorff are all topological invariants.

Theorem 2.1.8: If a topological vector space (X, τ) is convex, then it is connected.

The Heine-Borel Theorem 2.1.9: A subset $K \subseteq \mathbb{R}^n$ is compact if and only if it is closed and bounded.

Theorem 2.1.10: Let A be a subset of the topological space X. We say $x \in X$ is a limit point of A if every open set U containing x contains at least one point in A distinct from x. The subset A is closed if and only if it contains all its limit points.

Theorem 2.1.11: Let X be a compact space, and (Y, d) a metric space. The compact-open and uniform topologies coincide on the function space $\mathcal{C}(X, Y)$.

Theorem 2.1.12: Let X and Y be topological spaces, and let $\mathcal{C}(X, Y)$ be endowed with the compact-open topology. If $f : X \times Z \to Y$ is continuous, then so is the induced function $F : Z \to \mathcal{C}(X, Y)$ such that

$$\big(F(z)\big)(x) = f(x, z). \tag{2.6}$$

Theorem 2.1.13 Let $f : X \to \mathbb{R}$ be a continuous function where X is a compact space. Then the function f assumes its extrema at some x and y in X respectively.

2.2 Error Functions of Qubit Communication Protocols

This section introduces some general background information by defining our qubit communication protocols and the functions associated with their rate of error. We also include some examples of each and a brief discussion on what we call *QKD-like security measures*.

2.2.1 Qubit Communication Protocols

As there exists a plethora of diverse schemes for communication protocols, we restrict ourselves to qubit protocols of a particular kind. For us, a qubit communication protocol exchanges classical information between two parties (Alice and Bob) via a quantum channel and is composed of three elements: a compact, connected set X of n-tuples of communication bases, a point q of the n-simplex, and a procedure consisting of the following five steps:

(i) Alice and Bob choose an element $(x_1, x_2, ..., x_n)$ of X consisting of the n communication bases in which they will prepare and measure their qubits.

(ii) Alice then prepares each qubit such that the x_i basis is used with probability q_i.

(iii) Bob then receives and measures each qubit. He chooses the basis in which to measure according to the probabilities $(q_1, q_2, ..., q_n)$.

(iv) Via a classical channel, Alice then tells Bob which basis was used to prepare each qubit.

(v) Lastly, the two parties discard each qubit that Bob measured in a basis different from the one in which Alice prepared. The remaining qubits constitute shared data.

Since all qubits that are measured in the incorrect basis are discarded in step (v), the only way for the wrong bit to be obtained is if the qubit interacts with its environment while traveling between

Alice and Bob; see Figure 2.1 below.

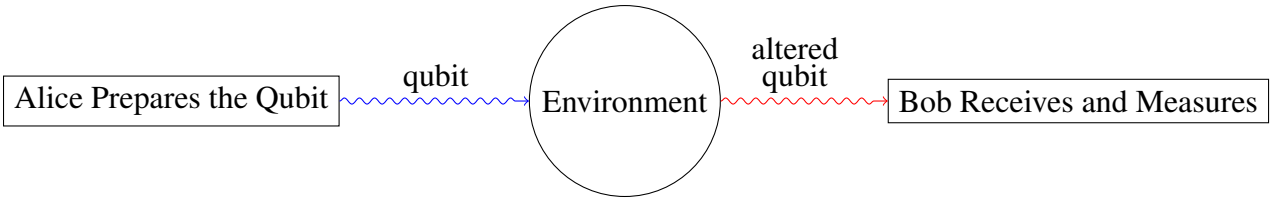

Figure 2.1: Pictorial representation of qubit communication.

Recall that all communication schemes of interest to us utilize the same procedure described above. What follows are some of our favorite examples.

2.2.2 Prepare and Send Protocol

The Prepare and Send protocol is arguably one of the simplest means of qubit communication. In this protocol Alice and Bob agree upon a single set of antipodal points in which to prepare and measure. Therefore, the Prepare and Send Protocol is characterized by the set $X = S^2$ and the point $q = (1)$.

We would like to note that the simplicity of this protocol enables us to omit steps (iv) and (v) of our procedure. That is, since we utilize only one communication basis, there is no way for Bob to perform an "incorrect measurement". Therefore, there is no need for a classical channel to broadcast which states to discard.

2.2.3 Quantum Key Distribution

Quantum Key Distribution (QKD) is a protocol in which a one-time pad is securely generated between two parties using quantum mechanics. This protocol requires that Alice and Bob agree upon a pair of orthogonal communication bases, each one utilized with a probability of $\frac{1}{2}$ [2]. That is, QKD is characterized by the set $X = \Big\{(x, y) \in S^2 \times S^2 \mid \langle x, y\rangle = 0\Big\}$ and the point $q = \big(\frac{1}{2}, \frac{1}{2}\big)$.

The requirement that the bases be orthogonal is of great importance. Let us assume there exists an eavesdropper $\big(\text{Eve}\big)$ that intercepts qubits sent by Alice, measures them, and then transmits the resulting qubit to Bob. Then any information Eve gains can introduce error in Bob's measurements because x and y are physically indistinguishable states. Let us explain this as follows:

Since Eve has no knowledge of the communication basis in which each qubit is prepared, she can only guess with a $\frac{1}{2}$ probability of accuracy when performing her own measurements. If she chooses correctly, then Eve will obtain the state sent by Alice and send it to Bob resulting in no error. However, if she chooses incorrectly, the state she obtains is not only in a different basis but random, and Bob will necessarily receive a state that is different from the one Alice prepared. That is, if Eve intercepts the state x but measures in the y basis, then she would send Bob either the state y or $-y$, each with probability

$$p(y|x) = \frac{1 + \langle x, y\rangle}{2} = \frac{1}{2}. \tag{2.7}$$

Assuming Bob measures the qubit in the same basis that Alice used to prepare (otherwise the qubit would have been discarded in the last step of the protocol), he would then obtain some result with a $\frac{1}{2}$ probability of error. Therefore, since Eve measures in the wrong basis with a probability of $\frac{1}{2}$, for each qubit sent Bob has an overall $\frac{1}{4}$ probability of error. Consequently, if the rate of error is too high, then Alice and Bob can agree to abort the protocol as an eavesdropper could be present.

Thus, we see that physically indistinguishable states can offer a means of security. Let us further explore protocols of this form.

2.2.4 Protocols with QKD-Like Security Measures

Before moving forward, we would like to take a moment to point out that the entirety of our study on protocols with a QKD-like security measure assumes no environmental effects. That is, we have assumed that as the qubit travels from Alice to Eve it does not interact with its environment; likewise between Eve and Bob. While this simplifies calculations and provides a good starting point for studying the effects of an eavesdropper, assuming no noise between each sender and receiver is not practical. The most reasonable scenario would involve a channel for each transmission of the qubit. That is, we would assume the qubit evolves according to some channel f when traveling to Eve, and another channel g as it makes its way to Bob; see Figure 2.2 below. However, environmental noise for the case of an eavesdropper is outside the scope of this work and will not be fully investigated.

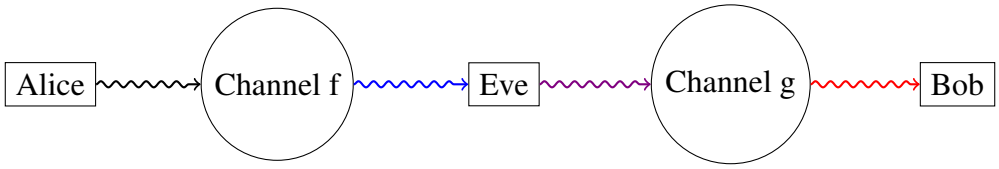

Figure 2.2: Pictorial representation of eavesdropper in qubit communication.

We say a qubit communication protocol has a "QKD-like security measure" if $q = \left(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right)$ and the components of each $x \in X$ are equiangular. That is, if $\left(x_1, x_2, ..., x_n\right)$ is in X, then for all $i \neq j$, $\langle x_i, x_j \rangle = \alpha$, for some fixed α. Then by the same arguments for the case of QKD, any information gained by Eve may result in an error in Bob's measurements. Let us now explicitly

calculate this error.

We begin by assuming Alice chooses to prepare a qubit in the x_i communication basis. Specifically, we say Alice prepares the qubit in the state x_i with probability p_1 and $-x_i$ with probability p_2. Then supposing Eve performs a measurement in the x_j basis (it is undetermined if $x_i = x_j$ or not), she would obtain the following results and send them to Bob:

When the qubit is prepared in the state x_i, after her measurement, Eve would then obtain and send the states x_j and $-x_j$ to Bob with probabilities $\frac{1+\langle x_j, x_i\rangle}{2}$ and $\frac{1-\langle x_j, x_i\rangle}{2}$ respectively. However, if Alice sends the state $-x_i$, then Eve would send Bob the states x_j and $-x_j$ with respective probabilities $\frac{1-\langle x_j, x_i\rangle}{2}$ and $\frac{1+\langle x_j, x_i\rangle}{2}$.

With Eve's effect established on the state Bob receives, we can now calculate the probability that Bob's measurement results in a state different from the one Alice sent. Assuming that Bob performs each measurement in the same communication basis used by Alice (This is safe to assume as the qubit would otherwise be discarded per our procedure), he would obtain the following results:

If Alice sends the state x_i, then after his measurement, Bob would obtain the state $-x_i$ with probability

$$\begin{aligned} p_j(-x_i|x_i) &= \left(\frac{1+\langle x_j, x_i\rangle}{2}\right)\left(\frac{1-\langle x_i, x_j\rangle}{2}\right) + \left(\frac{1-\langle x_j, x_i\rangle}{2}\right)\left(\frac{1+\langle x_i, x_j\rangle}{2}\right) \\ &= \frac{1-\langle x_j, x_i\rangle^2}{2}, \end{aligned} \tag{2.8}$$

where we have adopted the notation $p_j(-x_i|x_i)$ to denote the probability that Bob obtains the result $-x_i$ when Alice sends the state x_i and Eve performs her measurement in the x_j communication basis. Let us take a moment to discuss Eq. (2.8). In the first line, both products on the righthand side of the equation consist of two factors. The first factor being the probability that Eve sends a particular state $\big($either x_j or $-x_j\big)$ when Alice prepares the qubit as x_i. The second factor is the

probability that Bob measures the state sent by Eve and obtains $-x_i$.

By the same arguments, if Alice sends the state $-x_i$, then Bob's measurement would result in x_i with probability

$$p_j(x_i|-x_i) = \frac{1-\langle x_j, x_i\rangle^2}{2}. \tag{2.9}$$

Therefore, we have that $p_j(-x_i|x_i) = p_j(x_i|-x_i)$. Consequently, the probability Bob obtains an error when Alice sends either the state x_i or $-x_i$ and Eve measures in the x_j communication basis is

$$p_j(x_i) = p_1\, p_j(-x_i|x_i) + p_2\, p_j(x_i|-x_i) = p_j(-x_i|x_i), \tag{2.10}$$

recalling that p_1 and p_2 are the probabilities that Alice sends the states x_i and $-x_i$ respectively. Furthermore, noting that Eve chooses the x_j basis with probability $\frac{1}{n}$, then when Alice sends a qubit in the x_i communication basis, the overall error rate for Bob's measurement due to Eve is

$$\begin{aligned} p(x_i) &= \frac{1}{n}\sum_{j=1}^{n} p_j(x_i) \\ &= \frac{1}{n}\Bigg[\sum_{\substack{j=1 \\ j\neq i}}^{n} p_j(x_i) + p_i(x_i)\Bigg]. \end{aligned} \tag{2.11}$$

Recalling that our communication bases are equiangular, and noting

$$p_i(x_i) = \frac{1-\langle x_i, x_i\rangle^2}{2} = 0, \tag{2.12}$$

we have that

$$\begin{aligned} p(x_i) &= \frac{1}{n}\sum_{\substack{j=1\\ j\neq i}}^{n} p_j(x_i) = \frac{1}{n}\sum_{j\neq i}\frac{1-\langle x_j, x_i\rangle^2}{2} \\ &= \frac{1}{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{1-\alpha^2}{2} = \frac{(n-1)(1-\alpha^2)}{2n}. \end{aligned} \tag{2.13}$$

Wishing to finally calculate the overall probability of error for each qubit sent, we recall that Alice prepares the system in the x_i basis with probability $\frac{1}{n}$. Therefore, in a protocol with a QKD-like security measure, the rate of error for the measurement on each qubit due to an evaesdropper is

$$p(x) = \frac{1}{n}\sum_{i=1}^{n} p(x_i) = \frac{(n-1)(1-\alpha^2)}{2n}. \tag{2.14}$$

Let us quickly check Eq. (2.14) for the case of QKD where $n = 2$ and $\alpha = 0$. Then the probability of error is

$$\frac{(2-1)(1)}{2\cdot 2} = \frac{1}{4}, \tag{2.15}$$

which agrees with our previous discussion.

Remark: Eq. (2.14) tells us that for a fixed α the error due to an eavesdropper increases with n. However, we must not forget that each communication basis is characterized by a real 3-vector of unit length. Therefore, our n is bounded when we require that our communication bases are equiangular:

Lemma 2.2.1: The cardinality of a set of equiangular points on the unit sphere is bounded above by 4. That is, if

$$X = \left\{ x_i \in S^2 \;\middle|\; \langle x_i, x_j\rangle = \alpha \;\text{ for all }\; i \neq j \right\}, \tag{2.16}$$

then $\big|X\big| \leq 4$.

Proof: Since $SO(3)$ acts transitively on the two-shere and rotations preserve lengths and angles, we may assume that $x_1 = (1, 0, 0)$. Therefore, from the definition of X we must have that

$$x_i = \big(\alpha, b_i, c_i\big) \quad (\forall i \neq 1). \tag{2.17}$$

Moreover, if we apply a rotation of angle $\tan^{-1}(\frac{c_2}{b_2})$ about the e_1 axis to each element of X, then x_2 would assume the following form:

$$x_2 = \big(\alpha, b_2, 0\big). \tag{2.18}$$

Recalling that our unit vectors are defined such that

$$\langle x_2, x_3\rangle = \langle x_2, x_4\rangle = \alpha \tag{2.19}$$

we must have that $b_3 = b_4$; and since $||x_3|| = ||x_4||$, it must also be true that $c_3 = -c_4$. That is, we have assembled the following four unit vectors:

$$\begin{aligned} x_1 &= (1, 0, 0) \qquad x_2 = (\alpha, b_2, 0) \\ x_3 &= (\alpha, b_3, c_3) \quad x_4 = (\alpha, b_3, -c_3). \end{aligned} \tag{2.20}$$

We have now exhausted the elements of X. Simply put, if we wish to construct another x_i, some quick calculations show that it will necessarily be equal to one of the vectors we have already produced. Therefore, our set has cardinality four or less. Under this construction it is easy to verify that in the case of $n = 4$, our variables necessarily take on the following values:

$$\begin{aligned} \alpha &= -\frac{1}{3} \qquad b_2 = \frac{2\sqrt{2}}{3} \\ b_3 &= -\frac{\sqrt{2}}{3} \qquad c_3 = \sqrt{\frac{2}{3}}. \end{aligned} \tag{2.21}$$

Note, for the case of $n = 4$, the vectors given by Eqs. (2.20) and (2.21) can be used to define a symmetric, information complete, positive operator-valued measure, or SIC. For the curious reader, a detailed discussion on SICs can be found in [7].

□

While Lemma 2.2.1 is well-documented in the literature, we have included our proof for the sake of completeness and to provide the following original result.

Theorem 2.2.2: If a qubit communication protocol has a QKD-like security measure, then it must utilize four or less communication bases.

Proof: This is a direct result of Lemma 2.2.1.

□

Since the probability of error for each measurement due to Eve is given by

$$p(x) = \frac{(n-1)(1-\alpha^2)}{2n}, \tag{2.22}$$

it follows that $p(x) = 0$ when $n = 1$. Therefore, by Theorem 2.2.2 we have only three cases of interest. Namely, when n is equal to two, three, or four. When $n = 2$, the probability of error due to Eve is

$$\frac{1-\alpha^2}{4} \in \left[0, \tfrac{1}{4}\right]. \tag{2.23}$$

Similarly, for $n = 3$ the probability of error is

$$\frac{1-\alpha^2}{3} \in \left[0, \tfrac{1}{3}\right]. \tag{2.24}$$

In the case of $n = 4$, the probability is a constant $\frac{1}{3}$ as we would necessarily have that $\alpha = -\frac{1}{3}$.

Consequently, we have shown the new result that engaging in a qubit communication protocol with either three or four bases provides us with the potential to obtain the largest probability of error due to an eavesdropper; explicitly, $p(x) = \frac{1}{3}$.

In summary, we have described the qubit communication protocols we will be studying, provided two examples, and established a security measure for the case of an eavesdropper. We now aim to construct a function that computes the rate of error due to a unital qubit channel for each protocol. We are no longer assuming the presence of an eavesdropper, and therefore all error is assumed to originate solely from the environment. We do however note that there is a particular class of protocols with QKD-like security measures that provide the potential to simplify the calculations of error rates due to environmental noise. Specifically, we are referring to those protocols in which

$$X = \Big\{(rx_1, rx_2, ..., rx_n) \;\Big|\; r \in SO(3) \text{ and } (x_1, x_2, ..., x_n) \in X\Big\}. \tag{2.25}$$

Note, Eq. (2.25) is satisfied in QKD and the Prepare and Send Protocol. As a matter of fact, this is true for all protocols in which X contains every possible n-tuple of equiangular unit vectors for a fixed angle. That is, Eq. (2.25) is satisfied for every protocol where

$$X = \Big\{(x_1, x_2, ..., x_n) \in \prod^{n} S^2 \;\Big|\; \forall i \neq j \;\; \langle x_i, x_j \rangle = \alpha\Big\} \tag{2.26}$$

for some $\alpha \in [0, 1]$. Such protocols are said to have *ideal QKD-like bases*. Furthermore, when Eq. (2.25) is true and it is also the case that $q = \left(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right)$, we say that a protocol has an *ideal QKD-like security measure*. The significance of these conditions when calculating error rates will be made clear in Theorem 2.3.6.

2.2.5 The Error Functions

For us the error rate of a qubit communication protocol is the probability that Bob obtains a bit different from the one that Alice sends. Let us first consider a single basis with distinct states x and $-x$, each sent with probabilities p_1 and p_2 respectively. If the noisy environment is described by f, then from Eq. (1.23), the probability of Bob obtaining the wrong bit is

$$\begin{aligned} p(x) &= p_1\, p(-x|fx) + p_2\, p(x| - fx) \\ &= p_1\left(\frac{1+\langle -x, fx\rangle}{2}\right) + p_2\left(\frac{1+\langle x, -fx\rangle}{2}\right) \\ &= \frac{1-\langle x, fx\rangle}{2}, \end{aligned} \tag{2.27}$$

where $p(x|-fx)$ and $p(-x|fx)$ are the probabilities of obtaining the state x when measuring $-fx$ and $-x$ when measuring fx. More generally, if q_i is the frequency Alice prepares qubits in the x_i basis, then the overall error rate is

$$\sum_{i=1}^{n} q_i\, p(x_i). \tag{2.28}$$

Definition 2.1.3: An *error function* for an arbitrary qubit communication protocol is a function $E_{q,f} : X \to [0,1]$ such that

$$E_{q,f}(x) := \sum_{i=1}^{n} q_i \frac{1-\langle x_i, fx_i\rangle}{2}, \tag{2.29}$$

where X is a compact, connected set of n-tuples of communication bases and q is a point of the n-simplex. As discussed earlier, the two-sphere contains every representation of a qubit where the points x and $-x$ both correspond to the communication basis $\{x, -x\}$. Therefore, since X is an n-tuple of communication bases where each entry is a point in S^2, it is not surprising to see that

$$E_{q,f}\big((x_1, x_2, x_3, ..., x_n)\big) = E_{q,f}\big((\pm x_1, \pm x_2, \pm x_3, ..., \pm x_n)\big). \tag{2.30}$$

In other words, because we are representing the set of all communication bases with the two-sphere, we have that every error function is invariant under a sign change for any number of representations in its argument.

We would like to take a moment to address why X is required to be a compact, connected set. The error functions calculate probabilities associated with physically realizable protocols. Therefore, it is reasonable to apply physical constraints on the domain of these functions. We argue compactness due to the limitations of modern equipment, and assume connectedness as it ensures that we may apply the Intermediate Value Theorem.

Since X is a subset of a finite product of two-spheres, we get boundedness for free. Then by the Heine-Borel Theorem, X is compact if and only if it is closed. However, assuming X is not closed is equivalent to saying it does not contain all its limit points. Physically, this means that as we vary our communication bases in the lab (via a dial or some machine automated process) there exists at least one element of X that we can get arbitrarily close to and guarantee that we never obtain. Due to the limited precision of modern equipment, it is unrealistic to assume that we would be able to do such.

Furthermore, we require that X be connected as it, along with the continuity of $E_{q,f}$ (Theorem 2.2.2) and the total order on $\text{Im}(E_{q,f})$, allows us to apply the Intermediate Value Theorem (IVT). That is, if X is connected, then for every value r such that $E_{q,f}(a) < r < E_{q,f}(b)$ and $a, b \in X$, there exists a $c \in X$ where $E_{q,f}(c) = r$. Since the values of $E_{q,f}$ are error rates associated with physical protocols, one would expect that any value between two obtainable error rates would itself be achievable, and the connectedness of X is a natural way to ensure this. With an arbitrary qubit communication protocol described and its associated error function defined, we would now like to construct the error functions for the Prepare and Send Protocol and QKD.

2.2.6 *Error Functions for the Prepare and Send Protocol*

As a single basis is used for the Prepare and Send Protocol, we have $n = q = 1$. Therefore, assuming the channel f, the error rate is

$$E_f(x) = \frac{1 - \langle x, fx \rangle}{2}. \tag{2.31}$$

Due to the simplicity of the Prepare and Send Protocol's associated error functions, each one is referred to as a *simple function* and is denoted by E_f. The space of all simple functions is represented by $\mathbb{E}_s$.

2.2.7 *Error Functions for Quantum Key Distribution*

Since QKD calls for two orthogonal communication bases, each used with probability $\frac{1}{2}$, we have that $n = 2$ and $q = \left(\frac{1}{2}, \frac{1}{2}\right)$. Therefore, the error function for environmental noise f in QKD is

$$E_{q,f}(x, y) = \frac{1}{2}\left(\frac{1 - \langle x, fx \rangle}{2}\right) + \frac{1}{2}\left(\frac{1 - \langle y, fy \rangle}{2}\right), \tag{2.32}$$

where x and y are the predetermined bases such that $\langle x, y \rangle = 0$.

The error functions for the Prepare and Send Protocol and QKD will be the subject of many results in this dissertation. Specifically, in the next chapter we will introduce an adaptive algorithm that reduces the error rate for quantum communication and apply it to these two protocols. However, before doing such we first need to further establish the properties of our newly defined error functions.

2.3 Properties of Error Functions

Theorem 2.3.1: Let $E_{q,f}$ be an arbitrary error function,

$$E_{q,f}(x) = \sum_{i=1}^{n} q_i \frac{1 - \langle x_i, fx_i \rangle}{2}. \tag{2.33}$$

There exists a symmetric unital channel g such that $E_{q,f}(x) = E_{q,g}(x)$.

Proof: Each error function is determined by a sum of Euclidean inner products, explicitly the terms $\frac{1}{2} q_i \langle x_i, fx_i \rangle$. As shown in [21],

$$\begin{aligned} \langle x, fx \rangle &= \frac{1}{2}\Big(\langle x, fx \rangle + \langle x, fx \rangle \Big) \\ &= \frac{1}{2}\Big(\langle x, fx \rangle + \langle x, f^t x \rangle \Big) \\ &= \frac{1}{2} \langle x, (f + f^t) x \rangle \\ &= \langle x, \varphi(f) x \rangle, \end{aligned} \tag{2.34}$$

where $\varphi(f) \equiv \frac{f + f^t}{2}$ is necessarily symmetric and unital. Letting $g = \varphi(f)$, each inner product is left unchanged and the proof is complete.

□

In light of Theorem 2.3.1, when we talk about $\mathbb{E}$ we are referring to the space of error functions generated by $\mathcal{U}_s$, the set of symmetric unital channels.

Theorem 2.3.2: Each error function is continuous.

Proof: Again, as every error function can be written as the sum of the terms $\frac{1}{2} q_i \langle x_i, fx_i \rangle$ and the constant $\frac{1}{2}$, the continuity of the inner product implies the desired result.

□

Corollary 2.3.3: The image of each error function is a closed, bounded interval. That is, for any $E_{q,f} \in \mathbb{E}$, there are $a, b \in [0,1]$ with $a \leq b$ and $E_{q,f}(X) = [a,b]$.

Proof: By definition X is a compact, connected set. Then since $E_{q,f}$ is continuous, $E_{q,f}(X)$ is a compact, connected subset of $\mathbb{R}$. By connectedness the image of $E_{q,f}$ is an interval. Therefore, by compactness and the Heine-Borel Theorem, $\mathrm{Im}\big(E_{q,f}\big)$ is a closed, bounded interval.

□

Since the image of each error function is a closed, bounded interval, the achievable error rates are fully described by the values that occur between its extrema. The following results provide us with an explicit expression for the endpoints of the image of an arbitrary error function. We begin with the case of a simple function and remind the reader that we may assume each unital qubit channel is symmetric due to Theorem 2.3.1.

Theorem 2.3.4: Let f be a symmetric unital channel with eigenvalues $\lambda_3 \leq \lambda_2 \leq \lambda_1$. Then

$$\mathrm{Im}\big(E_f\big) = \Big[\tfrac{1}{2}(1-\lambda_1), \tfrac{1}{2}(1-\lambda_3)\Big]. \tag{2.35}$$

Proof: By Corollary 2.3.3, we only need to calculate the extrema of E_f; which by Theorem 2.1.13 we know are assumed since the image of each error function is a compact subset of $\mathbb{R}$. Furthermore, since E_f is a simple function given by

$$E_f(x) = \frac{1 - \langle x, fx\rangle}{2}, \tag{2.36}$$

we have by direct implication that

$$\max_{x\in X}\big(E_f\big) = \tfrac{1}{2} - \tfrac{1}{2}\min_{x\in X}\Big(\langle x, fx\rangle\Big) \tag{2.37}$$

and

$$\min_{x\in X}\big(E_f\big) = \tfrac{1}{2} - \tfrac{1}{2}\max_{x\in X}\Big(\langle x, fx\rangle\Big). \tag{2.38}$$

It is a well-documented result, that for a symmetric matrix A with eigenvalues $\lambda_{min} \leq \lambda \leq \lambda_{max}$,

$$\max_{||x||=1}\langle x, Ax\rangle = \lambda_{max} \quad \text{and} \quad \min_{||x||=1}\langle x, Ax\rangle = \lambda_{min}. \tag{2.39}$$

It then follows that Eqs. (2.37), (2.38), and (2.39) together give us the desired result,

$$\mathrm{Im}\big(E_f\big) = \Big[\tfrac{1}{2}(1-\lambda_1), \tfrac{1}{2}(1-\lambda_3)\Big]. \tag{2.40}$$

□

While Theorem 2.3.4 states the obtainable values for a simple function, this is a comparably small and unique subset of $\mathbb{E}$. For the sake of practicality, we therefore wish to obtain a similar result for an arbitrary error function. We aim to do so with the following two results.

Remark: If the definition of S in the following lemma seems exotic, we ask the reader to bear with us as the theorem that follows will explicitly show where such sets originate. Moreover, this set will be carefully examined in Chapter 3 for the case of QKD.

Lemma 2.3.5: Let $S = \Big\{p \in \prod^{3}[0,a] \;\Big|\; \sum_{i=1}^{3} p_i = 1\Big\}$ be a non-empty subset of Δ^3. Then for real

numbers $\lambda_3 \leq \lambda_2 \leq \lambda_1$ the extremum of $\phi : S \to \mathbb{R}$ where

$$\phi(p) = \sum_{i=1}^{3} p_i \lambda_i, \tag{2.41}$$

is given by:

$$\max_{p \in S} \big(\phi\big) = a\lambda_1 + b\lambda_2 + c\lambda_3 \quad \& \quad \min_{p \in S} \big(\phi\big) = c\lambda_1 + b\lambda_2 + a\lambda_3, \tag{2.42}$$

where

$$b = \begin{cases} 1 - a, & \text{if } a \geq \frac{1}{2} \\ a, & \text{otherwise,} \end{cases} \tag{2.43}$$

and $c = 1 - a - b$.

Proof: We begin by verifying that the maximum and minimum values are assumed by ϕ. Since S is the intersection of two closed, bounded sets, namely Δ^3 and $\prod^{3}[0, a]$, it follows that it is itself a closed and bounded subset of $\mathbb{R}^3$. Therefore, S is compact, and by Theorem 2.1.13 the continuous function ϕ assumes its extrema. Furthermore, noting that if $(a, b, c) \notin S$, then $S = \emptyset$, it follows that (a, b, c), and similary (c, b, a), are contained in S.

We approach this proof by assuming there exists a choice of p that obtains a larger value than $\phi(a, b, c)$ and draw a contradiction. Supposing that such a p exists then

$$\phi(p) = p_1\lambda_1 + p_2\lambda_2 + p_3\lambda_3 > a\lambda_1 + b\lambda_2 + c\lambda_3. \tag{2.44}$$

Utilizing the fact that the p_i's sum to 1, Eq. (2.44) is equivalent to

$$\big(p_2 - b\big)\big(\lambda_2 - \lambda_3\big) > \big(a - p_1\big)\big(\lambda_1 - \lambda_3\big). \tag{2.45}$$

Moreover, as $(a - p_1)$, $(\lambda_2 - \lambda_3)$, and $(\lambda_1 - \lambda_3)$ are all non-negative, it follows that $(p_2 - b) > 0$. Consequently, $\big(\lambda_1 - \lambda_3\big) \geq \big(\lambda_2 - \lambda_3\big)$ implies that

$$\big(p_2 - b\big)\big(\lambda_1 - \lambda_3\big) > \big(a - p_1\big)\big(\lambda_1 - \lambda_3\big). \tag{2.46}$$

Furthermore, we may assume $\lambda_1 \neq \lambda_3$; otherwise we have the trivial case where $\phi(p) = \lambda_1$ for all $p \in S$. Then dividing each side by $\big(\lambda_1 - \lambda_3\big)$, Eq. (2.46) becomes

$$p_1 + p_2 > a + b. \tag{2.47}$$

By the definition of b, we then have one of two possibilities:

(i) $\big(a \geq \frac{1}{2}\big) \quad \Rightarrow \quad p_1 + p_2 > a + b = a + 1 - a = 1.$

(ii) $\big(a < \frac{1}{2}\big) \quad \Rightarrow \quad p_1 + p_2 > a + b = 2a.$

For both cases $p \notin S$, and we have thus arrived at a contraction. The proof for the minimum follows by similar arguments. Assume there exists a $p \in S$ such that

$$\phi(p) = p_1\lambda_1 + p_2\lambda_2 + p_3\lambda_3 < c\lambda_1 + b\lambda_2 + a\lambda_3. \tag{2.48}$$

Recalling that the p_i's sum to 1, Eq. (2.48) becomes

$$(a - p_3)(\lambda_1 - \lambda_3) < (p_2 - b)(\lambda_1 - \lambda_2). \tag{2.49}$$

Once again, we have that $(p_2 - b)$ is positive, and we may therefore write

$$(a - p_3)(\lambda_1 - \lambda_3) < (p_2 - b)(\lambda_1 - \lambda_3). \tag{2.50}$$

Assuming $\lambda_1 \neq \lambda_3$, we can then divide by $(\lambda_1 - \lambda_3)$ on both sides and finally obtain

$$p_2 + p_3 > a + b. \tag{2.51}$$

We then have one of the following two possibilities:

(i) $\left(a \geq \frac{1}{2}\right) \quad \Rightarrow \quad p_2 + p_3 > a + b = a + 1 - a = 1.$

(ii) $\left(a < \frac{1}{2}\right) \quad \Rightarrow \quad p_2 + p_3 > a + b = 2a.$

In both cases p is not contained in S, and the proof is complete.

□

Theorem 2.3.6: Let $E_{q,f} \in \mathbb{E}$ have ideal QKD-like security bases, i.e. Eq. (2.25), where $f \in \mathcal{U}_s$ with eigenvalues $\lambda_3 \leq \lambda_2 \leq \lambda_1$, and the functions $p_j : X \to [0, 1]$ are given by

$$p_j(x) = \sum_{i=1}^{n} q_i (x_i)_j^2. \tag{2.52}$$

If $\max_{x \in X}(p_j) = \max_{x \in X}(p_k)$ for all $j \neq k$, then

$$\mathrm{Im}\big(E_{q,f}\big) = \left[\tfrac{1}{2}\big(1 - a\lambda_1 - b\lambda_2 - c\lambda_3\big), \tfrac{1}{2}\big(1 - c\lambda_1 - b\lambda_2 - a\lambda_3\big)\right] \tag{2.53}$$

where $a = \max_{x \in X}(p_i)$,

$$b = \begin{cases} 1 - a, & \text{if } a \geq \frac{1}{2} \\ a, & \text{otherwise,} \end{cases} \tag{2.54}$$

and $c = 1 - a - b$.

Proof: With so many variables floating around, it is important to know the origins of each one. The λ_i's are the eigenvalues of the symmetric channel f, while the p_j functions, and therefore $a, b,$ and c, are defined with respect to the error function. That is, the λ_i's are dependent on the channel f, where as a, b, and c come from the protocol itself.

We begin by noting that since f is symmetric it can be diagonalized by some rotation r. That is,

$$E_{q,f}(x) = \tfrac{1}{2}\left[1 - \sum_{i=1}^{n} q_i \langle rx_i, \lambda r x_i\rangle\right] \tag{2.55}$$

where

$$\lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}. \tag{2.56}$$

Since our protocol has ideal QKD-like bases, it follows from Eqs. (2.25) and (2.55) that our maximum and minimum values are given by

$$\max_{x\in X}\left(E_{q,f}\right) = \frac{1}{2}\left[1 - \min_{x\in X}\left(\sum_{i=1}^{n} q_i\langle x_i, \lambda x_i\rangle\right)\right] \tag{2.57}$$

and

$$\min_{x\in X}\left(E_{q,f}\right) = \frac{1}{2}\left[1 - \max_{x\in X}\left(\sum_{i=1}^{n} q_i\langle x_i, \lambda x_i\rangle\right)\right]. \tag{2.58}$$

Considering the summation in our extrema, we then have that

$$\begin{aligned}\sum_{i=1}^{n} q_i\langle x_i, \lambda x_i\rangle &= \sum_{i=1}^{n} q_i \sum_{j=1}^{3} (x_i)_j^2 \lambda_j \\ &= \sum_{j=1}^{3} \lambda_j \left(\sum_{i=1}^{n} q_i (x_i)_j^2 \right) \\ &= \sum_{j=1}^{3} p_j(x)\lambda_j.\end{aligned} \tag{2.59}$$

Checking that for all $x \in X$,

$$\sum_{j=1}^{3} p_j(x) = \sum_{i=1}^{n} q_i \sum_{j=1}^{3} (x_i)_j^2 = \sum_{i=1}^{n} q_i = 1, \tag{2.60}$$

then since $a = \max\limits_{x\in X} \left(p_j\right)$ for each j, it follows from Lemma 2.3.5 that

$$\max_{x\in X} \left(E_{q,f}\right) = \frac{1}{2}\Big[1 - c\lambda_1 - b\lambda_2 - a\lambda_3\Big] \tag{2.61}$$

and

$$\min_{x\in X} \left(E_{q,f}\right) = \frac{1}{2}\Big[1 - a\lambda_1 - b\lambda_2 - c\lambda_3\Big]. \tag{2.62}$$

□

It is essential that we address the assumptions of Theorem 2.3.6. In order to apply Lemma 2.3.5, we assumed that each p_j is bounded by the same value a. Physically, this means that for any $y \in S^2$ there exists an element $x \in X$ that contains y as one of its components. While this might seem reasonable on the surface, the practical application of a qubit protocol might call for different bounds.

One such example could arise in qubit communication between a satellite and Earth. In order to

reduce the payload, one might wish to limit the amount of equipment on the satellite. This would potentially restrict the states that could be detected. In turn, this could result in a restriction on the states in which the qubits are prepared, and ultimately end with different extremum for our p_j functions. Such a case is considered in the Appendix where a generalized version of Lemma 2.3.5 and Theorem 2.3.6 is provided. Furthermore, we also provide an additional version of Theorem 2.3.6, where we consider protocols that do not posses ideal QKD-like bases.

Theorem 2.3.7: Let $E_{q,f}$ be an error function such that Theorem 2.3.6 is applicable. Then there exists a diagonal unital channel g such that

$$\text{Im}(E_{q,f}) = \text{Im}(E_g). \tag{2.63}$$

Proof: We prove this result by constructing the channel g. We begin by recalling the following results, which are proven in Theorems 2.3.4 and 2.3.6:

$$\text{Im}\big(E_g\big) = \Big[\tfrac{1}{2}(1-\eta_1), \tfrac{1}{2}(1-\eta_3)\Big] \tag{2.64}$$

$$\text{Im}\big(E_{q,f}\big) = \Big[\tfrac{1}{2}\big(1-a\lambda_1-b\lambda_2-c\lambda_3\big), \tfrac{1}{2}\big(1-c\lambda_1-b\lambda_2-a\lambda_3\big)\Big] \tag{2.65}$$

where $\lambda_3 \leq \lambda_2 \leq \lambda_1$ and $\eta_3 \leq \eta_2 \leq \eta_1$ are the respective eigenvalues for f and the diagonal unital channel g. Therefore, if we assume $\text{Im}\big(E_g\big) = \text{Im}\big(E_{q,f}\big)$, it then follows that

$$\eta_1 = a\lambda_1 + b\lambda_2 + c\lambda_3 \tag{2.66}$$

$$\eta_3 = c\lambda_1 + b\lambda_2 + a\lambda_3. \tag{2.67}$$

Letting $\eta_2 = b\lambda_1 + a\lambda_2 + c\lambda_3$, we then complete the proof by showing g is unital and $\eta_3 \leq \eta_2 \leq \eta_1$. From [21], we know g, being a symmetric matrix, is unital if and only if its eigenvalues satisfy the

following inequalities:

$$(\forall i) \quad |\eta_i| \leq 1, \tag{2.68}$$

$$1 + \eta_1 + \eta_2 + \eta_3 \geq 0, \tag{2.69}$$

$$1 + \eta_1 - \eta_2 - \eta_3 \geq 0, \tag{2.70}$$

$$1 - \eta_1 + \eta_2 - \eta_3 \geq 0, \tag{2.71}$$

$$1 - \eta_1 - \eta_2 + \eta_3 \geq 0. \tag{2.72}$$

We then prove the unitality of g by showing that the matrix

$$g = \begin{pmatrix} \eta_1 & 0 & 0 \\ 0 & \eta_2 & 0 \\ 0 & 0 & \eta_3 \end{pmatrix} \tag{2.73}$$

satisfies inequalities (2.68-2.72). In order to show this, we begin by noting that $c \geq b \geq a \geq 0$, $a + b + c = 1$, and the eigenvalues of f, being a symmetric unital channel itself, satisfy the same inequalities that we desire of g. That is,

$$(\forall i) \quad |\lambda_i| \leq 1, \tag{2.74}$$

$$1 + \lambda_1 + \lambda_2 + \lambda_3 \geq 0, \tag{2.75}$$

$$1 + \lambda_1 - \lambda_2 - \lambda_3 \geq 0, \tag{2.76}$$

$$1 - \lambda_1 + \lambda_2 - \lambda_3 \geq 0, \tag{2.77}$$

$$1 - \lambda_1 - \lambda_2 + \lambda_3 \geq 0, \tag{2.78}$$

Therefore, since each η_i is a convex sum of the λ_i's, we have that for all i

$$-1 \leq \lambda_3 \leq \eta_i \leq \lambda_1 \leq 1. \tag{2.79}$$

It then follows that inequality (2.68) is satisfied. Considering the lefthand side of inequalities (2.69-2.72), we then have that

$$\begin{aligned} 1 + \eta_1 + \eta_2 + \eta_3 &= 1 + \lambda_1(a+b+c) + \lambda_2(a+b) + b\lambda_2 + \lambda_3(a+c) + c\lambda_3 \\ &= 1 + \lambda_1 + \lambda_2(1-c) + b\lambda_2 + \lambda_3(1-b) + c\lambda_3 \\ &= (1 + \lambda_1 + \lambda_2 + \lambda_3) + (b-c)(\lambda_2 - \lambda_3) \geq 0, \end{aligned} \tag{2.80}$$

$$\begin{aligned} 1 + \eta_1 - \eta_2 - \eta_3 &= a(1 + \lambda_1 - \lambda_2 - \lambda_3) + b(1 + \lambda_2 - \lambda_1 - \lambda_2) + c(1 + \lambda_3 - \lambda_3 - \lambda_1) \\ &= a(1 + \lambda_1 - \lambda_2 - \lambda_3) + b(1 - \lambda_1) + c(1 - \lambda_1) \geq 0, \end{aligned} \tag{2.81}$$

$$\begin{aligned} 1 - \eta_1 + \eta_2 - \eta_3 &= a(1 - \lambda_1 + \lambda_2 - \lambda_3) + b(1 - \lambda_2 + \lambda_1 - \lambda_2) + c(1 - \lambda_3 + \lambda_3 - \lambda_1) \\ &= a(1 - \lambda_1 + \lambda_2 - \lambda_3) + b(1 - \lambda_2) + b(\lambda_1 - \lambda_2) + c(1 - \lambda_1) \geq 0, \end{aligned} \tag{2.82}$$

and lastly,

$$\begin{aligned} 1 - \eta_1 - \eta_2 + \eta_3 &= a(1 - \lambda_1 - \lambda_2 + \lambda_3) + b(1 - \lambda_2 - \lambda_1 + \lambda_2) + c(1 - \lambda_3 - \lambda_3 + \lambda_1) \\ &= a(1 - \lambda_1 - \lambda_2 + \lambda_3) + b(1 - \lambda_1) + c(\lambda_1 - \lambda_3) + c(1 - \lambda_3) \geq 0. \end{aligned} \tag{2.83}$$

Note, the last step for inequalities (2.80-2.83) each result from the inequalities given in inequalities

(2.75-2.78) respectively. Lastly we show $\eta_3 \leq \eta_2 \leq \eta_1$. Explicitly,

$$\begin{aligned} \eta_1 - \eta_2 &= a\lambda_1 + b\lambda_2 + c\lambda_3 - a\lambda_2 - b\lambda_1 - c\lambda_3 \\ &= (a-b)(\lambda_1 - \lambda_2) \geq 0. \end{aligned} \tag{2.84}$$

Similarly, for η_2 and η_3,

$$\begin{aligned} \eta_2 - \eta_3 &= a\lambda_2 + b\lambda_1 + c\lambda_3 - a\lambda_3 - b\lambda_2 - c\lambda_1 \\ &= a(\lambda_2 - \lambda_3) + b(\lambda_1 - \lambda_2) + c(\lambda_3 - \lambda_1) \\ &\geq c(\lambda_2 - \lambda_3 + \lambda_1 - \lambda_2 + \lambda_3 - \lambda_1) = 0, \end{aligned} \tag{2.85}$$

where the final step follows directly from the fact that $0 \leq c \leq b \leq a$.

□

In conclusion, for each protocol with ideal QKD-like bases, the image of its error function can be calculated by a simple function with a diagonal unital channel (Note, this result can be further generalized by Theorem 1 in the Appendix)! Therefore, we may restrict our focus to the set $\mathbb{E}_s$. With this in mind, we now look at some of the properties of the function space $\mathbb{E}_s$ itself.

2.4 Properties of the Space of Error Functions

Lemma 2.4.1: The uniform topology and the compact-open topology coincide on the set $\mathbb{E}_s$.

Proof: By Theorem 2.1.11 since X and $[0,1]$ are compact metric spaces, the uniform topology and compact-open topology coincide on the function space $\mathbb{E}_s$.

□

Lemma 2.4.2: The convex function $G : \mathcal{U} \to \mathbb{E}_s$ where

$$\big(G(f)\big)(x) = E_f(x) \tag{2.86}$$

is continuous when $\mathcal{U}$ and $\mathbb{E}_s$ each have the uniform metric topology.

Proof: We prove this result by first showing that $E : \mathcal{U}_s \times S^2 \to \mathbb{R}$ where

$$E(f, x) = \frac{1 - \langle x, fx \rangle}{2}, \tag{2.87}$$

is continuous. Let (f_n) and (x_n) be sequences in $SO(3)$ and S^2 such that $f_n \to f$ and $x_n \to x$ respectively. We show that it then follows $f_n x_n \to fx$, which in turn implies the continuity of E. Note, this argument relies on the fact that every unital channel is contained in the convex closure of $SO(3)$.

Each element in $SO(3)$ is a linear operator between finite dimensional normed spaces. Thus each f in $SO(3)$, is continuous. It then follows that

$$\begin{aligned} ||f_n x_n - fx|| &= ||f_n x_n - f_n x + f_n x - fx|| \\ &\leq ||f_n||\, ||x_n - x|| + ||f_n x - fx||, \end{aligned} \tag{2.88}$$

where first we use the triangle inequality and then the fact that $||Ax|| \leq ||A||\, ||x||$ for any matrix A and vector x. Furthermore, since $||f_n|| = 1$ for all f_n in $SO(3)$, then

$$||f_n x_n - fx|| \leq 1 \cdot ||x_n - x|| + ||f_n x - fx||. \tag{2.89}$$

So as $n \to \infty$, we have that $||f_n x_n - fx|| \to 0$, i.e. $f_n x_n \to fx$, and we are left to conclude that E is continuous. Therefore, by Theorem 2.1.12, G is continuous in the compact-open topology.

Finally, applying Lemma 2.4.1, we have the desired result. The fact that G is convex is quickly verified by considering $G\big(pf + (1-p)g\big)$ where $f, g \in \mathcal{U}$ and $p \in [0,1]$. That is,

$$\begin{aligned} G\big(pf + (1-p)g\big)(x) &= \frac{1}{2}\Big(1 - \langle x, \big(pf + (1-p)g\big)x\rangle\Big) \\ &= \frac{1}{2}\Big(1 - p\langle x, fx\rangle - (1-p)\langle x, gx\rangle\Big) \\ &= \frac{1}{2}\Big(p\big[1 - \langle x, fx\rangle\big] + (1-p)\big[1 - \langle x, gx\rangle\big]\Big) \\ &= pE_f + (1-p)E_g. \end{aligned} \tag{2.90}$$

□

In light of Lemmas 2.4.1 and 2.4.2, from now on when we speak of $\mathbb{E}_s$ we will be assuming the uniform topology. Furthermore, when we speak of the simple function generated by the channel f, we are referring to $G(f) = E_f$.

Theorem 2.4.3: The set of simple functions $\mathbb{E}_s$ is a compact, convex space.

Proof: From Lemma 2.4.2 we know that $G : \mathcal{U} \to \mathbb{E}_s$ is a continuous, convex function. Therefore, since the set of unital channels is compact, it then follows that the same is true for the space $G(\mathcal{U}) = \mathbb{E}_s$.

□

With some of the more important properties of $\mathbb{E}_s$ established, we close this chapter by exploring a particular subset of symmetric unital channels and the error functions they generate. Letting $V = \big\{s_0, s_1, s_2, s_3\big\}$ where the s_i's are the identity and spin channels given by Eq. (1.25) we obtain the following theorem:

Theorem 2.4.4: The set of symmetric unitary channels is given by $U_s = \big\{rVr^t \big|\ r \in SO(3)\big\}$.

Proof: We begin by noting that $rs_ir^t \in SO(3)$ for all i, and

$$\left(rs_ir^t\right)^t = rs_i^tr^t = rs_ir^t. \tag{2.91}$$

Consequently, every element of U_s is a symmetric unitary channel. Conversely, if f is a symmetric unitary channel, then by the Spectral Theorem it is diagonalizable by some rotation r. That is, $r^tfr = \lambda$ is a diagonal member of $SO(3)$. This implies that λ is either the identity or a spin channel. Therefore, $f = r\lambda r^t \in U_s$.

□

Theorem 2.4.5: The simple functions generated by the set of symmetric unitary channels $G(U_s)$ are the extreme points of $\mathbb{E}_s$.

Proof: We begin by showing that every symmetric unitary channel generates an extreme point. Let

$$E_{rs_ir^t} = pE_f + (1-p)E_g \tag{2.92}$$

where $p \in (0,1)$. We then have that

$$\langle x, [rs_ir^t - pf - (1-p)g]x\rangle = 0 \tag{2.93}$$

for all $x \in S^2$. This is equivalent to requiring that the matrix $[rs_ir^t - pf - (1-p)g]$ is skew-symmetric. However, since rs_ir^t, f, and g are all symmetric, it follows that our matrix is both symmetric and skew-symmetric; which occurs if and only if it is the zero matrix. That is, we must have that $rs_ir^t = pf + (1-p)g$. Furthermore, it is shown in [21] that unitary channels are extreme points in $\mathcal{U}$. Consequently, $rs_ir^t = f = g$, which implies $E_{rs_ir^t} = E_f = E_g$. We are then left to conclude that $E_{rs_ir^t}$ is an extreme point of $\mathbb{E}_s$.

Conversely, let E_f be an extreme point. Since we can assume f is symmetric, there exists a rotation r such that $E_f = E_{r\lambda r^t}$ where λ is a diagonal unital channel. We recall from Proposition 1.5.2, that a diagonal matrix is unital if and only if it is a convex sum of the elements of V. Therefore,

$$E_f = E_{r\lambda r^t} = E_{\sum_i p_i r s_i r^t} = \sum_{i=0}^{3} p_i E_{r s_i r^t} \tag{2.94}$$

where $p \in \Delta^4$. Then since E_f is an extreme point, it follows that $E_f = E_{r s_i r^t}$ for some i, and the proof is complete.

□

With Theorem 2.4.5 we have now further characterized the set of simple functions generated by the symmetric unitary channels. Moreover, we note that for every $r \in SO(3)$ the subset $G\big(rVr^t\big)$ with the binary operation $*$ where $E_f * E_g = E_{fg}$ forms a representation of the *Klein four group* V defined at the beginning of Section 1.1.2. Furthermore, each convex closure $\big\langle G(rVr^t)\big\rangle$ forms a convex monoid of commutative error functions that is isomorphic to the free convex monoid over V, as shown in the following theorem.

Theorem 2.4.6: The convex closure $\langle V\rangle$ is isomorphic to $\big\langle G(rVr^t)\big\rangle$ for each $r \in SO(3)$.

Proof: It is shown in [24] that for every rotation r the convex closure $\langle rVr^t\rangle$ is isomorphic to the free convex monoid $\langle V\rangle$. Then the proof is complete by showing that G is an isomorphism, i.e. a convex, continuous injection on $\langle r\mathcal{S}r^t\rangle$.

From Lemma 2.4.2, we know that G is a convex, continuous function. We now prove that G is also injective, and therefore an isomorphism. Let f and g be symmetric unital channels, then

$G(f) = G(g)$ if and only if

$$\langle x, fx\rangle = \langle x, gx\rangle \tag{2.95}$$

$$\iff \langle x, [f-g]x\rangle = 0. \tag{2.96}$$

Since f and g are both symmetric, then by the same arguments in Theorem 2.4.5, it follows that $f - g$ is necessarily the zero matrix. Consequently, $f = g$ and we are left to conclude that G is an isomorphism.

□

Mathematically, for each rotation r the convex closure $\langle rVr^t\rangle$ is a representation of the free affine monoid, while physically, this same object is the given by the set of teleportation channels [24]. Although Theorem 2.4.6 appears to apply to only a particular subset of error functions, this is quite deceptive. We demonstrate this in step-by-step fashion:

If we begin with an arbitrary unital channel f, then from Theorem 2.3.1 we may replace the simple function E_f with that generated by $\varphi(f)$. Furthermore, since $\varphi(f)$ is symmetric it can be diagonalized such that $\varphi(f) = r\lambda r^t$, where r is some rotation and λ is a diagonal matrix. Utilizing [21], in which it is shown that a diagonal matrix is unital only if it is a convex sum of the identity and the spin channels, we then have that $\varphi(f) = \sum_{i=0}^{3} p_i r s_i r^t$ for some $p \in \Delta^4$. Therefore, for an arbitrary unital channel f, the simple function E_f has the same image as

$$\sum_{i=0}^{3} p_i E_{rs_ir^t} \in \langle G(rVr^t)\rangle. \tag{2.97}$$

Now that we have established some of the more important properties of error functions and the function space $\mathbb{E}_s$, we are prepared to investigate the physical application of our results. In the

next chapter we introduce a recently patented process that reduces the rate of error in quantum communication called *Adaptive Quantum Information Processing*; patent number US9838141 B2. We illustrate this process and its improvement via two well-known qubit communication protocols: the Prepare and Send Protocol and Quantum Key Distribution.

2.5 Summary

The main contribution of this chapter was the introduction of the error functions for qubit communication protocols. Furthermore, we were able to show that under certain assumptions the image of each one of these functions is obtained via a simple function generated by a systematically constructed diagonal channel. While all these results are abstract in conception, their application to physical protocols is far reaching and will now be the center of our attention.

CHAPTER 3: ADAPTIVE QUANTUM INFORMATION PROCESSING

3.1 The Adaptive Scheme

In quantum information an increase in the number of errors requires an increase in the number of qubits in order to successfully transmit information. This results in larger overhead and less efficiency for a given communication procotol. It is for this reason that the reduction of the error rate for a qubit communication protocol is of interest. Specifically, in Quantum Key Distribution, the transmitted information requires a key of the same size in order to encrypt a message. Therefore, reducing the error rate can speed up key generation [1,14].

As described in [21], *Adaptive Quantum Information Processing*, patent number US9838141 B2, aims to reduce the error rate associated with a quantum communication protocol. The scheme is as follows:

(i) Perform channel tomography in order to determine the qubit channel f.

(ii) Bob calculates the scope $s(f)$ and the eigenvector x that obtains its maximum.

(iii) Alice and Bob then engage in their protocol utilizing x.

Of course, the scheme depends on the protocol in question. For that reason, the remainder of this chapter is dedicated to the implementation of Adaptive Quantum Information Processing in the Prepare and Send Protocol and Quantum Key Distribution. We first wish to establish the bounds for the extent to which we may reduce the error rate for an arbitrary protocol.

We remind the reader that due to Theorem 2.3.1, any unital channel f generates the same error function as $\varphi(f) = \frac{1}{2}\big(f + f^t\big)$. Consequently, while each theorem will be a statement about a

potentially non-symmetric channel f, the conditions for all results will be dependent upon the symmetric matrix $\varphi(f)$. We consider non-symmetric channels in order to ensure a practical discussion. That is, while theoretical results allow us to consider only symmetric channels, nature is surely not expected to obey such restrictions. Therefore, it is important to keep the narrative straight in our mind and remember that we are using $\varphi(f)$ as a tool to uncover results about the original channel f. Lastly, we note that all results in this chapter consider protocols in which Theorem 2.3.6 is applicable. That is, all protocols will be assumed to have ideal QKD-like bases where each possible representation is the component of some $x \in X$.

Theorem 3.1.1: The largest possible reduction of the error rate for a general qubit communication protocol with unital noise f is

$$\tfrac{1}{2}(\lambda_1 - \lambda_3)(a - c), \tag{3.1}$$

where $\lambda_3 \leq \lambda_2 \leq \lambda_1$ are the eigenvalues of $\varphi(f)$ and a, b, and c are as defined in Theorem 2.3.6.

Proof: Since the image of an arbitrary error function is given by the interval

$$\Big[\min(E_{q,f}), \max(E_{q,f})\Big], \tag{3.2}$$

the largest possible improvement is given by $\max(E_{q,f}) - \min(E_{q,f})$. Simply put, the most one can reduce the error rate for an arbitrary qubit communication protocol is the length of the interval given by $\mathrm{Im}(E_{q,f})$. Then since we have shown in Theorem 2.3.6 that

$$\max_{x \in X}\big(E_{q,f}\big) = \frac{1}{2}\Big[1 - c\lambda_1 - b\lambda_2 - a\lambda_3\Big] \tag{3.3}$$

and

$$\min_{x \in X}\big(E_{q,f}\big) = \frac{1}{2}\Big[1 - a\lambda_1 - b\lambda_2 - c\lambda_3\Big], \tag{3.4}$$

it follows by subtracting Eq. (3.4) from (3.3) that the largest reduction in the rate of error is

$$\tfrac{1}{2}(\lambda_1 - \lambda_3)(a - c). \tag{3.5}$$

□

This is not to suggest that for any initial choice of bases one may reduce the rate of errors by $\frac{1}{2}(\lambda_1 - \lambda_3)(a - c)$. For example, if the initial choice of bases gives us an error rate A where

$$A < \max(E_{q,f}), \tag{3.6}$$

then the most one may reduce the initial error rate by is

$$A - \min(E_{q,f}) < \tfrac{1}{2}(\lambda_1 - \lambda_3)(a - c). \tag{3.7}$$

Therefore, the largest possible improvement is obtainable if and only if the initial choice of bases results in the maximum value of $E_{q,f}$.

Corollary 3.1.2: A given protocol has a constant error rate if and only if $a = b = c = \frac{1}{3}$ or $f = pI$ where $p \in [0, 1]$.

Proof: If the error rate is constant, then Eq. (3.1) must be equal to zero. That is,

$$\tfrac{1}{2}(\lambda_1 - \lambda_3)(a - c) = 0. \tag{3.8}$$

This occurs if and only if $\lambda_1 = \lambda_3$ or $a = c$; or equivalently, when $f = pI$ with $p \in [0, 1]$ or $a = b = c = \frac{1}{3}$ respectively.

□

With the bounds of improvement established for a general protocol, we now investigate the implementation of Adaptive Quantum Information Processing in the Prepare and Send Protocol.

3.2 Adaptive Prepare and Send Protocol

In the Prepare and Send Protocol Alice and Bob use a single basis in which they prepare and measure. Therefore, the implementation of our adaptive scheme is as follows: After Bob calculates the eigenvector x, described in step (ii), he would then simply engage in the Prepare and Send Protocol with Alice using the x communication basis to prepare and measure their qubits. The question then remains, "What does the maximum improvement look like for the Prepare and Send Protocol?"

Theorem 3.2.1: The most one may reduce the error rate for unital noise f in the Adaptive Prepare and Send Protocol is

$$\tfrac{1}{2}(\lambda_1 - \lambda_3), \tag{3.9}$$

where $\lambda_3 \leq \lambda_2 \leq \lambda_1$ are the eigenvalues of $\varphi(f)$.

Proof: As the maximum improvement for a general protocol is $\frac{1}{2}(\lambda_1 - \lambda_3)(a - c)$, we must first find the values of a and c, which are characterized by the protocol itself. In the Prepare and Send Protocol Alice and Bob encode their qubits in only one basis. Therefore, $n = q = 1$ and our p_j functions, defined by

$$p_j(x) = \sum_{i=1}^{n} q_i (x_i)_j^2, \tag{3.10}$$

reduce to

$$p_j = (x)_j^2. \tag{3.11}$$

It then follows that since $x \in S^2$, we have $a = \max(p_j) = 1$ for each p_j; which directly implies that $c = 0$. Plugging these values into Eq. (3.1), we then have that the largest possible reduction for the Prepare and Send Protocol is

$$\tfrac{1}{2}(\lambda_1 - \lambda_3). \tag{3.12}$$

□

Remark: While the improvement for the Adaptive Prepare and Send Protocol given in Eq. (3.12) is more easily obtained by calculating the difference

$$\max\left(E_f\right) - \min\left(E_f\right) = \tfrac{1}{2}(1 - \lambda_3) - \tfrac{1}{2}(1 - \lambda_1) = \tfrac{1}{2}(\lambda_1 - \lambda_3), \tag{3.13}$$

we provided a proof that utilizes the result for the improvement of a general protocol (Theorem 3.1.1). That is, we chose to prove this result by simplifying Eq. (3.1) as it follows a systematic approach that can be applied to any protocol without prior knowledge of the explicit values for the range of possible error rates.

We have then shown that the possible improvement is solely dependent on the eigenvalues of $\varphi(f)$ in the Prepare and Send Protocol! Moreover, we may further simplify Eq. (3.12) for the case of a unitary channel, and characterize the channels that generate a simple function with the image $[0, 1]$.

Theorem 3.2.2: For unitary noise $f \in SO(3)$, the largest improvement for the Adaptive Prepare and Send Protocol is $\sin^2 \frac{\theta}{2}$ where θ is the angle of rotation of f.

Proof: In the Bloch representation, every unitary channel is characterized by a member of $SO(3)$.

Therefore, f has the following eigenvalues and eigenvectors:

$$fv_1 = v_1, \tag{3.14}$$

$$fv_2 = e^{i\theta}v_2, \tag{3.15}$$

$$fv_3 = e^{-i\theta}v_3. \tag{3.16}$$

It then follows that multiplying each side by f^t we obtain

$$f^t v_1 = v_1, \tag{3.17}$$

$$f^t v_2 = e^{-i\theta}v_2, \tag{3.18}$$

$$f^t v_3 = e^{i\theta}v_3. \tag{3.19}$$

Finally, averaging Eqs. (3.14-3.16) with Eqs. (3.17-3.19), we obtain the eigenvalue equations for $\varphi(f)$:

$$\varphi(f)v_1 = v_1, \tag{3.20}$$

$$\varphi(f)v_2 = \cos\theta\ v_2, \tag{3.21}$$

$$\varphi(f)v_3 = \cos\theta\ v_3. \tag{3.22}$$

In other words, if $f \in SO(3)$, then $\sigma\big(\varphi(f)\big) = \{1, \cos\theta, \cos\theta\}$. Plugging these eigenvalues into Eq. (3.12), we then have that the largest possible improvement is

$$\tfrac{1}{2}(1 - \cos\theta) = \sin^2\left(\tfrac{\theta}{2}\right). \tag{3.23}$$

□

Corollary 3.2.3: For unitary noise $f \in SO(3)$,

$$\text{Im}(E_f) = \left[0, \sin^2(\tfrac{\theta}{2})\right]. \tag{3.24}$$

Proof: This is a direct consequence of Theorems 3.2.1 and 3.2.2.

□

Theorem 3.2.4: The image of a simple function E_f is the unit interval if and only if f is a non-trivial involutive rotation, where by non-trivial we mean $f \neq I$.

Proof: If the image of E_f is the unit interval $[0, 1]$, then the maximum improvement, given by $\max\big(E_f\big) - \min\big(E_f\big)$, is 1. That is, by Theorem 3.2.1, $\frac{1}{2}(\lambda_1 - \lambda_3) = 1$. It then follows that the $\text{Im}\big(E_f\big) = [0, 1]$ if and only if $\lambda_1, \lambda_3 \in \big\{+1, -1\big\}$. Furthermore, since $\varphi(f)$ is symmetric its eigenvalues must satisfy the following inequality:

$$1 - \lambda_1 - \lambda_2 + \lambda_3 \geq 0 \tag{3.25}$$

[21]. Therefore, $\lambda_2 \leq -1$. However, since $|\lambda_i| \leq 1$ for all i, it then follows that $\lambda_2 = -1$. In other words, $\sigma\big(\varphi(f)\big) = \{1, -1, -1\}$. Consequently, $\varphi(f) = rs_1r^t$ for some $r \in SO(3)$. Finally, since each unitary channel is an extreme point of $\mathcal{U}$, we have that $f = rs_1r^t$.

□

In summary, when implementing Adaptive Quantum Information Processing in the Prepare and Send Protocol the largest one may reduce the rate of error for unital noise is $\frac{1}{2}(\lambda_1 - \lambda_3)$. If the environment is instead described by a unitary channel, then the largest reduction is $\sin^2 \frac{\theta}{2}$. These results directly imply that the image of a simple function is the unit interval if and only if its associated channel is a non-trivial involutive unitary channel. We now turn our focus to the

implementation of Adaptive Quantum Information Processing in QKD.

3.3 Adaptive Quantum Key Distribution

While QKD is appealing since it provides a means of "detecting" an eavesdropper by examining the resulting error rate, there is a glaring concern that results from this security measure. What if the environment itself introduces a large error in Bob's measurements? This could result in any number of false abortions and drastically slow down, or even prevent the use of QKD! Therefore, minimizing the rate of error due to the environment is of great interest. Again, we remind the reader that these results assume no restrictions on the bases besides that they be orthogonal (i.e. Theorem 2.3.6 applies).

Theorem 3.3.1: The most one may reduce the error rate for unital noise in the Adaptive QKD is

$$\tfrac{1}{4}(\lambda_1 - \lambda_3). \tag{3.26}$$

Proof: From Theorem 3.1.1, we know the largest possible reduction of the error rate for a general protocol with unital noise f is

$$\tfrac{1}{2}(\lambda_1 - \lambda_3)(a - c). \tag{3.27}$$

Recall that QKD is characterized by the set $X = \left\{(x, y) \in S^2 \times S^2 \mid \langle x, y\rangle = 0\right\}$ and the point $q = (\frac{1}{2}, \frac{1}{2})$. Then by Theorem 2.3.6, it follows that

$$a = \max_{x,y}\left(q_1 x_j^2 + q_2 y_j^2\right) = \max_{x,y}\left(\frac{x_j^2 + y_j^2}{2}\right). \tag{3.28}$$

Therefore, in order to find the value of a we must first maximize the sum

$$x_j^2 + y_j^2, \tag{3.29}$$

where x and y are orthogonal unit vectors in $\mathbb{R}^3$. In order to find the maximum of Eq. (3.29), we note that for every pair of orthogonal vectors $x, y \in S^2$ there exists a rotation r where $x = re_j$ and $y = re_k$. Consequently,

$$x_j^2 + y_j^2 = \langle e_j, x\rangle^2 + \langle e_j, y\rangle^2 = \langle e_j, re_j\rangle^2 + \langle e_j, re_k\rangle^2 \tag{3.30}$$

$$= r_{jj}^2 + r_{jk}^2 \leq 1. \tag{3.31}$$

To this end, we have that

$$a = \max_{x,y}\left(\frac{x_j^2 + y_j^2}{2}\right) = \tfrac{1}{2}. \tag{3.32}$$

Lastly, plugging our value for a into Eq. (2.43), we see that $b = \frac{1}{2}$ and $c = 0$. Consequently, we are left to conclude that

$$\mathrm{Im}\big(E_{q,f}\big) = \left[\tfrac{1}{2}\Big(1 - \tfrac{1}{2}\lambda_1 - \tfrac{1}{2}\lambda_2\Big), \tfrac{1}{2}\Big(1 - \tfrac{1}{2}\lambda_2 - \tfrac{1}{2}\lambda_3\Big)\right], \tag{3.33}$$

and therefore, the largest one may reduce the rate of error in adaptive QKD is

$$\max(E_{q,f}) - \min(E_{q,f}) = \tfrac{1}{4}(\lambda_1 - \lambda_3). \tag{3.34}$$

□

Remark: As shown in Theorem 2.3.7, this interval is also obtained by the simple function that is

generated by the following diagonal unital channel:

$$g = \begin{bmatrix} \eta_1 & 0 & 0 \\ 0 & \eta_1 & 0 \\ 0 & 0 & \eta_3 \end{bmatrix} \tag{3.35}$$

where $\eta_1 = \frac{1}{2}(\lambda_1 + \lambda_2)$ and $\eta_3 = \frac{1}{2}(\lambda_2 + \lambda_3)$.

Theorem 3.3.2 For the case of unital noise, the error rate can not be reduced to less than half of the initial value in Adaptive QKD.

Proof: We begin by showing that $\frac{1}{2}\max\big(E_{q,f}\big) \leq \min\big(E_{q,f}\big)$. Recalling that $1-\lambda_1-\lambda_2+\lambda_3 \geq 0$, we have that

$$\begin{aligned} &\tfrac{1}{2}\max\big(E_{q,f}\big) - \min\big(E_{q,f}\big) \\ &\quad = \tfrac{1}{4}\Big(1 - \tfrac{1}{2}\lambda_2 - \tfrac{1}{2}\lambda_3\Big) - \tfrac{1}{2}\Big(1 - \tfrac{1}{2}\lambda_1 - \tfrac{1}{2}\lambda_2\Big) \\ &\quad = \tfrac{1}{4}\big(\lambda_1 - 1\big) + \tfrac{1}{8}\big(\lambda_2 - \lambda_3\big) \leq \tfrac{1}{8}\big(\lambda_1 - 1\big) \leq 0. \end{aligned} \tag{3.36}$$

Therefore,

$$\tfrac{1}{2}\max\big(E_{q,f}\big) \leq \min\big(E_{q,f}\big). \tag{3.37}$$

It then follows that if the initial error rate is given by A, and we assume there exists a change of bases that results in an error rate $B < \frac{1}{2}A$, then

$$B < \tfrac{1}{2}A \leq \tfrac{1}{2}\max\big(E_{q,f}\big) \leq \min\big(E_{q,f}\big). \tag{3.38}$$

Consequently, $B \notin \mathrm{Im}\big(E_{q,f}\big)$, and we have a contradiction.

□

Thus, we have shown that the best we can do is cut the error rate in half. With this in mind, we now establish the conditions in which this is possible for Adaptive QKD.

Theorem 3.3.3: The error rate for QKD can be reduced to half its initial value in Adaptive Quantum Information Processing if and only if $fe_2 = e_2$.

Proof: We begin by noting that if f is the identity, then it would fix every point and $E_{q,f}$ would be the constant function 0. Thus our result is trivially true when $f = I$. Let us now consider when f is not the identity.

It was shown in Theorem 3.3.2 that we can cut the error rate in half if and only if the initial error rate is equal to $\max\left(E_{q,f}\right)$ and $\min\left(E_{q,f}\right) = \frac{1}{2}\max\left(E_{q,f}\right)$. With this in mind, we prove our result by first showing that $\min(E_{q,f}) = \frac{1}{2}\max(E_{q,f})$ if and only if f fixes some unit vector x^*. The proof is then complete by showing that if $\min(E_{q,f}) = \frac{1}{2}\max(E_{q,f})$ and the initial error rate is $\max(E_{q,f})$, then $fe_2 = e_2$, and vice versa.

Assume that $\min(E_{q,f}) = \frac{1}{2}\max(E_{q,f})$. It can then be shown by combining like terms that this is equivalent to

$$1 - \lambda_1 = \tfrac{1}{2}(\lambda_2 - \lambda_3) \tag{3.39}$$

where $\lambda_3 \leq \lambda_2 \leq \lambda_1$ are the eigenvalues for $\varphi(f)$. Furthermore, since $\varphi(f)$ is a symmetric unital channel, its eigenvalues must satisfy the following inequality as shown in [21]:

$$1 - \lambda_1 \geq \lambda_2 - \lambda_3. \tag{3.40}$$

Therefore, by Eq. (3.39) and inequality (3.40), we are left to conclude that $\min\limits_{x\in X}\left(E_{q,f}\right) = \frac{1}{2}\max\limits_{x\in X}\left(E_{q,f}\right)$ if and only if

$$\tfrac{1}{2}(1 - \lambda_1) \geq (1 - \lambda_1). \tag{3.41}$$

However, since $|\lambda_i| \leq 1$ for all i, it then follows that inequalitiy (3.41) is true if and only if $\lambda_1 = 1$. That is, there exists some unit vector x^* such that $\varphi(f)x^* = x^*$. Moreover, since x^* is an extreme point of $\mathbb{B}^3$ it follows that $fx^* = x^*$. Additionally, since $f \neq I$, it follows that λ_2 and λ_3 are each less than 1 and satisfy the following inequalities:

$$1 - \lambda_1 - \lambda_2 + \lambda_3 \geq 0 \tag{3.42}$$

$$1 - \lambda_1 + \lambda_2 - \lambda_3 \geq 0, \tag{3.43}$$

[21]. Therefore, because $\lambda_1 = 1$, inequalities (3.42) and (3.43) are both true if and only if $\lambda_2 = \lambda_3$. In other words, $\min\limits_{x\in X}\big(E_{q,f}\big) = \frac{1}{2}\max\limits_{x\in X}\big(E_{q,f}\big)$ if and only if $fx^* = x^*$ for some $x^* \in S^2$ and $\sigma(\varphi(f)) = \{1, \lambda_2, \lambda_2\}$. Lastly, we show that if $\min\big(E_{q,f}\big) = \frac{1}{2}\max\limits_{x\in X}\big(E_{q,f}\big)$ and the initial choice of bases have an error rate of $\max\limits_{x\in X}(E_{q,f})$, then $x^* = e_2$.

In QKD it is conventional to use the bases e_1 and e_3. Consequently, it is safe to assume that when performing Adaptive QKD the initial error rate would be given by $E_{q,f}(e_1, e_3)$. With this last assumption, the largest possible improvement is obtainable if and only if

$$E_{q,f}(e_1, e_3) = \max_{x\in X}\big(E_{q,f}\big). \tag{3.44}$$

In other words, since $\lambda_2 = \lambda_3$, the initial error rate is $\max\limits_{x\in X}\big(E_{q,f}\big)$ if and only if

$$\tfrac{1}{2}\Big(1 - \tfrac{1}{2}\langle e_1, \varphi(f)e_1\rangle - \tfrac{1}{2}\langle e_3, \varphi(f)e_3\rangle\Big) = \tfrac{1}{2}\big(1 - \tfrac{1}{2}\lambda_2 - \tfrac{1}{2}\lambda_3\big) = \tfrac{1}{2}\big(1 - \lambda_2\big). \tag{3.45}$$

It then directly follows that Eq. (3.45) is equivalent to

$$\tfrac{1}{2}\langle e_1, \varphi(f)e_1\rangle + \tfrac{1}{2}\langle e_3, \varphi(f)e_3\rangle = \lambda_2. \tag{3.46}$$

Furthermore, because $\min_{x\in S^2}\langle x, \varphi(f)x\rangle = \lambda_2$, we have that

$$\langle e_1, \varphi(f)e_1\rangle = \langle e_3, \varphi(f)e_3\rangle = \lambda_2. \tag{3.47}$$

In other words,

$$\langle e_1, \varphi(f)e_1\rangle = \varphi(f)_{11} = \tfrac{1}{2}\big(f_{11} + f_{11}^t\big) = f_{11} = \lambda_2 \tag{3.48}$$

and similarly, $f_{33} = \lambda_2$. It then follows that

$$\text{tr}\big(\varphi(f)\big) = \text{tr}\big(f\big) = f_{11} + f_{22} + f_{33} = f_{22} + 2\lambda_2 = 1 + 2\lambda_2, \tag{3.49}$$

and we finally have that $f_{22} = 1$. Lastly, since all unital channels are nonexpansive, we have that $fe_2 = \lambda_1 e_2 = e_2$, and we are left to conclude that the error rate can be cut in half if and only if $fe_2 = \lambda_1 e_2 = e_2$.

□

3.4 Summary

In this chapter we applied results developed in the previous chapter to establish the largest on may reduce the rate of error when using Adaptive Quantum Information Processing. In particular, we have shown that when performing Adaptive QKD the most one can reduce the error rate is by half of its initial value. Furthermore, instances in which maximum improvement occur are characterized by nontrivial noise that fixes the e_2 basis. With this result in mind, the next step is to construct a physically significant channel that satisfies these conditions. Chapter 4 is a discussion of the physics for environmental noise due to gravity on an Earth orbiting qubit. We will explicitly state our set-up and show that it results in a unitary channel that satisfies Theorem 3.3.3. This

example is significant as there exists theoretical results that have been experimentally verified under comparable circumstances for the case of Earth to satellite QKD [3,34].

CHAPTER 4: GRAVITATIONAL NOISE

In this chapter we introduce the mathematics and physics needed to describe gravitational effects on qubits. The first four sections are a summary of the work by Lanzagorta [16] while the last two sections will examine the extent to which gravity affects a qubit, which we will henceforth refer to as *gravitational noise*. Lastly, we examine gravitational noise in Adaptive QKD. While this chapter is meant to introduce the origins of relativistic effects on qubits, for the sake of brevity it will fall short of being fully comprehensive. Therefore, citations will supplement most equations while the exposition will primarily serve as a guide for the overall narrative.

Many models in physics consist of a mathematical space that represents the possible states of some system and a set of mappings that describe physically realizable changes. In these models invariant quantities arise from certain mappings that preserve the form of the equations in physics, i.e. symmetries [28]. One such example is invariance under spatial translations which results in the conservation of linear momentum. We begin our discussion by introducing the space used to characterize position and time in relativistic theory. Those familiar with differential geometry may skip this section and simply keep in mind that spacetime is defined as a four-dimensional smooth Lorentzian manifold with signature $(1, 3)$.

4.1 Preliminaries: General Relativity

4.1.1 Spacetime

While this section may feel convoluted due to notation and the quantity of concepts, the overall motivation is very simple. In order to illustrate this, we first provide a discussion for the intuition or reasoning behind using differential geometry to study gravity.

General relativity is a non-Newtonian theory of gravity. Simply put, Newtonian gravity is described as a force exerted between bodies. Every body has an associated vector field dependent on the mass of the object that is used to calculate the force of gravity at each point in space. In this narrative forces between two bodies occur without any physical contact; a phenomenon referred to as "action at a distance". In turn, this means that bodies can exert forces on each other instantaneously. This is a concept with which many physicists disagree.

General relativity circumvents action at a distance by supposing that the energy of an object alters the geometry of both the space and time coordinates used to describe our system, referred to as *spacetime*. Under this assumption the gravitational force is no longer characterized by a vector field, but rather the geodesics of our "warped" spacetime. The problem then arises, "How do we define the geodesics and apply physics on a space that potentially does not have a Euclidean geometry?". This problem is resolved with the definition of a manifold.

A manifold is a locally Euclidean topological space. This means that at each point, if we stay in a small enough region, the space around us will look and behave as though it has a Euclidean geometry. A practical example of such is the surface of Earth. As we walk around on Earth, it appears as though we are dealing with a flat surface. However, when determining long distance trajectories, such as ballistic missiles, it becomes apparent that the surface is actually curved, and may be approximated as spherical. In such a case, we then discuss the physics of an object as it traverses long distances by considering the cumulative contributions of each locally Euclidean space. This is accomplished mathematically by modeling our system with a manifold.

Moreover, Newtonian mechanics deals with a non-negative metric, which means the distance between any two distinct points is positive. On the other hand, general relativity calls for a metric that may obtain negative values. This is due to Einstein's realization that the speed of light is invariant for inertial observers, which is best illustrated as follows:

Anyone moving at a constant velocity (inertial observer) will observe light to have the same speed. This is mathematically equivalent to requiring that the spacetime interval ds^2 between any two events be the same in each inertial frame. In turn, this condition is satisfied in $\mathbb{R}^4$ when the metric is Lorentzian. It is important to note that this physical assumption is disconnected from gravity, but rather a statement about how space and time are related to one another.

In summary, the speed of light is invariant for observers in constant motion and the energy of a system alters the geometry of space and time. This results in a theory that must account for a potentially non-trivial geometry that could easily complicate the physics. The solution is to model spacetime as a topological space that permits real analysis locally at each point and has a Lorentzian metric. That is, a smooth Lorentzian manifold.

Like our introduction of topology, this section is not meant to be a representation of differential geometry as a whole, and the proofs of many results will be left out for the sake of brevity. The entirety of this introduction is taken from [13], [26], and [36].

Definition 4.1.1 A Hausdorff topological space M is a n-dimensional *manifold* if it admits an open covering $\{U_\alpha\}$ such that each U_α is homeomorphic to an open subset of $\mathbb{R}^n$ via the functions h_α. We call $\{(U_\alpha, h_\alpha)\}$ the *atlas* of M, h_α the *charts*, and each pair (U_α, h_α) a *coordinate neighborhood* of M.

Since every h_α is a homeomorphism, we then have that each $p \in U_\alpha$ is determined by the n-tuple of real numbers $h_\alpha(p) = (x^1(p), ..., x^n(p))$, referred to as the set of *local coordinates* of the point p with respect to the coordinate neighborhood (U_α, h_α). Note, we index the components of $h_\alpha(p)$ with superscripts. The reasoning for this will be made clear later as we will require both subscripts and superscripts to specify different types of vectors.

Definition 4.1.2: An atlas is C^r if every *transition map* $h_\beta h_\alpha^{-1}$, is C^r on $h_\alpha\big(U_\alpha \cap U_\beta\big)$; that is all

derivates of order r and smaller exist and are continuous. In [36] Whitney shows that every C^r structure contains a C^∞ structure; therefore, we may exclusively consider C^∞ structures, which we will refer to as *smooth*. A *smooth manifold* is a second-countable Hausdorff manifold with a C^∞ atlas.

Let us explore these definitions via the following example from [26]. Consider the unit circle S^1 in $\mathbb{R}^2$ centered at the origin with the subspace topology $\tau_{S^1} = \big\{S^1 \cap U \mid U \in \tau\big\}$, where τ is the usual topology defined on $\mathbb{R}^2$. Then with the atlas $\big\{(U_1, \psi_1), (U_2, \psi_2), (V_1, \phi_1), (V_2, \phi_1)\big\}$ where

$$U_1 = \{(x, y) \in S^1 \mid y > 0\}, \tag{4.1}$$

$$V_1 = \{(x, y) \in S^1 \mid y < 0\}, \tag{4.2}$$

$$U_2 = \{(x, y) \in S^1 \mid x > 0\}, \tag{4.3}$$

$$V_2 = \{(x, y) \in S^1 \mid x < 0\}, \tag{4.4}$$

and

$$\psi_1(x, y) = x, \quad \phi_1(x, y) = x, \tag{4.5}$$

$$\psi_2(x, y) = y, \quad \phi_2(x, y) = y. \tag{4.6}$$

It then follows that S^1 is a 1-dimensional smooth manifold. This is easily verified as the maps ψ_i and ϕ_i are homeomorphisms, while the sets U_i and V_i are open. All that is left to show then is that the transition maps are C^∞.

The only transition maps that exist are those defined on the sets $U_1 \cap U_2$ and $V_1 \cap V_2$, which happen to be equal to the sets $\{(x, y) \in S^1 \mid x, y > 0\}$ and $\{(x, y) \in S^1 \mid x, y < 0\}$ respectively. Then $\psi_1(U_1 \cap U_2) = \psi_2(U_1 \cap U_2) = \{t \mid 0 < t < 1\}$. It then follows that the transition maps for the ψ_i functions are given by $\tau_{1,2} = \psi_2\big(\psi_1^{-1}(t)\big) = (1 - t^2)^{1/2}$ and $\tau_{2,1} = \psi_1\big(\psi_2^{-1}(t)\big) =$

$(1-t^2)^{1/2}$. Consequently, both transition maps are C^∞; the same can be shown for the transition maps generated by the ϕ_i charts. With a smooth manifold defined and an example given, we now wish to establish some significant properties that a function on a smooth structure can possess.

Definition 4.1.3: Let $f : M \to N$, where M and N are smooth manifolds. We say f is *smooth* if for the atlases $\{U_\alpha, h_\alpha\}$ and $\{V_\beta, g_\beta\}$ on M and N respectively, the maps $g_\beta f h_\alpha^{-1}$ are smooth wherever they are defined. If f is a smooth bijection that admits a smooth inverse, we call it a *diffeomorphism*.

In many branches of mathematics we define certain objects and study a special set of mappings that define when two such objects are equivalent. For example, in topology we have topological spaces and homeomorphisms; while in linear algebra we have vector spaces and linear isomorphisms. In differential geometry we have manifolds and diffeomorphisms. Succinctly put, each diffeomorphism is an equivalence relation between smooth structures.

Definition 4.1.4: Let $(a, b) \subseteq \mathbb{R}$. Then a differentiable map $\phi : (a, b) \to M$ is called a *differentiable curve* on the manifold M. Furthermore, if $\phi : (a, b) \to M$ is a differentiable curve, $t_0 \in (a, b)$, and $(x^1, ...x^n)$ is a local coordinate system of M in an open set of $\phi(t_0)$, then each

$$(\phi * x^i)(t) = \phi^i(t) \tag{4.7}$$

is a C^∞ function in the open set of t_0. The n-dimensional vector

$$\left(\left(\frac{d\phi^1}{dt} \right)\Big|_{t=t_0}, ..., \left(\frac{d\phi^n}{dt} \right)\Big|_{t=t_0} \right) \tag{4.8}$$

is called the *tangent vector to the curve* ϕ at $\phi(t_0)$ with respect to the local coordinate system $(x^1, ..., x^n)$.

From now on we shall denote the set of all real-valued C^∞ functions defined on some open set of p by $\mathcal{F}(p)$. Note, if $f, g \in \mathcal{F}(p)$, then $f + g$ and fg are defined on the intersection of the open sets U and V in which f and g are defined respectively. Furthermore, for any real number λ, λf is defined on U.

Definition 4.1.5: A map $v : \mathcal{F}(p) \to \mathbb{R}$ is called a *tangent vector* of M at p if for each $f, g \in \mathcal{F}(p)$ and $\lambda, \mu \in \mathbb{R}$ the following are true:

(i) $v(\lambda f + \mu g) = \lambda v(f) + \mu v(g)$

(ii) $v(fg) = v(f)g(p) + f(p)v(g)$.

Note, for tangent vectors v and w of M at p and $\lambda \in \mathbb{R}$, both $v + w$ and λv are also tangent vectors. Therefore, the set of all tangent vectors at p forms a vector space that we denote by $T_p(M)$. It can be shown that the dimension of the space $T_p(M)$ is the same as the dimension of M itself; and letting $(x^1, ..., x^n)$ be a local coordinate system on U, then at a point $p \in U$, the set

$$\left\{ \left(\frac{\partial}{\partial x^1}\right)_p, ..., \left(\frac{\partial}{\partial x^n}\right)_p \right\} \tag{4.9}$$

is a basis for $T_p(M)$.

This structure allows us to work in the vicinity of a point on a potentially non-Euclidean manifold. So we now have a means of discussing vector quantities and performing real analysis, which has been fundamental since the birth of Newtonian physics. However, as we have discussed earlier, it is possible to have multiple local coordinate systems at p, which means multiple bases of the form Eq. (4.9). Therefore, it is of interest to obtain the transformation rules between local coordinate systems.

For each $v \in T_p(M)$ we denote the components of v with respect to the local coordinate system

$(x^1, ..., x^n)$ by $v[x]$. Then supposing there exists a second local coordinate system $(\tilde{x}^1, ..., \tilde{x}^n)$, we wish to establish the transformation rules between $v[x]$ and $v[\tilde{x}]$. Recalling that both $\left\{ \left(\frac{\partial}{\partial x^1}\right)_p, ..., \left(\frac{\partial}{\partial x^n}\right)_p \right\}$ and $\left\{ \left(\frac{\partial}{\partial \tilde{x}^1}\right)_p, ..., \left(\frac{\partial}{\partial \tilde{x}^n}\right)_p \right\}$ form a basis of $T_p(M)$, we have that

$$\sum_{i=1}^{n} v[x]^i \left(\frac{\partial}{\partial x^i}\right)_p = v = \sum_{j=1}^{n} v[\tilde{x}]^j \left(\frac{\partial}{\partial \tilde{x}^j}\right)_p . \tag{4.10}$$

It can then be shown that the basis elements transform according to

$$\left(\frac{\partial}{\partial \tilde{x}^i}\right)_p = \sum_{j=1}^{n} \frac{\partial x^j}{\partial \tilde{x}^i}(p) \left(\frac{\partial}{\partial x^j}\right)_p \tag{4.11}$$

$$\left(\frac{\partial}{\partial x^i}\right)_p = \sum_{j=1}^{n} \frac{\partial \tilde{x}^j}{\partial x^i}(p) \left(\frac{\partial}{\partial \tilde{x}^j}\right)_p , \tag{4.12}$$

while the components of v satisfy the transformation rules

$$v[\tilde{x}]^i = \sum_{j=1}^{n} \frac{\partial \tilde{x}^j}{\partial x^i}(p) v[x]^j \tag{4.13}$$

$$v[x]^i = \sum_{j=1}^{n} \frac{\partial x^j}{\partial \tilde{x}^i}(p) v[\tilde{x}]^j . \tag{4.14}$$

We refer to vectors that transform according to Eqs. (4.11-4.12) as covariant vectors and denote such with subscripts; those that satisfy Eqs. (4.13-4.14) are called contravariant vectors and are denoted by superscripts. Due to the similarities between the transformation rules of contravariant vectors and the vectors of physical quantities in classical mechanics, physicists adopt the convention of using contravariant vectors to represent physically measurable quantities, such as position or momentum. In order to formalize the lengths between points on a manifold, we now examine the dual space of $T_p(M)$ which we will denote by $T_p^*(M)$.

Definition 4.1.6: Let $f \in \mathcal{F}(p)$ be defined on U. For an arbitrary $v \in T_p(M)$, let

$$df_p(v) = v(f). \tag{4.15}$$

Then df_p is a linear function from $T_p(M)$ into $\mathbb{R}$ that satisfies properties (i) and (ii) of Definition 4.1.5. That is, $df_p \in T_p^*(M)$ and the function df_p is called the *differential* of f at p. Considering the differential of an element in a local coordinate system at p

$$(dx^j)_p\left(\frac{\partial}{\partial x^i}\right)_p = \delta_i^j \tag{4.16}$$

where

$$\delta_i^j = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{otherwise.} \end{cases} \tag{4.17}$$

Therefore, $\left\{(dx^1)_p, ..., (dx^n)_p\right\}$ forms a basis for $T_p^*(M)$, and given any df_p

$$df_p = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(p)(dx^i)_p. \tag{4.18}$$

With a basis defined on the dual space of $T_p^*(M)$ we can now introduce what is arguably the most important mathematical object in general relativity, the metric tensor.

Definition 4.1.7: Let M be a smooth manifold. A *metric tensor* g on M assigns to each point p of M a bilinear, symmetric, nondegenerate function $g(p) : T_p(M) \times T_p(M) \to \mathbb{R}$ called the *metric* such that $g(p)$ is a smooth function of p.

At any point p in a smooth manifold, with the metric tensor we now have a means of defining the length of a vector v. That is, using the differentials of a local coordinate system at p, it then follows

that

$$\begin{aligned}(dx^i)_p(v) &= (dx^i)_p \sum_{j=1}^{n} v[x]^j \left(\frac{\partial}{\partial x^j}\right)_p \\ &= \sum_{j=1}^{n} v[x]^j \frac{\partial x^i}{\partial x^j}(p) = \sum_{j=1}^{n} v[x]^j \delta^i_j = v[x]^i.\end{aligned} \tag{4.19}$$

Simply put, $(dx^i)_p(v)$, or dx^i for short, is the i^{th} component of v with respect to $(x^1, \dots, x^n)$. Consequently, we may write the length of a vector $v \in T_p(M)$ as

$$||v||^2 = \sum_{i=1}^{n} g_{ij}(p)(dx^i)_p(v)(dx^j)_p(v), \tag{4.20}$$

where $g_{ij}(p) = g(p)\Big(\left(\frac{\partial}{\partial x^i}\right)_p, \left(\frac{\partial}{\partial x^j}\right)_p\Big)$.

Notation 4.1.8: When an index appears as both a superscript and subscript it is assumed to be summed over, this is referred to as the Einstein summation convention. With this notation in mind, Eq.(4.20) becomes

$$||v||^2 = g_{ij}(p)(dx^i)_p(v)(dx^j)_p(v). \tag{4.21}$$

Allowing v to represent the vector between two points defined in the local tangent space of p, we now have a means of examining the distance between points. The distance between two points is referred to as the *spacetime interval* and is often denoted as ds^2. But what if the two points in question are separated far enough apart that there does not exist a local coordinate system that contains them both? In such a case we consider the contributions from each local coordinate system. That is, letting ϕ be a differentiable curve defined on (a, b) on a smooth manifold M with a metric tensor g, and v a tangent vector of ϕ at $\phi(t)$, then if $a < c < d < b$, it follows that

$$L(c, d) = \int_c^d ||v|| dt \tag{4.22}$$

is the length of ϕ between $\phi(c)$ and $\phi(d)$ where v depends on ϕ as described by Eq. (4.8). In other words, Eq. (4.22) is simply an integral of the contributions given by Eq. (4.21) at each point along the curve from $\phi(c)$ to $\phi(d)$. Consequently, we now have a means of calculating the distance between points that do not share a local coordinate system.

With a smooth manifold defined and the necessary structure established to perform vector calculus at each point, we end our brief introduction to the mathematics of spacetime with the following two definitions.

Definition 4.1.9: The signature of a metric $g_{ij}(p)$ is the triple (n_0, n_-, n_+), where n_0 is the number of 0s and $n_{\pm 1}$ is the number of ± 1s in $g_{ij}(p)$'s diagonal form. Since we defined a metric to be nondegenerate we have that $n_0 = 0$. Therefore, the signature of each metric is sufficiently characterized by the 2-tuple (n_-, n_+). If each metric $g_{ij}(p)$ has the same signature independent of p, then it is called the *signature* of the metric tensor g.

Definition 4.1.10: A *smooth Lorentzian* manifold is a smooth manifold with a metric tensor g of signature $(1, 3)$; g is referred to as a Lorentzian metric. One may equivalently impose the signature $(3, 1)$.

In general relativity we define a *spacetime* to be a real 4-dimensional smooth Lorentzian manifold, and each point p of the the manifold is referred to as an *event*. Each event consists of one time component and three spatial components. While this appears similar to Newtonian gravity, it is the very structure of a Lorentzian manifold that complicates things. Concisely put, Newtonian mechanics is performed on the manifold $\mathbb{R}^4$ with a metric of signature $(0, 4)$ where each chart is the identity, while general relativity allows wildly different manifolds with a metric of signature $(1, 3)$.

4.1.2 Isometries of Relativity

As we have previously discussed, a physical model typically consists of a space that describes the system and a set of mappings that preserve certain laws of physics. With spacetime described by a smooth Lorentzian manifold, we now turn our focus to the equations we wish to preserve in general relativity. We begin this discussion with what is known as the principle of general covariance.

Principle of General Covariance: A physical equation remains valid in the presence of gravity if the following two conditions are satisfied:

(1) The equation holds in the absence of gravity; that is, when the metric tensor g_{ij} reduces to the Minkowski metric

$$\eta_{ij} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.23}$$

(2) The equation is *general covariant*. That is, our equation is preserved under general spacetime coordinate transformations $x \to \tilde{x}$. This is important as the laws of physics should be invariant for particles at each point of spacetime.

In other words, our special set of mappings is characterized by the group of real invertible 4×4 matrices $GL(4, \mathbb{R})$ that preserves the spacetime interval $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$. However, this only describes global actions, or actions that span across the manifold itself. So what about the actions within the tangent spaces defined at each point? Again, this is important as we are using these tangent spaces to perform real analysis.

As we have briefly addressed, in every tangent space we must have that the spacetime interval $ds^2 = \eta_{ij} dx^i dx^j$ is invariant for each inertial observer, independent of their respective velocities.

Therefore, it can be shown that any local transformations must be of the form

$$\tilde{x}^i = \Lambda^i_j x^j + a^i, \tag{4.24}$$

where $\Lambda^i_j = \frac{\partial \tilde{x}^i}{\partial x^j}$ and a^i is a real constant four vector [16]. To simplify our equations, greek indices will correspond to the global coordinates of the manifold while latin indices are associated with the local coordinate systems defined at each point.

Transformations of this form are called *Poincare transformations*, and together they form the *Poincare group*. Upon inspection we see that the a^i component characterizes translations within the tangent space. In this work we will be considering the cases in which the a^i's vanish. While this might seem negligent, it has a purpose. Recalling that each tangent space $T_p(U)$ is only defined on a neighborhood U of the point p, translations could result in an event that is not contained in U. Therefore, we only consider the subgroup of transformations

$$\tilde{x}^i = \Lambda^i_j x^j, \tag{4.25}$$

called the *Lorentz group*. Having now defined the space and mappings used to discuss general relativity, we are ready to finally investigate gravitational noise in quantum information. We begin with the dynamics of massive bodies in free fall.

4.2 The Dynamics of Free Fall in General Relativity

The physical description of a free falling particle in general relativity begins with the *principle of equivalence*.

Principle of Equivalence: At every point p in an arbitrary gravitational field, it is possible to choose a locally inertial coordinate system such that, in a small region around p, the laws of motion take the same form as that in a space with no gravity.

We begin by considering a free massive particle moving much slower than the speed of light, c. Then Newtonian mechanics tells us that the equations of motion are

$$\frac{d^2x^\mu}{d\tau^2} = 0, \tag{4.26}$$

where in natural units $(c = 1)$ the *proper time* τ is defined in terms of the spacetime interval $ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu$ such that

$$\tau = \int_{\mathcal{P}} \sqrt{-ds^2} \tag{4.27}$$

and $\mathcal{P}$ is a time-like path with $ds^2 < 0$. If we then consider the presence of gravity, by Newtonian mechanics we would have that

$$\frac{d^2x^\mu}{d\tau^2} \neq 0. \tag{4.28}$$

However, by the principle of equivalence, there exists a local inertial coordinate system $\xi^a(x^\mu)$ such that in a small region around the particle

$$\frac{d^2\xi^\alpha}{d\tau^2} = 0. \tag{4.29}$$

Since our local coordinate system is dependent on the global coordinate system x^μ, we may rewrite our equations of motion as

$$\begin{aligned} \frac{d^2\xi^\alpha}{d\tau^2} &= \frac{d}{d\tau}\left(\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{dx^\mu}{d\tau}\right) \\ &= \frac{\partial\xi^\alpha}{\partial x^\mu}\frac{d^2x^\mu}{d\tau^2} + \frac{\partial^2\xi^\alpha}{\partial x^\mu \partial x^\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0. \end{aligned} \tag{4.30}$$

Defining the *affine connection*

$$\Gamma^{\alpha}_{\mu\nu} = \frac{\partial x^{\mu}}{\partial \xi^{\alpha}} \frac{\partial^2 \xi^{a}}{\partial x^{\mu} \partial x^{\nu}}, \tag{4.31}$$

the equations of motion for a free falling particle in the presence of gravity are then given by

$$\frac{d^2 x^{\alpha}}{d\tau^2} + \Gamma^{\alpha}_{\mu\nu} \frac{dx^{\mu}}{d\tau} \frac{dx^{\nu}}{d\tau} = 0. \tag{4.32}$$

In the next section we discuss the equation that gives a relativistic description for the evolution of a spin-$\frac{1}{2}$ massive particle. This is because we are not only interested in the possible trajectories of spin-$\frac{1}{2}$ qubits, but also the gravitational effects on their states as they travel along some path. However, before doing so we would like to make further use of our newly defined affine connection. If we consider some contravariant four-vector v^{μ}, then it can be shown that its standard derivative does not transform according to the covariant and contravariant rules for general coordinate transformations in Eqs. (4.11-4.14). That is,

$$\frac{\partial v^{\mu}}{\partial x^{\nu}} \rightarrow \frac{\partial \tilde{x}^{\mu}}{\partial x^{\lambda}} \frac{\partial x^{\rho}}{\partial \tilde{x}^{\nu}} \frac{\partial v^{\lambda}}{\partial x^{\rho}} + \frac{\partial^2 \tilde{x}^{\mu}}{\partial x^{\lambda} \partial x^{\rho}} \frac{\partial x^{\rho}}{\partial \tilde{x}^{\nu}} v^{\lambda}. \tag{4.33}$$

However, the *covariant derivative* of a contravariant four-vector

$$\frac{\mathcal{D} v^{\mu}}{\mathcal{D} x^{\nu}} = \frac{\partial v^{\mu}}{\partial x^{\nu}} + \Gamma^{\mu}_{\nu\rho} v^{\rho} \tag{4.34}$$

does adhere to our transformations rules:

$$\frac{\mathcal{D} v^{\mu}}{\mathcal{D} x^{\nu}} \rightarrow \frac{\partial \tilde{x}^{\mu}}{\partial x^{\lambda}} \frac{\partial x^{\rho}}{\partial \tilde{x}^{\nu}} \frac{\mathcal{D} v^{\lambda}}{\mathcal{D} x^{\rho}}. \tag{4.35}$$

Similarly, we define the covariant derivative for a covariant four-vector as:

$$\frac{\mathcal{D} v_{\mu}}{\mathcal{D} x^{\nu}} = \frac{\partial v_{\mu}}{\partial x^{\nu}} - \Gamma^{\mu}_{\nu\rho} v_{\rho}. \tag{4.36}$$

Note, in the absence of gravity, the affine connection $\Gamma^{\mu}_{\nu\rho}$ vanishes and our covariant derivative reduces to the standard derivative. The covariant derivative is necessary in order for the principle of general covariance to apply to any equation that is dependent on differentiation. Explicitly, Eq. (4.34) results in the equations of motion for a free falling particle in the presence of gravity when we take our four vector to be the position of the particle x^{μ} and take our covariant derivative with respect to the proper time τ. That is, using the covariant derivative, our equations of motion are

$$\frac{\mathcal{D}^2 x^{\alpha}}{\mathcal{D}\tau^2} = 0. \tag{4.37}$$

4.3 The Dirac Equation

For convenience we will not be using the Bloch representation of a qubit until the last two sections of this chapter. This is favorable as the physics is simplified in the Hilbert space representation for quantum states. Explicitly, we consider the quantum state of a massive spin-$\frac{1}{2}$ free particle, $\Psi_{p,\alpha}(x)$, where α denotes its spin degrees of freedom and p is its definite four-momentum.

In the absence of gravity, the Dirac equation is

$$(i\eta_{\mu\nu}\gamma^{\mu}\partial^{\nu} - m)\Psi_{p,\alpha}(x) = 0, \tag{4.38}$$

where the Dirac matrices are

$$\gamma^0 = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \quad \boldsymbol{\gamma} = \begin{bmatrix} 0 & -\boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{bmatrix} \tag{4.39}$$

such that I is the 2×2 identity and $\boldsymbol{\sigma} = [\sigma_1, \sigma_2, \sigma_3]^t$ [16].

Understandably, one might be tempted to simply replace the Minkowski metric in Eq. (4.38) with

some general metric $g_{\mu\nu}$, and the standard derivative with our covariant derivative. However, this would not prove fruitful. The downfall of this approach is two-fold. First off, the state $\Psi_{p,\alpha}(x)$ does not transform under members of $GL(4,\mathbb{R})$, but rather the Lorentz group. Therefore, since the covariant derivative is defined on objects that transform according to $GL(4,\mathbb{R})$ the term $\mathcal{D}^{\nu}\Psi_{p,\alpha}(x)$ is not well-defined. Additionally, the Dirac matrices γ^{μ} may depend on the spacetime coordinates; that is, $\gamma^{\mu}(x)$ [16].

We resolve these issues by using local inertial frames defined at each point of the global manifold. We define these local reference frames via tetrad fields $e^{\mu}_{\ a}(x)$ such that

$$g^{\mu\nu} = e^{\mu}_{\ a}(x)e^{\nu}_{\ b}(x)\eta^{ab}, \tag{4.40}$$

where $g^{\mu\nu}$ is the metric for the spacetime and η^{ab} is the Minkowski metric of the local inertial reference frame. Recall that the latin indices refer to the local coordinates while the greek indices denote the global system [13,33]. It is important to note that the tetrad field is a set of four covariant vector fields, and is therefore both a local and a global object that transforms under general coordinates as follows:

$$\tilde{e}^{a}_{\ \mu}(\tilde{x}) = \frac{\partial \tilde{x}^{\nu}}{\partial \tilde{x}^{\mu}} e^{a}_{\ \nu}(x) \tag{4.41}$$

$$\tilde{e}^{a}_{\ \mu}(\tilde{x}) = \Lambda^{a}_{\ b}(x)e^{b}_{\ \mu}(x), \tag{4.42}$$

where $\Lambda^{a}_{\ b}(x)$ is a local Lorentz transformation at the point x. With a tetrad field defined over our entire spacetime, we can then discuss the dynamics of massive spin-$\frac{1}{2}$ particle states with reference to local inertial reference frames whose isometries are described by the Lorentz group. Applying our tetrad field to the Dirac matrices, their form in the local inertial frame is given by

$$\gamma^{\mu} = e^{\mu}_{\ a}\gamma^{a}. \tag{4.43}$$

Furthermore, as shown in [16] we define the covariant derivate on $\Psi_{p,\alpha}(x)$ as

$$\mathcal{D}_\mu \Psi = (\partial_\mu - \Gamma_\mu)\Psi_{p,\alpha}(x) \tag{4.44}$$

where Γ_μ is called the *spinorial affine connection* and the covariant derivative on a Dirac matrix is

$$\mathcal{D}_\mu \gamma_\nu(x) = 0 \tag{4.45}$$

[13,16]. We will not derive the spinorial affine connection Γ_μ here for the sake of brevity, but it is shown by Lanzagorta in [16] to have the following form:

$$\Gamma_\mu(x) = -i\Sigma_{ab}\omega_\mu^{\ ab}, \tag{4.46}$$

where

$$\omega_\mu^{\ ab} = e^{a\alpha}(x)\nabla_\mu e^b_{\ \alpha}(x) \quad \text{and} \quad \Sigma_{ab} = -\frac{1}{8}[\gamma_a, \gamma_b]. \tag{4.47}$$

Lastly, the coordinate covariant derivative of the tetrad fields is

$$\nabla_\mu e^b_{\ \alpha}(x) = \partial_\mu e^b_{\ \alpha}(x) - \Gamma^\nu_{\ \mu\alpha} e^b_{\ \nu}(x) \tag{4.48}$$

[13,16,33]. Therefore, the Dirac equation in curved spacetime has the following form:

$$(i\gamma^\mu(x)\mathcal{D}_\mu - m)\Psi_{p,\alpha} = 0 \tag{4.49}$$

$$\Rightarrow (i\gamma^a\partial_a - m)\Psi = i\gamma^a\Gamma_a\Psi_{p,\alpha}, \tag{4.50}$$

where the lefthand side of Eq.(4.50) describes the mechanics of a free particle and the righthand side corresponds to the contribution from gravity [16,33].

Since we are studying qubits via local inertial reference frames defined at each point of our space-time, the evolution of the state of our system will result as a cumulative effect due to the transformations from each local coordinate system along the particle's trajectory. With this in mind, we wish to now discuss the effects of each local Lorentz transformation on the state of a massive spin-$\frac{1}{2}$ particle.

4.4 The Wigner Rotation

As shown in [16], the Poincare group has a unitary representation such that for an arbitrary Poincare transformation on the state of a spin-$\frac{1}{2}$ particle

$$\tilde{\Psi}_{p,\alpha} = \hat{U}(\Lambda, a)\Psi_{p,\alpha} \tag{4.51}$$

where $\hat{U}(\Lambda, a)$ is the unitary representation of our Poincare transformation with its translation described by a [16,27]. We denote the Lorentz transformation $\Lambda^{\mu}_{\ \rho}$ with Λ when it is the argument of another function in order to reduce the number of indices. Again, we note that since these transformations are only defined on the local space, it is possible that a translation could result in an event outside of our neighborhood. Therefore, we will only consider pure Lorentz transformations where $a = 0$. That is, we consider the transformations $\hat{U}(\Lambda)$. It then follows that assuming the spin degrees of freedom form a complete basis for the state space, we have

$$\hat{U}(\Lambda)\psi_{p,\alpha} = \sum_{\beta} C_{\alpha\beta}(\Lambda, p)\Psi_{\Lambda p,\beta}, \tag{4.52}$$

where $C_{\alpha\beta}$ denotes the complex coefficients of the expanded form of the state [16].

Note, in Eq. (4.52) we see that our Lorentz transformations can change the momentum of our

qubit; specifically, from p to Λp. However, since the square of the four-momentum is invariant under Lorentz transformations, we can assign states to specific classes identified by a unique four-momentum k^μ [16]. In other words, we may express any four-momentum p^μ as

$$p^\mu = L^\mu_{\ \nu}(p)k^\nu, \tag{4.53}$$

where L is used to denote a Lorentz transformation in order to avoid confusion with the already used Λ. Now any quantum state with four-momentum p^μ and characteristic momentum k^μ, can be written as

$$\Psi_{p,\alpha} = N(p)\hat{U}(L(p))\psi_{k,\alpha} \tag{4.54}$$

where $N(p)$ is a normalization constant. Then as shown by Lanzagorta in [16], applying our Poincare transformation to our state

$$\hat{U}(\Lambda)\Psi_{p,\alpha} = N(p)\hat{U}(L(\Lambda p))\hat{U}(W)\Psi_{k,\alpha} \tag{4.55}$$

where

$$W = W(\Lambda, p) = L^{-1}(\Lambda p)\Lambda L(p) \tag{4.56}$$

transforms the characteristic momentum as follows:

$$k^\mu \xrightarrow{L} p^\mu \xrightarrow{\Lambda} \Lambda^\mu_{\ \nu}p^\nu \xrightarrow{L^{-1}} k^\mu \ \text{ such that } \ W^\mu_{\ \nu}k^\nu = k^\mu, \tag{4.57}$$

leaving the four-momentum invariant. This transformation is called the *Wigner rotation* [16,33]. We note that the set of Wigner rotations form a subgroup of the Poincare group called the *little group*. This overall effect on a quantum state is characterized by a rotation in Hilbert space [16].

Therefore, when acting on a spin-$\frac{1}{2}$ particle, the resulting state is

$$\hat{U}(W)\psi_{k,\alpha} = \sum_{\beta} D_{\alpha\beta}\Psi_{k,\beta}, \tag{4.58}$$

where the $D_{\alpha\beta}$ coefficients represent the little group [16,33]. It is important to note that what we have written is the Wigner rotation experienced in a single local inertial reference frame. That is, we have established the effects on our state when it is confined to one local inertial reference frame. This is significant since in general an orbiting mass may take a path that traverses spacetime on a global scale, thus necessitating more than one local frame. In the following subsection we consider such a case and establish the overall evolution of a state due to the contribution of the Wigner rotations experienced in each local frame defined along our path.

4.4.1 The Cumulative Wigner Rotation

The momenta of our particles can be described in the local inertial frame via the tetrad fields,

$$p^a(x) = p^\mu(x)e^a{}_\mu(x). \tag{4.59}$$

Then as shown in [16], the infinitesimal change in our momentum is given by

$$\delta p^a(x) = \delta p^\mu(x)e^a{}_\mu(x) + p^\mu(x)\delta e^a{}_\mu(x) \tag{4.60}$$

where $\delta p^\mu(x)$ refers to external forces other than gravity and

$$\begin{aligned} \delta e^a{}_\mu(x) &= (\nabla_\nu e^a{}_\mu(x))u^\nu d\tau \\ &= -u^\nu(x)\omega_\nu{}^a{}_b(x)e^b{}_\mu(x)d\tau. \end{aligned} \tag{4.61}$$

It then follows that as a particle moves through a non-trivial spacetime, the momentum will experience an infinitesimal local Lorentz transformation

$$p^a(x) = \Lambda^a_{\ b}(x)p^b(x), \tag{4.62}$$

where $\Lambda^a_{\ b} = \delta^a_{\ b} + \lambda^a_{\ b}(x)d\tau$ and

$$\lambda^a_{\ b}(x) = -\frac{1}{m}(a^a(x)p_b(x) - p^a(x)a_b(x)) - u^\nu(x)\omega_\nu^{\ a}{}_b(x). \tag{4.63}$$

Furthermore, when we consider a particle free of external forces $a^\mu = 0$, our transformation reduces to

$$\lambda^a_b(x) = -u^\nu(x)\omega_\nu^{\ a}{}_b(x) \tag{4.64}$$

[16]. It is then shown by Lanzagorta that the associated infinitesimal Wigner rotation is

$$W^a_{\ b}(x) = \delta^a_{\ b} + \vartheta^a_{\ b}(x)d\tau, \tag{4.65}$$

where $\vartheta^a_{\ b}(x)$ is

$$\vartheta^a_{\ b}(x) = \left[\lambda^a_{\ b}(x) + \frac{\lambda^a_{\ 0}(x)u_b(x) - \lambda_{b0}(x)u^a(x)}{u^0(x) + 1}\right]d\tau. \tag{4.66}$$

It then follows that when we consider each local Wigner rotation as the particle moves across our spacetime, the cumulative effect over a finite proper time interval is

$$W^a_{\ b}(x_f; x_i) = \mathcal{T}e^{\int_{\tau_i}^{\tau_f} \vartheta^a{}_b(x(\tau))d\tau} \tag{4.67}$$

where $\mathcal{T}$ is a time-ordered product [16]. Finally, the unitary operator representation of the overall Wigner rotation is given by

$$\hat{D} = \mathcal{T} \exp\left(\frac{i}{2}\int_{\tau_i}^{\tau_f} \sigma \cdot (\vartheta^2{}_3, \vartheta^3{}_1 \vartheta^1{}_2)\right) \tag{4.68}$$

[16,33]. As shown by Bonior in [4], for the case of the the Schwarzchild metric, the gravitational effects on a massive spin-$\frac{1}{2}$ particle orbiting a large mass are fully described by the angle of rotation about the e_2 axis; that is, $\vartheta^3{}_1$. Therefore, our operator reduces to

$$\begin{aligned} \hat{D} =& e^{\frac{i}{2}\sigma_y \int_{\tau_i}^{\tau f} \vartheta_{31} d\tau} \\ =& e^{\frac{i}{2}\sigma_y \Omega}, \end{aligned} \tag{4.69}$$

where $\Omega = \int_{\tau_i}^{\tau_f} \vartheta^3{}_1 d\tau$. Thus far we have been working in the Hilbert representation of quantum mechanics. Now that we have found the explicit form of gravitational noise Eq. (4.69), we will return to the Bloch characterization for the remainder of this chapter. Again, we choose to do so as it greatly simplifies calculations in information theory.

4.5 Adaptability of Gravitational Noise

Upon inspection of Eq.(4.69), the unitary representation of the gravitational noise for an orbiting spin-$\frac{1}{2}$ particle in the Schwarzchild metric is

$$\hat{\mathcal{D}} = \begin{pmatrix} \cos\frac{\Omega}{2} & \sin\frac{\Omega}{2} \\ -\sin\frac{\Omega}{2} & \cos\frac{\Omega}{2} \end{pmatrix} \tag{4.70}$$

where Ω is the Wigner rotation. Therefore, its Bloch representation has the following form:

$$f = \begin{pmatrix} \cos\Omega & 0 & -\sin\Omega \\ 0 & 1 & 0 \\ \sin\Omega & 0 & \cos\Omega \end{pmatrix} \tag{4.71}$$

[16]. The matrix in Eq.(4.71) is a principal rotation about the e_2-axis in $\mathbb{R}^3$. Therefore, it follows from Theorem 3.3.3 that for a non-trivial Wigner rotation Ω, the error rate can be cut in half for Adaptive QKD between Earth and an orbiting satellite. Explicitly,

$$\mathrm{Im}\big(E_{q,f}\big) = \left[\tfrac{1}{2}\sin^2\tfrac{\Omega}{2}\ ,\ \sin^2\tfrac{\Omega}{2}\right], \tag{4.72}$$

and therefore, assuming the conventional communication bases for QKD (e_1 & e_3) are the initial choice of bases, we have that

$$E_{q,f}(e_1, e_3) = \frac{1}{2}\left[\frac{1-\cos\Omega}{2} + \frac{1-\cos\Omega}{2}\right] = \sin^2\tfrac{\Omega}{2} = \max\big(E_{q,f}\big). \tag{4.73}$$

Consequently, changing to the pair of bases $\{e_1, e_2\}$ or $\{e_2, e_3\}$, we obtain a reduction of the error rate by half its original value;

$$E_{q,f}(e_1, e_2) = E_{q,f}(e_2, e_3) = \frac{1}{2}\sin^2\frac{\Omega}{2}. \tag{4.74}$$

Let us further exemplify gravitational noise by explicitly calculating the Wigner rotation of a circular orbit in the Schwarzchild metric.

4.6 An Example: The Schwarzschild Metric

This section is a summary of the work previously done by Bonior in [4]. We will be modeling gravity due to Earth with the Schwarzschild metric that is defined by the following spacetime interval:

$$ds^2 = -\left(1 - \frac{2\mu}{r}\right)dt^2 + \left(1 - \frac{2\mu}{r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2, \tag{4.75}$$

where $\mu = GM$, G is the universal gravitational constant, and M is the mass of the Earth. Restricting our calculations to the case of a circular orbit, the first step is to solve the equations of motion dictated by our metric in Eq.(4.75). For convenience we will assume unity mass of our spin-$\frac{1}{2}$ particle, equatorial orbits ($\theta = \frac{\pi}{2}$), and introduce the following variables:

$$r_s = 2\mu \quad \text{and} \quad f = 1 - \frac{r_s}{r}. \tag{4.76}$$

The spacetime interval then reduces to

$$ds^2 = -f dt^2 + f^{-1} dr^2 + r^2\big(d\theta^2 + d\phi^2\big). \tag{4.77}$$

Therefore, considering the alternative form of the affine connections

$$\Gamma^{\sigma}_{\mu\nu} = \frac{1}{2} g^{\sigma\rho}\big(\partial_\nu g_{\rho\mu} + \partial_\mu g_{\rho\nu} - \partial_\rho g_{\mu\nu}\big), \tag{4.78}$$

it then follows that the only non-zero entries are given by:

$$\Gamma^{t}_{tr} = \frac{r_s}{2fr^2}, \quad \Gamma^{r}_{tt} = \frac{f r_s}{2r^2}, \quad \Gamma^{r}_{rr} = \frac{-r_s}{2fr^2}, \quad \Gamma^{r}_{\theta\theta} = -rf, \tag{4.79}$$

$$\Gamma^{r}_{\phi\phi} = -rf, \quad \Gamma^{\theta}_{r\theta} = \frac{1}{r}, \quad \text{and} \quad \Gamma^{\phi}_{r\phi} = \frac{1}{r} \tag{4.80}$$

[4,13,16]. Recalling that we are considering the motion of a particle in free fall, we have that our equations of motion are given by

$$\frac{\mathcal{D}u^\mu}{\mathcal{D}\tau} = \frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\tau}\frac{dx^\beta}{d\tau} = 0 \tag{4.81}$$

where u^μ is the four-velocity of the spin-$\frac{1}{2}$ particle. Consequently, we have the following three equations of motion:

$$\frac{du^t}{d\tau} + \frac{r_s}{fr^2}u^t u^r = 0, \tag{4.82}$$

$$\frac{du^r}{d\tau} + \frac{fr_s}{2r^2}u^t u^t - rfu^\phi u^\phi - \frac{r_s}{2fr^2}u^r u^r = 0, \tag{4.83}$$

$$\frac{du^\phi}{d\tau} + \frac{2}{r}u^\phi u^r = 0; \tag{4.84}$$

the solutions of which result in the geodesic equations:

$$u^t = \frac{C}{f}, \tag{4.85}$$

$$u^r = \sqrt{C^2 - \frac{J^2 f}{r^2} - kf}, \tag{4.86}$$

$$u^\phi = \frac{J}{r^2}, \tag{4.87}$$

where J, k, and C are integration constants. Since we are assuming unity mass, our four-momentum is $p^\mu = u^\mu$. It then follows from Eqs. (4.85) and (4.87) that J is associated with the angular momentum and C the total energy of the particle [4,13,16]. Furthermore, since we must have a time-like geodesic, we are forced to conclude that $-\frac{ds^2}{d\tau^2} = k = 1$. Noting that

$$\frac{dr}{d\tau} = \frac{dr}{d\phi}\frac{d\phi}{d\tau} = \frac{dr}{d\phi}\frac{J}{r^2}, \tag{4.88}$$

then defining the variable $u = r^{-1}$, by Eq. (4.86) we have that

$$\left(\frac{du}{d\phi}\right)^2 + u^2 f + \frac{f}{J^2} = \frac{C^2}{J^2}. \tag{4.89}$$

Taking the derivative of each side with respect to ϕ and simplifying our expression, we then have that

$$\frac{d^2u}{d\phi^2} + u - \tfrac{3}{2} r_s u^2 = \frac{r_s}{2J^2}. \tag{4.90}$$

Recalling that we are considering a circular orbit, it follows that $\frac{du}{d\phi} = 0$. Therefore, solving for J^2, we have that

$$J^2 = \frac{r_s r^2}{2r - 3r_s}. \tag{4.91}$$

Similarly, noting that r is constant for a circular orbit, and consequently $u^r = 0$, then Eq.(4.86) implies

$$C^2 = \left(\frac{J^2}{r^2} + 1\right) f. \tag{4.92}$$

Finally, plugging in our values for C and J, the equations of motion for a circular orbit are given by

$$u^t = \frac{1}{\sqrt{1 - \frac{3}{2}\frac{r_s}{r}}} \quad \& \quad u^\phi = \frac{1}{r}\sqrt{\frac{r_s}{2r - 3r_s}}. \tag{4.93}$$

In order to finally calculate the resulting Wigner rotation, we must first determine the local Lorentz transformations along our orbit. However, our local transformations are dependent on the spinorial connections and our tetrad field. With this in mind, we first note that

$$g_{\mu\nu} = \eta_{mn} e^m_{\ \mu} e^n_{\ \nu}. \tag{4.94}$$

It then follows that

$$e^0{}_t = \sqrt{f}, \qquad e^1{}_r = \frac{1}{\sqrt{f}}, \tag{4.95}$$

$$e^2{}_\theta = r, \qquad e^3{}_\phi = r. \tag{4.96}$$

and by Eq. (4.47)

$$\omega^0{}_t = \omega^1{}_t = \frac{r_s}{2r^2}, \tag{4.97}$$

$$\omega^1{}_\theta = -\omega^2{}_\theta = -\sqrt{f}, \tag{4.98}$$

$$\omega^1{}_\phi = -\omega^3{}_\phi = -\sqrt{f}. \tag{4.99}$$

Finally, plugging our values into Eq. (4.63), we find that the non-zero local Lorentz transformations are given by

$$\lambda^1{}_0 = -\frac{r_s}{2r^2\sqrt{1-\frac{3r_s}{2r}}} \quad \& \quad \lambda^3{}_1 = \frac{\sqrt{f}}{r}\sqrt{\frac{r_s}{2r-3r_s}} \tag{4.100}$$

where $\lambda^1{}_0$ is a boost over the x^1-axis and $\lambda^3{}_1$ is a rotation about the x^2-axis. Substituting these transformations into Eq. (4.66), we then obtain the non-zero infinitesimal Wigner rotation

$$\vartheta^3{}_1 = \frac{J\sqrt{f}}{r^2}\left(1 - \frac{Cr_s}{2rf}\frac{1}{C+\sqrt{f}}\right). \tag{4.101}$$

It then follows that upon integration over the proper time interval of the orbit, the total angle of rotation is

$$\Omega = \int_{\tau_i}^{\tau_f} \vartheta^3{}_1 d\tau = \vartheta^3{}_1 \frac{r^2}{J}\int_0^{2\pi} d\phi = 2\pi\vartheta^3{}_1\frac{r^2}{J} \tag{4.102}$$

However, after the particle has transversed an entire revolution around the Earth, it will have rotated by an angle of 2π independent from the effects of gravity. Therefore, we must subtract this angle

to obtain the Wigner rotation produced exclusively by gravitational effects. That is, the Bloch representation of our gravitational noise is

$$f = \begin{pmatrix} \cos\Omega & 0 & -\sin\Omega \\ 0 & 1 & 0 \\ \sin\Omega & 0 & \cos\Omega \end{pmatrix} \tag{4.103}$$

where the angle of the Wigner rotation can be rewritten in terms of the physical quantities r, G, and M as follows:

$$\Omega = 2\pi\left(\sqrt{1-\frac{2GM}{r}} - \frac{GM}{\sqrt{r}\left(\sqrt{r-2GM}+\sqrt{r-3GM}\right)} - 1\right). \tag{4.104}$$

Remark: Though we only consider massive particles for the sake of simplifying the mathematics, similar results hold for photons, i.e. a photon would also experience a Wigner rotation.

In conclusion, this means that we can cut the rate of error due to gravitational noise in half when performing Adaptive QKD with an Earth orbiting satellite. Explicitly, the minimum rate of error due to gravitational noise is

$$\tfrac{1}{2}\sin^2\left[\pi\left(\sqrt{1-\frac{2GM}{r}} - \frac{GM}{\sqrt{r}\left(\sqrt{r-2GM}+\sqrt{r-3GM}\right)} - 1\right)\right]. \tag{4.105}$$

4.7 Summary

This chapter was a brief introduction to gravitational effects on orbiting qubits. We laid out the foundations for general relativity and utilized previous results in order to obtain the explicit form of gravitational noise on a spin-$\frac{1}{2}$ particle subject to the Schwarzschild metric in a circular orbit around the Earth. The importance of this specific example is two-fold. Firstly, there is a practical

motivation for studying this set-up as it has been both theoretically and experimentally (for photons) investigated in previous research [34]; and secondly, gravitational noise is an example that offers the largest possible reduction for the error rate in Adaptive QKD. Simply put, gravitational noise is relevant to current communication technologies and there exists an algorithm that will cut its rate of error in half when performing QKD.

CHAPTER 5: DOMAIN THEORY AND SYMMETRIC UNITAL CHANNELS

While this chapter will stand apart from the rest of this work in terms of narrative, it aims to further our fundamental understanding of the underlying structure of the set of symmetric unital channels. The motivation for applying domain theory to $\mathcal{U}_s$, is founded on its success in many other areas of physics. For example, in [18] Martin shows that domain theory is rich enough in structure to give rise to globally hyperbolic spacetimes, and therefore, provides a means of discussing general relativity! As practical applications have already been verified, our goal is to now uncover a deeper understanding of the set of unital channels, and by extension provide, or create, a foundation for further applications.

Following the same format of our previous chapters, we will provide a brief introduction to the definitions necessary for our discussion. However, it is not uncommon for those new to the subject to lose grasp of the purpose of domain theory when first presented with the slew of axioms that define it. For this reason, we begin by presenting a conceptual rendition of a domain and its characteristic properties.

5.1 Preliminaries: Domain Theory

5.1.1 The Conceptual View

Domain theory was first introduced as a mathematical model of computation. The motivation for developing such, lies in the desire to discuss computations on different representations of information in a more fundamental way than that obtained by operational means. This concept is best

described by Dana Scott, "The point is that, mathematically speaking, functions are independent of their means of computation and hence are "simpler" than explicitly generated, step-by-step evolved sequences of operations on representations.", [31]. Dana Scott developed domain theory in 1970 in order to formalize these ideas, and since then this theory has been further developed at the hands of both himself and others; in the context of this dissertation, most notably Keye Martin. With this in mind, we begin by highlighting the core attributes of a *domain* in a constructive manner.

In computer science, before constructing an algorithm, let alone executing it, one must first establish a type for the data they wish to consider. That is, we must first provide an interpretation for the values and variables of the data we plan to manipulate. As such, we begin our discussion by considering the set D consisting of all objects of an arbitrary *data type*. However, the structure of a set is too simple to fully capture the properties of the elements of D; specifically it does not contain the notion of approximation.

In order to establish some notion of approximation, we begin by placing a partial order $\sqsubseteq$ on our set D that provides an informatic comparison between certain elements. That is, the statement $x \sqsubseteq y$ is taken to mean that y contains at least as much information as x. Furthermore, in the context of a computation, it is possible that y is a more accurate approximation. Of course, the question at hand then becomes "What does it mean for an element of D to approximate another?".

The notion of *approximation* is formalized as a special case of the informatic order $\sqsubseteq$. Specifically, x approximates y, denoted by $x \ll y$, if all informatic paths to or beyond y must pass through x. Intuitively, this means that x carries essential information about y. In all cases of practical intent, if a process generates a sequence of values (y_n) with supremum y, then $x \ll y$ if $x \sqsubseteq y_n$ for all but a finite number of y_n [23].

Lastly, a domain has a notion of *completeness*. This is equivalent to certain processes having "limits", or outcomes that they appear to be "going towards". For example, considering a process

that generates a sequence (y_n) that is increasing in the informatic order, then the completeness of a domain refers to the fact that the supremum $\bigsqcup y_n$ exists. Intuitively, $\bigsqcup y_n$ is the final result of the process that generates the sequence (y_n) [23].

Succinctly put, a domain is a partially ordered set of objects of a particular type that has the intrinsic notions of approximation and completeness. With some of the basic concepts of a domain established, we now part from our conceptual discussion and begin the preliminary introduction of domain theory's mathematical foundations.

5.1.2 Mathematical Foundations

Definition 5.1.1: Let X be an arbitrary set. A *binary relation* $\sqsubseteq$ is a subset of $X \times X$. We use the notation $x \sqsubseteq y$ to mean $(x, y) \in \sqsubseteq$ and call $(X, \sqsubseteq)$ a *partially ordered set*, or *poset*, if the following three properties hold for all $x, y, z \in X$:

(i) $x \sqsubseteq x$,

(ii) If $x \sqsubseteq y$ and $y \sqsubseteq z$ then $x \sqsubseteq z$,

(iii) If $x \sqsubseteq y$ and $y \sqsubseteq x$ then $x = y$.

A basic example of a poset is some collection of sets with the partial order $\sqsubseteq$ taken to be the set inclusion relation. Note, this example illustrates that not all elements of a poset necessarily compare, as it is possible to construct two sets in which neither one is contained in the other. Alternatively, we could also consider the poset $(\mathbb{R}, \leq)$, where $\leq$ is the standard order on the real numbers. This is an example of a poset in which every element compares. Objects such as these are referred to as *totally ordered sets*.

Definition 5.1.2: Let $(X, \sqsubseteq)$ be a poset. For any subset $A \subseteq X$, if $y \sqsubseteq x$ for all $y \in A$, then x is called an *upper bound* of the set A. Furthermore, if $x \sqsubseteq z$ for any other upper bound z of A, then

x is referred to as the *supremum* of A and is denoted by $\sup(A)$, or $\bigsqcup A$.

Definition 5.1.3: Let X be a poset. A nonempty subset $A \subseteq X$ is *directed* if for every $x, y \in A$ there exists $z \in X$ such that $x, y \sqsubseteq z$. If every directed subset of X has a supremum, then X is called a *directed complete poset*, or *dcpo*.

One of the simplest examples of a dcpo is the closed interval $[0, 1]$ under the usual order $\leq$. In this case, 1 is an upper bound of any nonempty subset $A \subseteq [0, 1]$. It then follows by Dedekind completeness that $\bigsqcup A$ exists.

Intuitively, if we were to perform a computation on the space X, a directed set can be seen as the set of partial answers obtained as our algorithm converges to a final outcome. The final outcome then being the supremum of our directed set. This provides a means of assigning intrinsic information between elements. We can think of x being intrinsic to y, if during the course of any computation that has the answer y, we must first compute x along the way.

Definition 5.1.4: Let $(X, \sqsubseteq)$ be a *dcpo* with $x, y \in X$. If for every directed subset A where $y \sqsubseteq \bigsqcup A$, we have $x \sqsubseteq z$ for some $z \in A$, then we say that x *approximates* y, or $x \ll y$.

We note that if $x \ll y$, then when computing y, it is essential to first compute x. Therefore we would expect $x \ll y$ to imply $x \sqsubseteq y$. This is easily verified as the collection of all elements below y forms a directed set; and by definition $x \ll y$ implies that there exists an element below y that is also above x. Therefore, by the transitive property, $x \sqsubseteq y$.

Definition 5.1.5: Let $(X, \sqsubseteq)$ be a dcpo.

$\uparrow x \equiv \{y \in X \mid x \sqsubseteq y\}$ (*principal upper set* of x)

$\downarrow x \equiv \{y \in X \mid y \sqsubseteq x\}$ (*principal lower set* of x)

$\Uparrow x \equiv \{y \in X \mid x \ll y\}$ (*way above set* of x)

$\Downarrow x \equiv \{y \in X \mid y \ll x\}$ (*way below set* of x)

Definition 5.1.6: Let $(X, \sqsubseteq)$ be a poset. Then a subset $U \subseteq X$ is *Scott open* if

(i) U is an upper set: $x \in U \ \ \& \ \ x \sqsubseteq y \ \Rightarrow \ y \in U$,

(ii) U is inaccessible by directed suprema: For every directed subset $S \subseteq X$ which has a supremum,

$$\bigsqcup S \in U \ \ \Rightarrow \ \ S \cap U \neq \emptyset. \tag{5.1}$$

The collection of all Scott open sets forms the *Scott topology* which denote by τ_S. Furthermore, it is shown in [38] that the collection $\{\uparrow x \mid x \in X\}$ is a basis for τ_S on a *continuous* poset X.

Definition 5.1.7: Let $(X, \sqsubseteq)$ be a poset. We say that X is *continuous* at $x \in X$ if $\downarrow x$ is directed with $\bigsqcup \downarrow x = x$. If X is a dcpo that is continuous at all $x \in X$, then we call $(X, \sqsubseteq)$ a *domain*.

We conclude this section with three final definitions of particular mappings between domains.

Definition 5.1.8: Let $\phi : X \to Y$ be a function between posets. Then ϕ is *monotone* if for all $x, y \in X$, $x \sqsubseteq y$ implies $\phi(x) \sqsubseteq \phi(y)$. Furthermore, we say a monotone function ϕ is *strictly monotone* if $\phi(x) = \phi(y)$ and $x \sqsubseteq y$ implies $x = y$. We say ϕ preserves suprema if for every directed subset $A \subseteq X$ such that $\bigsqcup A$ exists, then $\bigsqcup \phi(A)$ exists and $\bigsqcup \phi(A) = \phi(\bigsqcup A)$. A function is *Scott continuous* if it is monotone and preserves suprema.

Definition 5.1.9: Let X, Y be sets, each with its own binary relation. We then call an invertible function $\phi : X \to Y$ an order isomorphism if both it and its inverse are monotone.

Order isomorphisms are, the "order version" of homeomorphisms in topology; or a diffeomorphism in differential geometry. Simply put, each order isomorphism represents an equivalence between order structures.

Definition 5.1.10: Let $\mu : X \to Y$ be a Scott continuous function between domains. Then the

ϵ-approximations of x are given by $\mu_\epsilon(x) = \{y \sqsubseteq x \mid \epsilon \ll \mu y\}$ where $\epsilon \in Y$. If $y \ll x$ and we can find some ϵ such that $z \in \mu_\epsilon(x)$ implies $y \ll z$, then we say μ *measures the content of* x. If μ measures the content of ker$\mu = \{x \in X \mid \mu x \in \max Y\}$, then μ is a *measurement*.

With some of the basic foundations of domain theory established, we now aim to define an informatic partial order on the set of simple functions $\mathbb{E}_s$.

5.2 The Informatic Order of Error Functions

Let E_f and E_g be two arbitrary simple functions. Then

$$E_f(x) \sqsubseteq E_g(x) \quad \iff \quad E_f(x) \leq E_g(x) \tag{5.2}$$

for all $x \in S^2$, where $\leq$ is the standard order on the real numbers. This order is of interest as it provides a means of comparing the values of two simple functions that are generated by different channels. That is, we are comparing simple functions based on the values of their images, which are in turn determined by the channels that generate them. Wishing to extend a similar order to the set of symmetric unital channels themselves, we say that for any $f, g \in \mathcal{U}_s$

$$f \sqsubseteq g \quad \iff \quad \langle x, \big[f - g\big]x\rangle \geq 0, \tag{5.3}$$

for all $x \in S^2$. While it appears that we have defined two different ordered sets in Eqs. (5.2) and (5.3), we have in fact described the same object via two different representations.

Theorem 5.2.1: $(\mathbb{E}_s, \sqsubseteq)$ is order isomorphic to $(\mathcal{U}_s, \sqsubseteq)$.

Proof: From Theorem 2.4.6, we know that $G : \mathcal{U}_s \to \mathbb{E}_s$ is a continuous, bijection. Thus, if we can show that G and its inverse are monotone, the proof is complete. By definition, $E_f \sqsubseteq E_g$ implies

that

$$\frac{1}{2}\big[1-\langle x, fx\rangle\big] \leq \frac{1}{2}\big[1-\langle x, gx\rangle\big] \tag{5.4}$$

for all $x \in S^2$. It then follows by basic arithmetic that E_f is below E_g if and only if

$$\langle x, \big[f-g\big]x\rangle \geq 0. \tag{5.5}$$

Therefore, we are left to conclude that $E_f \sqsubseteq E_g$ if and only if $f \sqsubseteq g$, and the proof is complete.

□

Theorem 5.2.2: The binary relation $\sqsubseteq$ is a partial order on the set of symmetric unital channels.

Proof: For every symmetric unital channel f

$$\langle x, \big[f-f\big]x\rangle = 0 \geq 0, \tag{5.6}$$

for all $x \in S^2$. Therefore, $\sqsubseteq$ is reflexive. Furthermore, if $f \sqsubseteq g$ and $g \sqsubseteq h$, then for all $x \in S^2$

$$\langle x, \big[f-h\big]x\rangle = \langle x, \big[f-g\big]x\rangle + \langle x, \big[g-h\big]x\rangle \geq 0. \tag{5.7}$$

Consequently, $\sqsubseteq$ is also transitive. Finally, we show that our order is antisymmetric by considering the case in which $f \sqsubseteq g$ and $g \sqsubseteq f$. Under this assumption,

$$\langle x, \big[f-g\big]x\rangle \leq 0 \leq \langle x, \big[f-g\big]x\rangle, \tag{5.8}$$

which is true if and only if

$$\langle x, \big[f-g\big]x\rangle = 0. \tag{5.9}$$

Then, by Eq. (5.9), and the fact that f and g are each symmetric, we arrive at the conclusion

that $[f - g]$ is both symmetric and skew-symmetric. This directly implies that $[f - g] = 0$; or equivalently $f = g$. And with this last step we have shown that the order relation defined in Eq. (5.3) forms a partially ordered set on the collection of symmetric unital channels.

□

In summary, Theorems 5.2.1 and 5.2.2 let us study the order relation defined on the set of simple functions in Eq. (5.2) via the partially ordered set $(\mathcal{U}_s, \sqsubseteq)$. Therefore, for the remainder of this chapter, when we write $\sqsubseteq$, we mean the partial order on $\mathcal{U}_s$ defined in Eq. (5.3). With this in mind, the objective of this chapter is to further characterize the structure of $(\mathcal{U}_s, \sqsubseteq)$. We begin our investigation by first establishing some significant properties of the symmetric unital channels themselves. However, before shifting our focus we would like to make one more observation about our order that will prove to simplify many arguments that will be made in the future.

Theorem 5.2.3: In the poset $(\mathcal{U}_s, \sqsubseteq)$, for any symmetric unital channels f and g, $f \sqsubseteq g$ if and only if $rfr^t \sqsubseteq rgr^t$ for all $r \in SO(3)$.

Proof: By definition of our order relation, $f \sqsubseteq g$ if and only if

$$\langle x, [f - g]x \rangle \geq 0 \tag{5.10}$$

for all $x \in S^2$. It then follows that since every $r \in SO(3)$ maps S^2 bijectively onto itself, f is below g if and only if

$$\langle r^t x, [f - g]r^t x \rangle \geq 0 \tag{5.11}$$

for all $x \in S^2$ and $r \in SO(3)$. Therefore, by the definition of an adjoint operator, and since $r^\dagger = r^t$, we have that $f \sqsubseteq g$ if and only if

$$\langle x, r[f - g]r^t x \rangle \geq 0; \tag{5.12}$$

or equivalently, $rfr^t \sqsubseteq rgr^t$.

□

Remark: As This result will be utilized fairly regularly, we will often refer to Theorem 5.2.3 with the statement "conjugation by a rotation is an order isomorphism".

5.2.1 Properties of Symmetric Unital Channels

Theorem 5.2.4: The identity matrix I is the least element in $(\mathcal{U}_s, \sqsubseteq)$.

Proof: If f is a symmetric unital channel, then the magnitude of each of its eigenvalues $\lambda_3 \leq \lambda_2 \leq \lambda_1$ is less than or equal to 1. It then follows that for all unit vectors x

$$-1 \leq \lambda_3 \leq \langle x, fx \rangle \leq \lambda_1 \leq 1. \tag{5.13}$$

Furthermore, since $\langle x, Ix \rangle = 1$ for all $x \in S^2$, we have that

$$\langle x, \big[I - f\big]x \rangle = \langle x, Ix \rangle - \langle x, fx \rangle = 1 - \langle x, fx \rangle \geq 0. \tag{5.14}$$

□

Theorem 5.2.5: For every symmetric unital channel $f \neq I$, there exists a unique $p \in [0, 1)$ and a unique $m \in \mathcal{U}_s$ with $\text{tr}(m) = -1$ such that

$$f = pI + (1 - p)m. \tag{5.15}$$

Proof: We begin by noting that for the case of the identity f can still be written is the same form

as Eq. (5.15). However, the choice of m does not matter and p is trivially 1. Continuing on, if we assume $f \neq I$, then from Proposition 1.5.2 we know that there exists some $r \in SO(3)$ and $q \in \Delta^4$ such that

$$f = \sum_{i=0}^{3} q_i r^t s_i r, \tag{5.16}$$

where the s_i's are the Bloch representations of the identity and spin channels given in Eq. (1.25). Therefore, we may write

$$f = q_0 I + (1 - q_0) \sum_{i=1}^{3} \frac{q_i}{1 - q_0} r^t s_i r. \tag{5.17}$$

Checking that the following coefficients sum to 1

$$\sum_{i=1}^{3} \frac{q_i}{1 - q_0} = 1, \tag{5.18}$$

we then have that the summation on the righthand side of Eq. (5.17) is a convex sum. That is, our summation forms a symmetric unital channel with

$$\text{tr}\left(\sum_{i=1}^{3} \frac{q_i}{1 - q_0} r^t s_i r \right) = -\sum_{i=1}^{3} \frac{q_i}{1 - q_0} = -1. \tag{5.19}$$

Therefore, letting $p = q_0$ and $m = \sum_{i=1}^{3} \frac{q_i}{1-q_0} r^t s_i r$, each symmetric unital channel f may be written in the following form:

$$f = pI + (1 - p)m, \tag{5.20}$$

where $p \in [0, 1)$ and $m \in \mathcal{U}_s$ with $\text{tr}(m) = -1$. To show the uniqueness of m and p, we assume there exists distinct pairs (m_1, p_1) and (m_2, p_2) that satisfy Eq. (5.20). That is,

$$p_1 I + (1 - p_1)m_1 = f = p_2 I + (1 - p_2)m_2. \tag{5.21}$$

Under this assumption, we then have that the following statements are equivalent:

$$\begin{aligned} \text{tr}\big(p_1 I + (1-p_1)m_1\big) &= \text{tr}\big(p_2 I + (1-p_2)m_2\big) \\ p_1 \,\text{tr}(I) + (1-p_1)\,\text{tr}(m_1) &= p_2 \,\text{tr}(I) + (1-p_2)\,\text{tr}(m_2) \\ 3p_1 - (1-p_1) &= 3p_2 - (1-p_2) \\ p_1 &= p_2. \end{aligned} \tag{5.22}$$

Therefore, p is unique. Furthermore, calling $p = p_1 = p_2$, we have from Eq. (5.21) that

$$(1-p)m_1 = (1-p)m_2. \tag{5.23}$$

Consequently, since $p_1 \in [0,1)$, we are left to conclude that $m_1 = m_2$, and therefore m is also unique.

□

We would like to take a moment to discuss the significance of Theorem 5.2.5. While it may not yet be clear, this result is an important tool; as we will see in Corollary 5.2.6. However, the most interesting results that follow from Theorem 5.2.5 have to do with the overall structure of $(\mathcal{U}_s, \sqsubseteq)$ itself. Let us now explain this in more detail.

We begin by noting that as a result of Theorem 5.2.5 there exists a canonical way to split the set of non-identity symmetric unital channels into equivalence classes. That is, for each trace -1 symmetric unital channel m, we form the equivalence class

$$[m] = \Big\{ f \in \mathcal{U}_s \;\Big|\; \exists p \in [0,1) \text{ where } f = pI + (1-p)m \Big\}. \tag{5.24}$$

There are several interesting implications that follow from the equivalence classes given in Eq.

(5.24). First off, each equivalence class $[m]$ is a half-open line segment with endpoints m and I, where the latter end point is not included. That is, letting M denote the set of trace -1 symmetric unital channels,

$$\mathcal{U}_s = \{I\} \cup \bigcup_{m \in M} [m]. \tag{5.25}$$

We can alternatively think of $\mathcal{U}_s$ as being the union of every closed line segment from the identity to some trace -1 symmetric unital channel. However, in this description, each line segment would have the identity as a common point. Furthermore, there is an even more significant characterization of $\mathcal{U}_s$.

We begin by defining the following sets for each $p \in [0, 1]$:

$$A_p = \Big\{ f \in \mathcal{U}_s \;\Big|\; \exists m \in M \text{ where } f = pI + (1-p)m \Big\}. \tag{5.26}$$

We note that each A_p forms an antichain, i.e. any two distinct elements in A_p do not compare. This is because the trace function is a strictly monotone map on $\mathcal{U}_s$, as shown in Lemma 5.2.8. Explicitly, for all $f, g \in A_p$ we have that $\text{tr}(f) = \text{tr}(g)$, and therefore $f \sqsubseteq g$ if and only if $f = g$. Furthermore, since all channels in $[m]$ compare, this means that each equivalence class $[m]$ has one and only one element in each antichain A_p. Therefore, the poset $(\mathcal{U}_s, \sqsubseteq)$ can be visualized as either the union of an infinite number of line segments, or the union of the antichains defined by Eq. (5.26).

Now returning to the properties of symmetric unital channels, the following corollary is a direct result of Theorem 5.2.5.

Corollary 5.2.6: For every symmetric unital channel f:

(i) $-1 \leq \text{tr}(f) \leq 3$.

(ii) There exists a $g \in \mathcal{U}_s$ where $f \sqsubseteq g$ and $\text{tr}(g) = -1$.

(iii) f has a non-zero fixed point if and only if f is the identity or there exists a rotation r in $SO(3)$ such that $f = pI + (1-p)r^t s_1 r$ for some unique $p \in [0, 1)$.

Proof:

(i) From Theorem 5.2.5 either $f = I$, in which case $\text{tr}(f) = 3$, or $f = pI + (1-p)m$ where m has trace -1 and $p \in [0, 1)$. In the latter case

$$\text{tr}(f) = p\,\text{tr}(I) + (1-p)\text{tr}(m) = 3p - (1-p) = 4p - 1. \tag{5.27}$$

Therefore, since $p \in [0, 1)$, we have $-1 \leq \text{tr}(f) \leq 3$ where the minimum and maximum values are obtained when $p = 0$ and $f = I$ respectively.

(ii) Again, writing $f = pI + (1-p)m$, if we let $g = m$, then

$$f - g = p(I - m) \geq 0, \tag{5.28}$$

where we used the fact that I is the least element (Theorem 5.2.4).

(iii) When $f = I$, our result is trivial as each of its eigenvalue would be 1. Assuming $f \neq I$, then because f is non-expansive we have that for all $x \in S^2$

$$\begin{aligned} \langle x, fx \rangle &= ||x||\,||fx|| \cos(\theta) \\ &= ||fx|| \cos(\theta) \leq ||x|| \cos(\theta) = \cos(\theta), \end{aligned} \tag{5.29}$$

where θ is the angle between the vectors x and fx. It then follows from Eq. (5.29) that there exists some $x \in S^2$ where $fx = x$ if and only if $\langle x, fx \rangle = 1$. In other words, f has a

non-zero fixed point if and only if there exists some $x \in S^2$,

$$\begin{aligned} \langle x, fx \rangle &= p\langle x, Ix \rangle + (1-p)\langle x, mx \rangle \\ &= p + (1-p)\langle x, mx \rangle = 1 \end{aligned} \tag{5.30}$$

Thus, noting that $p \in [0,)$, by basic arithmetic the following statements are equivalent:

$$\begin{aligned} p + (1-p)\,\langle x, mx \rangle &= 1 \\ (1-p)\,\langle x, mx \rangle &= 1-p \\ \langle x, mx \rangle &= 1 \end{aligned} \tag{5.31}$$

Therefore, since each unital channel is non-expansive it follows that if f has a non-zero fixed point, then $1 \in \sigma(m)$. Consequently, letting m have eigenvalues $\alpha_3 \leq \alpha_2 \leq \alpha_1$, we have that

$$\text{tr}(m) = 1 + \alpha_2 + \alpha_3 = -1, \tag{5.32}$$

or equivalently, $\alpha_2 + \alpha_3 = -2$. It then follows that since the magnitude of each eigenvalue is less than or equal to 1, we have that $\sigma(m) = \{1, -1, -1\}$. Therefore, there exists a rotation r such that $rmr^t = s_1$; or equivalently, $m = r^t s_1 r$.

□

Remark: In Corollary 5.2.6, result (i) can be extended to the set of unital channels $\mathcal{U}$ since $\text{tr}(\varphi(f)) = \text{tr}(f)$. Furthermore, in part (iii) s_1 can be replaced with any spin channel. This is because every spin channel has the same spectrum. In other words, for each $i \in \{1, 2, 3\}$, $\exists r_i \in SO(3)$ such that $m = r_i^t s_i r_i$.

The set of trace -1 symmetric unital channels has greater significance than simplifying the elements of $\mathcal{U}_s$ and providing an element in each principal upper set. However, in order to fully

appreciate the role of this set of channels we must first introduce a new definition and monotone function.

Definition 5.2.7: Let x be an element in the poset $(X, \sqsubseteq)$. We say x is a *maximal* element of X if for all $y \in X$, $x \sqsubseteq y$ implies $x = y$. We denote the set of maximal elements for a set X by $\max(X)$.

Lemma 5.2.8: The trace function $\text{tr} : \mathcal{U}_s \to \mathbb{R}$ is strictly monotone, where the standard order on $\mathbb{R}$ is reversed. That is, for all $a, b \in \mathbb{R}$, we say $a \sqsubseteq b$ if $b \leq a$.

Proof: If $f \sqsubseteq g$, then by definition

$$f - g \geq 0. \tag{5.33}$$

Furthermore, since $f - g$ is symmetric and $\langle x, [f-g]x\rangle \geq 0$, we know that the smallest eigenvalue of $f - g$ is non-negative. Therefore, we have that $\text{tr}(f - g) \geq 0$, as the trace of a matrix is simply the sum of its eigenvalues. With this in mind, it follows that

$$\text{tr}(f - g) = \text{tr}(f) - \text{tr}(g) \geq 0. \tag{5.34}$$

Consequently, $\text{tr}(f) \geq \text{tr}(g)$, and we have that $\text{tr}(f) \sqsubseteq \text{tr}(g)$. Additionally, when $\text{tr}(f) = \text{tr}(g)$ and $f \sqsubseteq g$, we then have that both $\text{tr}(f - g) = 0$ and $f - g \geq 0$. It then follows that not only do the eigenvalues of the matrix $[f - g]$ sum to 0, but they are each non-negative. Therefore, every eigenvalue must be 0. In other words, $[f - g]$ is the zero matrix, and we are forced to conclude that the trace function is strictly monotone.

□

Theorem 5.2.9: Any symmetric unital channel f is maximal in $(\mathcal{U}_s, \sqsubseteq)$ if and only if $\text{tr}(f) = -1$.

Proof: Let f be a maximal element in $(\mathcal{U}_s, \sqsubseteq)$. By Corollary 5.2.6, there exists some symmetric unital channel g such that $\text{tr}(g) = -1$ and $f \sqsubseteq g$. Then since f is maximal, it follows that $f = g$.

Conversely, let $\text{tr}(f) = -1$. In Lemma 5.2.8 we showed that the trace function is strictly monotone. Therefore, if there exists some element $g \in \mathcal{U}_s$ where $f \sqsubseteq g$, then $\text{tr}(g) \leq \text{tr}(f) = -1$. However, from Corollary 5.2.6, we know that $-1 \leq \text{tr}(g) \leq 3$, and consequently we conclude that $\text{tr}(g) = -1$. Finally, because the trace function is strictly monotone, it follows that that $f = g$ and the proof is complete.

□

Corollary 5.2.10: The poset $(\mathcal{U}_s, \sqsubset)$ does not have a greatest element $\top$.

Proof: We prove this result by contradiction. Let $\top$ be the greatest element in $(\mathcal{U}_s, \sqsubset)$. Therefore, $\top$ is above every symmetric unital channel, including every maximal element. That is, $s_1 \sqsubseteq \top$ and $s_2 \sqsubseteq \top$. However, this would imply that $\top = s_1 = s_2$, which we know is false by the definition of s_1 and s_2. Consequently, $(\mathcal{U}_s, \sqsubset)$ does not contain a greatest element.

□

Corollary 5.2.11: There does not exist a surjective, order reversing map A where $A : \mathcal{U}_s \to \mathcal{U}_s$.

Proof: Assuming A exists, it then follows that since $I \sqsubseteq f$ for all $f \in \mathcal{U}_s$, we would have that $A(f) \sqsubseteq A(I)$. That is, $A(I)$ would be the greatest element. However, we showed in Corollary 5.2.10 that $(\mathcal{U}_s, \sqsubseteq)$ does not have a greatest element, and we are therefore left to conclude that A does not exist.

□

Corollary 5.2.12: The antipodal map $a : S^2 \to S^2$ where $ax = -x$ is not a unital channel.

Proof: Let $A : \mathcal{U}_s \to \mathcal{U}_s$ such that $Af = a \circ f$. Therefore, if we assume that the antipodal map is a unital channel, then since the set of unital channels is closed under composition, we would have that $Af = -f \in \mathcal{U}_s$ for all $f \in \mathcal{U}_s$. It then follows that $f \sqsubseteq g$ would imply $Ag \sqsubseteq Af$, and every channel would be contained in the image of A. Consequently, A would be a surjective, order reversing map on $\mathcal{U}_s$. However, this contradicts Corollary 5.2.11, and we are left to conclude that $a \notin \mathcal{U}_s$.

□

With several properties for the elements of $\mathcal{U}_s$ established, we now turn our attention to the order theoretic structure of the poset $(\mathcal{U}_s, \sqsubseteq)$ for the remainder of this chapter.

5.2.2 The Directed-Complete Poset of Symmetric Unital Channels

Theorem 5.2.13: Let f be a symmetric unital channel. The principal lower and upper sets $\downarrow f$ and $\uparrow f$ are each closed in $(\mathcal{U}_s, \tau)$ where τ is the uniform metric topology.

Proof: Let the principal lower set $\downarrow f$ contain the sequence (y_n) such that $y_n \to y$ in τ. Then since $y_n \in \downarrow f$, it follows that $y_n - f \geq 0$ for all n. Therefore, $\lim(y_n - f) = y - f \geq 0$. That is, $y \in \downarrow f$. Similarly, if (z_n) is some sequence in $\uparrow f$ such that $z_n \to z$ in τ, then $\lim(z_n - f) = z - f \leq 0$ and $z \in \uparrow f$. Moreover, since $(\mathcal{U}_s, \tau)$ is a metric space, every limit point is the limit of a sequence. In other words, we have shown that both $\downarrow f$ and $\uparrow f$ contain all their limit points. Therefore, by Theorem 2.1.10 we are left to conclude that the principal upper and lower sets are closed in $(\mathcal{U}_s, \tau)$.

□

Theorem 5.2.14: The partially ordered set $(\mathcal{U}_s, \sqsubseteq)$ is directed-complete.

Proof: This proof directly follows from Theorem 5.2.13 and the results of "A new fixed point

theorem in Domain Theory" by Martin and Feng [25]. In this paper, it is shown in Theorem 3.2 that every poset with a compact, Hausdorff topology in which the principal upper and lowers sets are closed, is a directed-complete partially ordered set in which every filtered subset has an infimum. Therefore, since $\mathcal{U}_s$ is a closed subset of $\mathcal{U}$, we have that $(\mathcal{U}_s, \tau)$ is a compact metric space in which $\uparrow f$ and $\downarrow f$ are closed and the proof is complete. Note, we have used the fact that every metric space is Hausdorff.

□

In summary, we have shown that the set of symmetric unital channels with the order given by Eq. (5.3) has the following properties:

(i) I is the least element and there is no greatest element.

(ii) The set of maximal elements is given by $\max\big(\mathcal{U}_s\big) = \big\{f \in \mathcal{U}_s \mid \text{tr}(f) = -1\big\}$.

(iii) The antipodal map $a : S^2 \to S^2$ where $ax = -x$, is not a unital channel.

(iv) Every non-trivial symmetric unital channel f is characterized by a unique maximal element m and unique $p \in [0, 1)$ such that

$$f = pI + (1 - p)m. \tag{5.35}$$

(v) And lastly, $(\mathcal{U}_s, \sqsubseteq)$ is a directed-complete partially ordered set in which the principle upper and lower sets are closed with respect to the uniform metric topology.

Recall that a domain is characterized as having the intrinsic notions of completeness and approximation. Therefore, our next objective is to establish an approximation relation on the dcpo $(\mathcal{U}_s, \sqsubseteq)$, in order to verify whether or not it forms a domain.

5.3 The Approximation Relation

We begin our efforts to find a approximation relation by making several initial observations about the informatic order on the set of symmetric unital channels.

Theorem 5.3.1 (Martin & Panangaden): Every directed subset $A \subseteq \mathcal{U}_s$ contains a convergent cofinal net whose limit is $\bigsqcup A$ in the uniform metric topology τ.

Proof: We begin by noting that this result has been previously shown by Martin. That is, the proof provided here is simply a reiteration of Theorem 6.1 in [18]. In particular, this result is a specific case of the exact arguments made by Martin and Panangaden for which they considered a more generalize type of poset.

Let $A \subseteq \mathcal{U}_s$ be a directed subset. It then follows that the set $\{\uparrow f \mid f \in A\}$ has the finite intersection property. Furthermore, since $\mathcal{U}_s$ is compact in the uniform topology, the set of upper bounds of A, given by $\bigcap_{f \in A} \uparrow f$, is nonempty. We denote the upper bound of A by x.

Fixing any element $1 \in A$, then for all $f, g \in \uparrow 1 \bigcap A$ we have that $1 \sqsubseteq f, g$ and $f, g \in A$. Then since A is directed, there exists some h in A such that $f, g \sqsubseteq h$. Consequently, $h \in \uparrow f$. That is, there exists a channel h in $\uparrow 1 \bigcap A$ such that $f, g \sqsubseteq h$. Therefore, we have that $\uparrow 1 \bigcap A$ is also a directed subset. Furthermore, $\uparrow 1 \bigcap A$ has a supremum if and only if A does. This is due to the fact that the set of upper bounds of these two sets coincide. Therefore, we may instead consider the set $\uparrow 1 \bigcap A$, which we will now rename A. Note, we may also assume that A has a least element, namely 1.

Letting $F : A \to \mathcal{U}_s \ :: \ a \mapsto a$, we have F is a net in $(\mathcal{U}_s, \sqsubseteq)$. Since $\uparrow 1 \bigcap \downarrow x$ is τ-closed subset of $\mathcal{U}_s$, A is contained in a compact set. Consequently, F has a convergent subnet $G : I \to A$. Defining $T = G(I) \subseteq A$, we then have that T is directed, and by the definition of a subnet, cofinal in A. We

will now show that $\bigsqcup T = \lim(T)$. Note, we have placed no assumptions on the cofinal subnet G beyond that it be convergent. That is, all following results about G apply to any convergent cofinal subnet of F.

First we show $\lim(T)$ is an upper bound of T. If there exists a $t \in T$ such that $t \not\sqsubseteq \lim(T)$, then $\lim(T) \in \big(\mathcal{U}_s \backslash \uparrow t\big)$. Furthermore, since $\big(\mathcal{U}_s \backslash \uparrow t\big)$ is open in $\mathcal{U}_s$, there exists $\alpha \in I$ such that for all $\beta \in I$ we have that $\alpha \leq \beta$. This then implies that $g(\beta) \in \big(\mathcal{U}_s \backslash \uparrow t\big)$. Therefore, since I is directed and g is a subnet, if we let $u = g(\alpha)$ and $t = g(\gamma)$, then there exists some $\beta \in I$ such that $\alpha, \gamma \leq \beta$. It then follows that $g(\beta) \in \big(\mathcal{U}_s \backslash \uparrow t\big)$ and $t = g(\alpha) \sqsubseteq g(\beta)$, where the second inequality results from the fact that subnets are monotone by definition. This is a contradiction, and therefore implies that $t \sqsubseteq \lim(T)$ for all $t \in T$.

In order to show that $\lim(T) = \bigsqcup T$, we begin by letting u be an upper bound of T. Then if $\lim(T) \not\leq u$, it follows that $\lim(T) \in \big(\mathcal{U}_s \backslash \downarrow u\big)$. As this is an open subset, we then have that $T \bigcap (\mathcal{U}_s \backslash \downarrow u) \neq \emptyset$. This of course contradicts our assumption that u is an upper bound of T. Consequently, $\lim(T) = \bigsqcup T$.

Lastly, we show that $\bigsqcup A = \lim(T)$. Let $a \in A$, then since T is cofinal in A, there must exists some $t \in T$ such that $a \leq t$. That is, $a \leq t \leq \lim(T)$, and therefore $\lim(T)$ is an upper bound of A. Furthermore, since any upper bound of A is also one for T, it follows that each one must be above $\lim(T)$. In other words, $\lim(T) = \bigsqcup A$.

We have then shown that every directed subset of $A \subseteq \mathcal{U}_s$ contains the the image of the convergent cofinal net $G : I \to A$, such that $T = G(I) \subseteq A$ is a cofinal directed set with $\lim(T) = \bigsqcup T = \bigsqcup A$.

□

As every increasing sequence is a directed set, we immediately see that Theorem 5.3.1 applies to

any increasing sequence in $\mathcal{U}_s$. Furthermore, in such a case we may assume that each convergent cofinal subnet is a convergent infinite subsequence. Consequently, Theorem 5.3.1 directly implies the following corollary:

Corollary 5.3.2: Every increasing sequence $g_n \in \mathcal{U}_s$ has a convergent infinite subsequence such that $\lim(g_{n_k}) = \bigsqcup g_{n_k} = \bigsqcup g_n$.

Theorem 5.3.3: Every increasing sequence $g_n \in \mathcal{U}_s$ has a limit given by its supremum, explicitly $\lim(g_n) = \bigsqcup g_n$.

Proof: If (g_n) is an increasing sequence in $\mathcal{U}_s$, then by Corollary 5.3.2, we have that there exists a convergent infinite subsequence where $\lim(g_{n_k}) \to \bigsqcup g_n$. Furthermore, in Theorem 5.3.1 it was implicitly shown that all convergent infinite subsequence of (g_n) have the same limit. Lastly, since $\mathcal{U}$ is compact and $\mathcal{U}_s$ is a closed subset, it then follows that $\mathcal{U}_s$ is a compact Hausdorff space in the uniform metric topology. With this in mind, we invoke Lemma 3.1 in [20], where Martin shows that a sequence in a compact Hausdorff space converges to g if and only if all its convergent infinite subsequences converge to g. Consequently, we have that $\lim(g_n) = \lim(g_{n_k}) = \bigsqcup g_n$.

□

Theorem 5.3.4: Let $[0, \infty)^*$ denote the set of non-negative real numbers with the reverse standard order. Then the map $\mu : \mathcal{U}_s \to [0, \infty)^*$ given by

$$\mu(f) = \frac{1 + \text{tr}(f)}{4} \tag{5.36}$$

is a strictly monotone, Scott continuous function.

Proof: We prove this result by showing that μ is a monotone and strictly monotone map that preserves the supremum of every increasing sequence in $\mathcal{U}_s$. The proof is then complete by invoking

part (ii) of Theorem 2.2.1 from [17], which shows that these conditions imply Scott continuity.

From Lemma 5.2.8, we know that the trace function $\text{tr} : \mathcal{U}_s \to [0, \infty)^*$ is monotone and strictly monotone. The same is then true of μ, as it is simply $1 + \text{tr}$ divided by the constant 4. Furthermore, it then follows that

$$\begin{aligned} \mu(f) &= \tfrac{1}{4}\Big(1 + \text{tr}(f)\Big) \\ &= \tfrac{1}{4}\Big(1 + \sum_{i=1}^{3} f_{ii}\Big) = \tfrac{1}{4}\Big(1 + \sum_{i=1}^{3} \langle e_i, f e_i \rangle\Big), \end{aligned} \tag{5.37}$$

where for each $x \in S^2$ the inner product $\langle x, fx \rangle$ is a continuous function from $\mathcal{U}_s$ to $\mathbb{R}$. Consequently, μ is the sum of continuous functions, and is therefore also continuous.

With the continuity of μ in mind, we consider the increasing sequence $g_n \in \mathcal{U}_s$. We then have from Theorem 5.3.3 that the sequence is convergent with $\lim(g_n) = \bigsqcup g_n$. Therefore, by the continuity of μ, we have that

$$\mu\big(\bigsqcup g_n\big) = \mu\big(\lim g_n\big) = \lim\big(\mu g_n\big). \tag{5.38}$$

However, since μg_n is a bounded, monotone sequence in $\mathbb{R}$, it follows that $\lim\big(\mu g_n\big) = \bigsqcup(\mu g_n\big)$. Consequently,

$$\mu\big(\bigsqcup g_n\big) = \lim\big(\mu g_n\big) = \bigsqcup \mu(g_n). \tag{5.39}$$

Thus we have shown that μ is a monotone and strictly monotone map that preserves the supremum of increasing sequences. Therefore, by [17], μ is a Scott continuous function.

□

Remark: Let us take a moment to connect the function μ with some of our previous discussions found in section 5.2. By Theorem 5.2.5 we know that for every symmetric unital channel $f \neq I$ there exists a unique $p \in [0, 1)$ and a unique $m \in \max(\mathcal{U}_s)$ such that $f = pI + (1 - p)m$. It then

follows that

$$\begin{aligned}\mu(f) &= \tfrac{1}{4}\big[1 + \text{tr}(f)\big] = \tfrac{1}{4}\big[1 + \text{tr}(pI + (1-p)m)\big] \\ &= \tfrac{1}{4}\big[1 + 3p - (1-p)\big] = \tfrac{1}{4}4p = p.\end{aligned} \tag{5.40}$$

Therefore, for all $f \in \mathcal{U}_s$ we may replace Eq. (5.15) in Theorem 5.2.5 with the following statement:

$$f = (\mu f)I + (1 - \mu f)m. \tag{5.41}$$

With the strictly monotone Scott continuous map μ, we have the following result:

Corollary 5.3.5 (Martin): In the domain $(\mathcal{U}_s, \sqsubseteq)$ we may work with increasing sequences and their limits in lieu of the supremum of directed sets.

Proof: This result directly follows from Theorem 2.2.1 of [17] and Theorem 5.3.3 and 5.3.4 of this section.

□

Corollary 5.3.5 might be the most significant tool that will be utilized in this section, as it allows us to work with sequences and their limits instead of directed sets. If this does not appear paramount, we implore the reader to attempt the remaining proofs of this chapter using only directed sets. While this feat may be possible, doing so will leave one with a great appreciation for Martin's result. Let us now continue to further explore the structure of our partial order and ultimately characterize the approximation relation on the dcpo $(\mathcal{U}_s, \sqsubseteq)$.

Proposition 5.3.6: Let f be a symmetric unital channel with a non-zero fixed point. Then the channel f has a degenerate eigenvalue.

Proof: This is immediate from part (iii) of Corollary 5.2.6.

□

Before proceeding to the next two results, and ultimately our approximation relation, we would like to make a significant observation about Proposition 5.3.6. This result implicitly relies on the assumption that a quantum channel is a completely positive map. While this might seem insignificant, its importance is made apparent by the following implications. That is, Proposition 5.3.6 implies the next result, which in turn is used to characterize the approximation relation on the set of symmetric unital channels. In other words, *complete positivity of quantum channels results in our ability to characterize the approximation relation in* $(\mathcal{U}_s, \sqsubseteq)$.

Theorem 5.3.7: Let f and g be symmetric unital channels. If $f \sqsubseteq g$ and g has a non-zero fixed point, then f is a convex sum of g and the identity.

Proof: We begin by noting that when $g = I$, then because I is the least element $f \sqsubseteq g$ implies that $f = g$, and the desired result is immediate. The case where $f = I$ also follows trivially. Consequently, we will continue under the assumption that neither f or g are the identity. Therefore, since g has a non-zero fixed point, by part (iii) of Corollary 5.2.6 there exists some $r \in SO(3)$ and $p \in [0, 1]$ such that

$$g = pI + (1 - p)rs_1r^t. \tag{5.42}$$

Furthermore, since conjugation by any rotation is an order isomorphism, we may assume that g is diagonal. In other words, since $f \sqsubseteq g$ if and only if $r^t f r \sqsubseteq r^t g r$ for all $r \in SO(3)$, we may consider the channels $r^t f r$ and $r^t g r$ which we rename f and g respectively for the sake of brevity. It then follows from Proposition 5.3.6 that since g has a non-zero fixed point its spectrum is given

by $\sigma(g) = \{1, \mu, \mu\}$. Therefore, we may assume that

$$g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix}. \tag{5.43}$$

Recalling that $f \sqsubseteq g$, it then follows from direct calculation that

$$\langle e_1, [f - g]e_1 \rangle = f_{11} - g_{11} = f_{11} - 1 \geq 0, \tag{5.44}$$

which is true if and only if $f_{11} \geq 1$. Moreover, since f is non-expansive $\big($i.e. $||fe_1|| \leq ||e_1||\big)$ we have that

$$\begin{aligned} \langle e_1, fe_1 \rangle &= ||e_1||\,||fe_1|| \cos(\theta) \\ &= ||fe_1|| \cos(\theta) \leq ||e_1|| \cos(\theta) = \cos(\theta), \end{aligned} \tag{5.45}$$

where θ is the angle between the vectors e_1 and fe_1. Therefore, $f_{11} = \langle e_1, fe_1 \rangle \leq 1$ where equality is assumed if and only if $fe_1 = e_1$. This directly implies that $f_{21} = f_{31} = 0$, otherwise $fe_1 = [1, f_{21}, f_{31}]^t$ and $fe_1 \neq e_1$. Lastly, since $f = f^t$, we are left to conclude that f has the following form:

$$f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & b & c \end{pmatrix}. \tag{5.46}$$

Upon calculation, the eigenvalues of f are given by:

$$\begin{aligned} \lambda_1 &= 1, \\ \lambda_2 &= \tfrac{1}{2}\Big(a + c + \sqrt{(a-c)^2 + 4b^2}\Big), \\ \lambda_3 &= \tfrac{1}{2}\Big(a + c - \sqrt{(a-c)^2 + 4b^2}\Big). \end{aligned} \tag{5.47}$$

However, again utilizing Proposition 5.3.6, it follows that $1 \in \sigma(f)$ implies $\lambda_2 = \lambda_3$. Therefore, setting $\lambda = \lambda_2 = \lambda_3$, the following statements are equivalent:

$$\begin{aligned} \tfrac{1}{2}\Big(a + c + \sqrt{(a-c)^2 + 4b^2}\Big) &= \tfrac{1}{2}\Big(a + c - \sqrt{(a-c)^2 + 4b^2}\Big) \\ \sqrt{(a-c)^2 + 4b^2} &= -\sqrt{(a-c)^2 + 4b^2} \\ \sqrt{(a-c)^2 + 4b^2} &= 0. \end{aligned} \tag{5.48}$$

Squaring both sides, we have that $(a-c)^2 + 4b^2 = 0$. Moreover, since $a, b, c \in \mathbb{R}$, it follows that that $(a-c)^2$ and $4b^2$ are both non-negative. Therefore, their sum is zero if and only if they are both zero. That is, $a = c$ and $b = 0$. Consequently,

$$f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \tag{5.49}$$

Lastly, since $f \sqsubseteq g$ $\big($i.e. $\langle x, [f-g]x\rangle \geq 0$ for all $x \in S^2\big)$, it then follows that $\lambda \geq \mu$. Therefore, letting $p = \frac{\lambda - \mu}{1 - \mu} \in [0, 1]$, we have that $(1-p) = \frac{1-\lambda}{1-\mu}$ and

$$\begin{aligned} \lambda(1-\mu) &= \lambda(1-\mu) + \mu - \mu = \lambda - \mu + \mu - \lambda\mu \\ &= (\lambda - \mu) + (1-\lambda)\mu = (1-\mu)\big[p + (1-p)\mu\big]. \end{aligned} \tag{5.50}$$

Dividing each side by $(1-\mu)$, Eq. (5.50) implies that $\lambda = p + (1-p)\mu$ and we finally have that

$$\begin{aligned} f &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} p+(1-p) & 0 & 0 \\ 0 & p+(1-p)\mu & 0 \\ 0 & 0 & p+(1-p)\mu \end{pmatrix} \\ &= p \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (1-p) \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix} = pI + (1-p)g. \end{aligned} \tag{5.51}$$

□

Corollary 5.3.8: Let $f, g \in \mathcal{U}_s$. If $1 \in \sigma(g)$ and $f \sqsubseteq g$, then any matrix that diagonalizes g also diagonalizes f.

Proof: This result is immediate from the proof of Theorem 5.3.7 as it depends only of the rotation that diagonalizes g.

□

Theorem 5.3.9: The approximation relation for each symmetric unital channel $g \neq I$ is characterized as follows:

(i) If $1 \notin \sigma(g)$, then $f \ll g \iff f - g > 0$,

(ii) If $1 \in \sigma(g)$, then $f \ll g \iff f \neq g$ and $f \sqsubseteq g$,

where $f - g > 0$ denotes the condition that $\langle x, [f-g]x \rangle > 0$ for all $x \in S^2$.

Proof:

(i) $\underline{1 \notin \sigma(g)}$:

($\Leftarrow$) : Recall from Corollary 5.3.5, instead of using directed subsets and their supremum we may instead work with the limits of increasing sequences. With this in mind, considering an increasing sequence $(y_n) \to \lim(y_n)$ where $g \sqsubseteq \lim(y_n)$, we show that $f - g > 0$ implies that there exists some y_k such that $f \sqsubseteq y_k$, i.e. $f \ll g$. We begin by defining the function $\phi : \mathcal{U}_s \times S^2 \to \mathbb{R}$ where

$$\phi(g, x) = \langle x, [f - g]x \rangle. \tag{5.52}$$

In Lemma 2.4.2 we have implicitly showed that ϕ is continuous, and therefore, by Corollary 2.10 in reference [22], the map $\underline{\phi} : \mathcal{U}_s \to \mathbb{R}$

$$\underline{\phi}(g) = \inf_{x \in S^2} \langle x, [f - g]x \rangle \tag{5.53}$$

is also continuous. Note, since S^2 is compact, we know that the infimum of $\langle x, [f - g]x \rangle$ is assumed by some $x \in S^2$. It then follows that because $f - g > 0$ and $g - \lim(y_n) \geq 0$, we have

$$\begin{aligned} \langle x, [f - \lim(y_n)]x \rangle &= \langle x, [f - g + g - \lim(y_n)]x \rangle \\ &= \langle x, [f - g]x \rangle + \langle x, [g - \lim(y_n)]x \rangle > 0 \end{aligned} \tag{5.54}$$

for all $x \in S^2$. Consequently, by Eqs. (5.53) and (5.54), and the fact that the continuous function $\underline{\phi}$ assumes its absolute infimum on the compact set S^2,

$$\lim\big(\underline{\phi}(y_n)\big) = \underline{\phi}\big(\lim(y_n)\big) > 0. \tag{5.55}$$

It then follows that $\underline{\phi}(y_n) > 0$, i.e. $\langle x, [f - y_n]x \rangle > 0$, for most n. Thus, we have shown that the condition $f - g > 0$ implies that for any increasing sequence (y_n) where $g \sqsubseteq \lim(y_n) = \bigsqcup y_n$, there exists some y_k such that $f \sqsubseteq y_k$. Therefore, $f \ll g$.

($\Rightarrow$) : We begin by noting that by the definition of approximation, if $f \ll g$, then for every increasing sequence (y_n) where $g \sqsubseteq \lim(y_n) = \bigsqcup y_n$, there exists some y_k such that $f \sqsubseteq y_k$ (since the set of all y_n is directed itself). With this in mind, let $y_n = \frac{1}{n}I + (1 - \frac{1}{n})g$. We then have that

$$\begin{aligned} y_n - y_{n+1} &= \big(\tfrac{1}{n} - \tfrac{1}{n+1}\big)I + \big(1 - \tfrac{1}{n} - 1 + \tfrac{1}{n+1}\big)g \\ &= \big(\tfrac{1}{n} - \tfrac{1}{n+1}\big)I + \big(\tfrac{1}{n+1} - \tfrac{1}{n}\big)g = \big(\tfrac{1}{n} - \tfrac{1}{n+1}\big)(I - g) > 0 \end{aligned} \tag{5.56}$$

where the last inequality follows from the fact that $g = g^t$ and $1 \notin \sigma(g)$ implies that $\langle x, gx\rangle$ is less than g's largest eigenvalue which is necessarily less than 1 while $\langle x, Ix\rangle = 1$, for all $x \in S^2$. Furthermore,

$$\begin{aligned} y_n - g &= \tfrac{1}{n}I + \big(1 - \tfrac{1}{n}\big)g - g \\ &= \tfrac{1}{n}I + \big(1 - \tfrac{1}{n} - 1\big)g = \tfrac{1}{n}(I - g) > 0, \end{aligned} \tag{5.57}$$

where the last inequality follows by the same arguments used in inequality (5.56). That is, (y_n) is an increasing sequence with $\bigsqcup y_n = \lim(y_n) = g$ and $g \sqsubseteq \bigsqcup y_n$. Then since $f \ll g$, there exists some y_k where $f \sqsubseteq y_k$, i.e. $f - y_k \geq 0$. Consequently, we have

$$f - g = f - y_k + y_k - g = (f - y_k) + (y_k - g) > 0 \tag{5.58}$$

where the last inequality is true because $f - y_k \geq 0$ and $y_k - g > 0$ by Eq. (5.57). Therefore, we have shown that if $1 \notin \sigma(g)$, then $f \ll g$ if and only if $f - g > 0$.

(ii) $\underline{1 \in \sigma(g)}$:

We begin by noting that if $1 \in \sigma(g)$, then there does not exist a $f \in \mathcal{U}_s$ where $f - g > 0$.

Otherwise, for the point x^* where $gx^* = x^*$, we would have that

$$\langle x^*, [f-g]x^* \rangle = \langle x^*, fx^* \rangle - 1 > 0. \tag{5.59}$$

However, recalling Eq. (5.45), which results from the fact that the unital channels are non-expansive, we have that $\langle x^*, fx^* \rangle \leq \cos(\theta) \leq 1$, where θ is the angle between the vectors x^* and fx^*. Therefore, we are left to conclude that $f - g \not> 0$. With this in mind, it is clear that we need a different relation for the set of unital channels with non-zero fixed points. Otherwise, nothing would approximate these channels, which contradicts the fact that the least element I approximates every channel.

($\Leftarrow$): We use some of the same arguments for this proof as those found in part (i). In particular, we use Corollary 5.3.5 which tells us that we can work with increasing sequences and their limits instead of supremum of directed subsets. With this in mind, we assume $f \sqsubseteq g$, $f \neq g$, and let (y_n) be an increasing sequence where $g \sqsubseteq \lim(y_n) = y$. We then arrive at one of two cases: $1 \notin \sigma(y)$ or $1 \in \sigma(y)$. We begin with the former.

(a) $\underline{1 \notin \sigma(y)}$:

We remind the reader that we are now in the case where $1 \in \sigma(g)$. Since conjugation by a rotation is an order isomorphism, we may consider the case where g is diagonal, in particular

$$g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix}. \tag{5.60}$$

Therefore, if $f \neq g$, and $f \sqsubseteq g$, we have by Corollary 5.3.8 that

$$f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \tag{5.61}$$

where $\lambda > \mu$. It then follows that since $f - g \geq 0$ and $g - y \geq 0$, we have for all $x \in S^2$

$$\langle x, [f - y]x \rangle = \langle x, [f - g]x \rangle + \langle x, [g - y]x \rangle \geq 0. \tag{5.62}$$

Consequently, since $\langle x, [f-g]x \rangle$ and $\langle x, [g-y]x \rangle$ are both non-negative, so is their sum. Therefore, there exists some $x \in S^2$ where $\langle x, [f - y]x \rangle = 0$ if and only if

$$\langle x, [f - g]x \rangle = \langle x, [g - y]x \rangle = 0. \tag{5.63}$$

With this in mind, because g and f are diagonal matrices given by Eqs. (5.60) and (5.61) respectively and $\lambda > \mu$, it follows that

$$\langle x, \big[f - g\big]x \rangle = (x_2^2 + x_3^2)(\lambda - \mu) = 0 \tag{5.64}$$

if and only if $x = e_1$. However, since $ge_1 = e_1$ and $y_{11} = \langle e_1, ye_1 \rangle$ is less than or equal to its largest eigenvalue $\big($which is less than 1 since y is non-expansive and does not have a fixed point in S^2 $\big)$, it is also true that

$$\langle e_1, [g - y]e_1 \rangle = 1 - y_{11} \neq 0. \tag{5.65}$$

In other words, $\langle e_1, [g - y]e_1 \rangle > 0$. Consequently, there does not exist an $x \in S^2$ where $\langle x, [f - g]x \rangle = \langle x, [g - y]x \rangle = 0$, and we are left to conclude that $\langle x, [f - y]x \rangle > 0$ for

all $x \in S^2$, i.e. $f - y = f - \lim(y_n) > 0$. We then have by the definition of $\underline{\phi}$ that

$$\underline{\phi}\big(\lim y_n\big) = \inf_{x \in X} \langle x, [f - \lim y_n]x \rangle > 0, \tag{5.66}$$

where $\underline{\phi}$ assumes its absolute infimum on the compact set S^2. It then follows from the continuity of $\underline{\phi}$ that $\underline{\phi}\big(\lim y_n\big) = \lim \phi(y_n) > 0$, and therefore, $f - y_n > 0$ for most n. Thus, when $1 \notin \sigma(y)$, we have that $f \ll g$.

(b) $\underline{1 \in \sigma(y)}$:

As this proof is long, we remind the reader once again that we are considering an increasing sequence $(y_n) \to y$ where $1 \in \sigma(g)$, $f \neq g$, $f \sqsubseteq g$, and $g \sqsubseteq \bigsqcup y_n = y$. Therefore, $f \sqsubseteq g \sqsubseteq y$ and $y_n \sqsubseteq y$. It then follows from Corollary 5.3.8 that if $1 \in \sigma(y)$, then any rotation r that diagonalizes y also diagonalizes f, g, and each y_n. Consequently, since conjugation by a rotation is an order isomorphism, by Theorem 5.3.7 we may consider the following:

$$f = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}, \quad g = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{bmatrix}, \quad y = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{bmatrix}, \quad y_n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha_n & 0 \\ 0 & 0 & \alpha_n \end{bmatrix} \tag{5.67}$$

where $\alpha_n \to \alpha$. Furthermore, since $y_n \sqsubseteq y$ and $f \sqsubseteq g \sqsubseteq y$, we have that $\alpha_n \geq \alpha$ and $\lambda > \mu \geq \alpha$, where the strict inequality follows from the fact that $f \neq g$. It then follows that

$$\lambda - \alpha = \lambda - \lim(\alpha_n) = \lim(\lambda - \alpha_n) > 0, \tag{5.68}$$

and therefore, $\lambda - \alpha_n > 0$ for most n. Consequently, there exists some y_k such that $f \sqsubseteq y_k$, and we are left to conclude that $f \ll g$.

Thus we have shown that when $1 \in \sigma(g)$, if $f \neq g$ and $f \sqsubseteq g$, then $f \ll g$.

($\Rightarrow$): Conversely, let $f \ll g$ and $1 \in \sigma(g)$. By the definition of approximation we immediately have that $f \sqsubseteq g$. Therefore, we need only show that $f \neq g$. With this in mind, we once again consider the sequence $y_n = \frac{1}{n}I + (1 - \frac{1}{n})g$. Consequently, this sequence is increasing where $\lim(y_n) = \bigsqcup y_n = g$. Therefore, since $f \ll g$, there exists some y_k such that $f \sqsubseteq y_k$. Furthermore, $f = g$ if and only if

$$f - g = \big(f - y_k\big) + \big(y_k - g\big) = 0, \tag{5.69}$$

where we know that $f - y_k \geq 0$ and $y_k - g \geq 0$. It then follows that $f = g$ if and only if $f = y_k$ and $g = y_k$. On the other hand,

$$\begin{aligned} y_k - g &= \tfrac{1}{k}I + \big(1 - \tfrac{1}{k}\big)g - g \\ &= \tfrac{1}{k}I + \big(1 - \tfrac{1}{k} - 1\big)g = \tfrac{1}{k}(I - g) \geq 0 \end{aligned} \tag{5.70}$$

where the last inequality follows from the fact that I is the least element. Therefore, $g = y_k$ if and only if $g = I$. However, by assumption $g \neq I$, and it then follows that $f \neq g$ and $f \sqsubseteq g$.

Thus, when $1 \in \sigma(g)$, we have shown that $f \sqsubseteq g$ and $f \neq g$ if and only if $f \ll g$.

□

Corollary 5.3.10: A symmetric unital channel approximates itself if and only if it is the identity.

Proof: If $f = I$, then it is the least element and for every directed subset $A \subseteq \mathcal{U}_s$ it follows that $f \sqsubseteq \bigsqcup A$. Furthermore, for every element $g \in A$, it is also the case that $f \sqsubseteq g$. Therefore, $f \ll f$. However, when $f \neq I$, then either $1 \notin \sigma(f)$ or $1 \in \sigma(f)$. By Theorem 5.3.9, we have that when $1 \notin \sigma(f)$, then $f \ll f$ if and only if $f - f = 0 > 0$, which is never true. Similarly, when $1 \in \sigma(f)$,

then $f \ll f$ if and only if $f \neq f$ and $f \sqsubseteq f$, which again is never true. Thus we have shown that $f \ll f$ if and only if $f = I$.

□

Corollary 5.3.11: Every symmetric unital channel either approximates a maximal element of is a maximal element itself.

Proof: If $f = I$, then it approximates everything in $\mathcal{U}_s$, and the result is trivial. However, if $f \neq I$, then by Theorem 5.2.5, there exists a unique $p \in [0, 1)$ and a unique $m \in \max\left(\mathcal{U}_s\right)$ such that

$$f = pI + (1 - p)m. \tag{5.71}$$

Therefore, when $f \notin \max(\mathcal{U}_s)$, we have that $f \neq m$ and

$$f - m = pI + (1 - p)m - m = p(I - m) \geq 0. \tag{5.72}$$

Since I is the least element, when $1 \in \sigma(f)$, it then follows that $f \sqsubseteq m$ and $f \neq m$. In other words, by Theorem 5.3.9, $f \ll m$. Similarly, when $1 \notin \sigma(f)$, then by Eqs. (5.71) and (5.72), $1 \notin \sigma(m)$, and $f - m = p(I - m) > 0$. That is, we have again that $f \ll m$. Thus we have shown that for every symmetric unital channel $f \notin \max\left(\mathcal{U}_s\right)$, there exists a $m \in \max\left(\mathcal{U}_s\right)$ where $m \in \Uparrow f$.

□

With an approximation relation established, we may now finally verify whether or not $(\mathcal{U}_s, \sqsubseteq)$ is a continuous dcpo. We start with the following lemma, whose proof is admitted as it is well-documented in the literature.

Lemma 5.3.12 (Abramsky & Jung): Let X be a dcpo. If $A \subseteq \downarrow x$ where A is a directed subset and $\bigsqcup A = x$, then $\downarrow x$ is directed with $\bigsqcup \downarrow x = x$.

Theorem 5.3.13: The set of symmetric unital channels is a continuous dcpo.

Proof: Let $f = I$. Then $\downarrow f = \emptyset$ and the result follows trivially. Therefore, for the remainder of the proof we assume that $f \neq I$. Let (y_n) be a sequence of functions $y_n : \mathcal{U}_s \to \mathcal{U}_s$ where

$$y_n(f) = \tfrac{1}{n} I + \big(1 - \tfrac{1}{n}\big) f. \tag{5.73}$$

Then for every symmetric unital channel, $y_n(f)$ is a sequence in $\mathcal{U}_s$ that converges to f. In particular, since I is the least element and

$$\begin{aligned} y_n(f) - y_{n+1}(f) &= \tfrac{1}{n} I + \big(1 - \tfrac{1}{n}\big) f - \tfrac{1}{n+1} I - (1 - \tfrac{1}{n+1}) f \\ &= \big(\tfrac{1}{n} - \tfrac{1}{n+1}\big) I + \big(\tfrac{1}{n+1} - \tfrac{1}{n}\big) f \\ &= \big(\tfrac{1}{n} - \tfrac{1}{n+1}\big)(I - f), \end{aligned} \tag{5.74}$$

$y_n(f)$ is an increasing sequence for all $f \in \mathcal{U}_s$, and therefore, a directed set. Furthermore,

$$y_n(f) - f = \tfrac{1}{n} I + \big(1 - \tfrac{1}{n}\big) f - f = \tfrac{1}{n}(I - f). \tag{5.75}$$

Consequently, since $f \neq I$ by assumption, we have that $f \neq y_n(f)$ for all n. Therefore, when $1 \in \sigma(f)$, we have that $y_n(f) \neq f$ and $y_n(f) \sqsubseteq f$, which implies $y_n(f) \ll f$.

On the other hand, when $1 \notin \sigma(f)$, since $\langle x, [I - f]x\rangle > 0$ for all $x \in S^2$, it then follows that $y_n(f) - f > 0$. In other words, we still have that $y_n(f) \ll f$. Therefore, in either case $y_n(f) \ll f$ for all n.

We then have by Theorem 5.3.3, that $\lim(y_n(f)) = \bigsqcup y_n(f) = f$, and we have thus shown that

the sequence $(y_n(f))$ is a directed subset of $\downarrow f$ with $\bigsqcup y_n(f) = f$. It then follows from Lemma 5.3.12, that for all $f \in \mathcal{U}_s$, the way below set $\downarrow f$ is directed with supremum f. In other words, the dcpo $(\mathcal{U}_s, \sqsubseteq)$ is continuous.

□

Having now shown that $(\mathcal{U}_s, \sqsubseteq)$ is a domain, we would like to conclude this chapter by further exploring the properties of the function $\mu : \mathcal{U}_s \to [0, \infty)^*$, where

$$\mu(f) = \frac{1 + \text{tr}(f)}{4}. \tag{5.76}$$

In particular, while the informatic order $\sqsubseteq$ on $\mathcal{U}_s$ provides a qualitative view of content, we will see that the *measurement* μ will provide a corresponding quantitative notion.

5.4 A Measurement on the Set of Unital Channels

Theorem 5.4.1: The strictly monotone Scott continuous map $\mu : \mathcal{U}_s \to [0, \infty)^*$ given by

$$\mu(f) = \frac{1 + \text{tr}(f)}{4} \tag{5.77}$$

measures all of $\mathcal{U}_s$.

Proof: Let $f \in \uparrow g$ where f and g are arbitrary symmetric unital channels. We prove that there exists an $\epsilon \in [0, \infty)$ such that $f \in \mu_\epsilon(f) \subseteq \uparrow g$. We will do so by considering the two cases where $1 \in \sigma(f)$ and $1 \notin \sigma(f)$.

(i) $1 \in \sigma(f)$:

We show that for all $g \in \downarrow f$, there exists an ϵ where $h \in \mu_\epsilon(f) = \{h \sqsubseteq f \mid \epsilon \ll \mu(h)\}$

implies that $g \ll h$. Since $1 \in \sigma(f)$ and $g, h \sqsubseteq f$, then because conjugation by a rotation is an order isomorphism, by Corollary 5.3.8 we may consider f, g, and h to be given by the following matrices:

$$f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \quad g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \beta \end{pmatrix}, \tag{5.78}$$

where $\lambda < \alpha$ and $\lambda \leq \beta$. The strict inequality follows from Theorem 5.3.9 where we have shown that when $1 \in \sigma(f)$, then $g \ll f$ if and only if $g \sqsubseteq f$ and $g \neq f$. Then letting

$$\epsilon = \frac{1+\alpha}{2}, \tag{5.79}$$

for any $h \in \mu_\epsilon(f)$ we have that

$$\epsilon = \frac{1+\alpha}{2} > \mu(h) = \frac{1+\text{tr}(h)}{4} = \frac{1+\beta}{2}. \tag{5.80}$$

Consequently, $\alpha > \beta$, and it then follows that $g \sqsubseteq h$ and $g \neq h$. Lastly, because $1 \in \sigma(h)$, we have again from Theorem 5.3.9 that $g \ll h$. Furthermore, since $\lambda < \alpha$, it follows that

$$\mu(f) = \frac{1+\lambda}{2} < \frac{1+\alpha}{2}. \tag{5.81}$$

Therefore, $\epsilon > \mu(f)$, and we also have that $f \in \mu_\epsilon(f)$. Thus we have shown that if $1 \in \sigma(f)$, then for each $g \in \downarrow f$, there exists an ϵ such that $f \in \mu_\epsilon(f) \subseteq \uparrow g$.

(ii) $\underline{1 \notin \sigma(f)}$:

For this case we show that for all $g \in \downarrow f$, if

$$\epsilon = \mu(f) + \tfrac{1}{4} \min_{x \in S^2} \langle x, [g - f]x \rangle, \tag{5.82}$$

then $f \in \mu_\epsilon(f) \subseteq \uparrow g$. First off, because $\langle x, [g - f]x \rangle$ is continuous, its minimum value is assumed by some point in the compact set S^2. Furthermore, since $g \ll f$ with $1 \notin \sigma(f)$, we have that $g - f > 0$, and thus $\min_{x \in S^2} \langle x, [g - f]x \rangle > 0$. It then follows that $\epsilon > \mu(f)$, and therefore, $f \in \mu_\epsilon(f)$.

We begin by assuming $h \in \mu_\epsilon(f)$, and let $\eta_3 \leq \eta_2 \leq \eta_1$ be the eigenvalues of the matrix $h - f$. Since $h \in \mu_\epsilon(f)$, we have that $h \sqsubseteq f$, i.e. $h - f$ is positive semi-definite. Therefore, $\eta_i \geq 0$ for all i. It then follows that since $h - f$ is symmetric,

$$\begin{aligned}
\max_{x \in S^2} \langle x, [h - f]x \rangle = \eta_1 &= \text{tr}(h - f) - \eta_2 - \eta_3 \\
&= \text{tr}(h) - \text{tr}(f) - \eta_2 - \eta_3 \\
&\leq \text{tr}(h) - \text{tr}(f) \\
&= 4\Big[\mu(h) - \mu(f)\Big] \\
&< 4\Big[\epsilon - \mu(f)\Big] \\
&= 4\Big[\tfrac{1}{4} \min_{x \in S^2} \langle x, [g - f]x \rangle + \mu(f) - \mu(f)\Big] \\
&= \min_{x \in S^2} \langle x, [g - f]x \rangle,
\end{aligned} \tag{5.83}$$

where the first inequality follows from the fact that $\eta_2, \eta_3 \geq 0$ while the second inequality is due to the assumption that $\mu(h) < \epsilon$, whose value is given in Eq. (5.82). Simply put, we have shown that

$$\min_{x \in S^2} \langle x, [g - f]x \rangle - \max_{x \in S^2} \langle x, [h - f]x \rangle > 0. \tag{5.84}$$

Then by basic arithmetic, we finally have that

$$\begin{aligned}
\min_{x\in S^2}\langle x,[g-h]x\rangle &= \min_{x\in S^2}\Big[\langle x,[g-f]x\rangle + \langle x,[f-h]x\rangle\Big] \\
&\geq \min_{x\in S^2}\langle x,[g-f]x\rangle + \min_{x\in S^2}\langle x,[f-h]x\rangle \\
&= \min_{x\in S^2}\langle x,[g-f]x\rangle + \min_{x\in S^2}\Big[-\langle x,[h-f]x\rangle\Big] \\
&= \min_{x\in S^2}\langle x,[g-f]x\rangle - \max_{x\in S^2}\langle x,[h-f]x\rangle \\
&> 0,
\end{aligned} \tag{5.85}$$

where the last inequality follows directly from inequality (5.84). Therefore, we have that $g-h>0$. Of course, this implies that $1\notin\sigma(h)$, otherwise the inner product $\langle x,[g-h]x\rangle\leq 0$ when $hx=x$. Consequently, by Theorem 5.3.9 we have that $g\ll h$. Thus, $f\in\mu_\epsilon(f)\subseteq\uparrow g$, as desired.

Therefore, the map μ measures the content of every symmetric unital channel.

□

We conclude this chapter with an interesting observation about the way above sets $\uparrow f$ in $(\mathcal{U}_s,\sqsubseteq)$, which form a basis for the Scott topology. We start with the following lemma:

Lemma 5.4.2: For every $r\in SO(3)$ such that $re_1=e_1$, there exists a sequence of rotations (r_n) that converges to r where $r_ne_1\neq e_1$ for all n.

Proof: It is a well-documented result that every element of $SO(3)$ can be written as a multiplication of principal rotations about the standard basis elements in $\mathbb{R}^3$. Therefore, we may consider the sequence $r_n=r_1(\alpha)r_2(\beta_n)r_3(\gamma_n)$ where each r_i is a principal rotation about e_i. Furthermore, if we let $\beta_n\to 0$, and $\gamma_n\to 0$, it then follows that, $r_n\to r=r_1(\alpha)$, where $re_1=e_1$. On the other

hand, for each n

$$r_n e_1 = r_x(\alpha) r_y(\beta_n) r_z(\gamma_n) e_1 = \begin{bmatrix} \cos\beta_n \cos\gamma_n \\ \cos\alpha \sin\gamma_n + \sin\alpha \sin\beta_n \cos\gamma_n \\ \sin\alpha \sin\gamma_n + \cos\alpha \sin\beta_n \cos\gamma_n \end{bmatrix}. \tag{5.86}$$

It then follows that for any angle α, if we let $0 < \beta_n, \gamma_n < \pi$, then $\cos\beta_n \cos\gamma_n \neq 1$, and we therefore have that $r_n e_1 \neq e_1$ for each r_n. In other words, (r_n) is a sequence of rotations that do not fix e_1 but whose limit does. Note, this argument can be generalized from e_1 to any non-zero vector in $\mathbb{R}^3$.

□

Theorem 5.4.3: Let f be a symmetric unital channel.

(i) If $1 \notin \sigma(f)$, then $\uparrow\!\!\uparrow f$ is open.

(ii) If $1 \in \sigma(f)$, then $\uparrow\!\!\uparrow f$ is both open and closed, or neither.

In particular, for case (ii) we have that if $f \in \max\left(\mathcal{U}_s\right)$, then $\uparrow\!\!\uparrow f = \emptyset$, and if $f = I$, then $\uparrow\!\!\uparrow f = \mathcal{U}_s$.

Proof:

(i) $1 \notin \sigma(f)$:

We begin by showing that when $1 \notin \sigma(f)$ and $g \in \uparrow\!\!\uparrow f$, then $1 \notin \sigma(g)$. This result is significant since it, along with Theorem 5.3.9, implies that when $1 \notin \sigma(f)$, then $g \in \uparrow\!\!\uparrow f$ if and only if $f - g > 0$.

If $1 \notin \sigma(f)$, then for any $g \in \mathcal{U}_s$ with $1 \in \sigma(g)$, for the point $x^* \in S^2$ where $gx^* = x^*$

$$\langle x^*, [f-g]x^* \rangle = \langle x^*, fx^* \rangle - 1 < 0 \tag{5.87}$$

due to the fact that $\langle x, fx \rangle$ is bounded from above by the largest eigenvalue of f, which is less than 1 since $1 \notin \sigma(f)$. Consequently, this implies that $f \not\sqsubseteq g$, which contradicts the assumption that $f \ll g$. Thus, when $1 \notin \sigma(f)$, if $g \in \Uparrow f$ then $1 \notin \sigma(g)$. Therefore, because $1 \notin \sigma(g)$ for any $g \in \Uparrow f$, we know from Theorem 5.3.9 that $g \in \Uparrow f$ if and only if $f - g > 0$. Then for the function $\underline{\phi} : \mathcal{U}_s \to \mathbb{R}$

$$\underline{\phi}(g) = \inf_{x \in S^2} \langle x, [f-g]x \rangle, \tag{5.88}$$

we have that $\underline{\phi}(g) > 0$ when $g \in \Uparrow f$ and the infimum of $\langle x, [f-g]x \rangle$ is assumed at some x in the compact set S^2. Furthermore, by the continuity of $\underline{\phi}$ (proven in the arguments of Theorem 5.3.9) there exists an open set U where $g \in U \subseteq \underline{\phi}^{-1}(0, \infty)$. This implies that if $h \in U$, then $f - h > 0$, and as seen earlier, $1 \notin \sigma(h)$. Therefore, by Theorem 5.3.9, $h \in \Uparrow f$ and it follows that $g \in U \subseteq \Uparrow f$. Consequently, the set $\Uparrow f$ is open in the uniform metric topology τ.

(ii) $\underline{1 \in \sigma(f)}$:

If f is maximal, then $\Uparrow f = \emptyset$. Consequently, $\Uparrow f$ is both open and closed in every topology. On the other hand, if $f = I$, then by Theorem 5.3.9 and Corollary 5.3.10 $\Uparrow f = \mathcal{U}_s$, which is also both open and closed. Therefore, when f is a maximal element or the identity, the desired result is immediate.

We now consider the case where f is neither the identity or a maximal element. Because $1 \in \sigma(f)$, by Corollary 5.2.6 we may assume that $f = pI + (1-p)rs_1r^t$ where $p \in [0, 1)$. Furthermore, since conjugation by a rotation is an order isomorphism, we may also assume

that

$$f = pI + (1-p)s_1. \tag{5.89}$$

Lastly, by Lemma 5.4.2 there exists a sequence $(r_n) \to r$ where each r_n does not fix e_1 and $re_1 = e_1$. Therefore, if we let $q < p$ and

$$g_n = qI + (1-q)r_n^t s_1 r_n, \tag{5.90}$$

then g_n fixes the point $r_n^t e_1 \neq e_1$ while f does not, and $g_n \to g = qI + (1-q)s_1$. Consequently, for each n

$$\langle r_n^t e_1, [f - g_n] r_n^t e_1 \rangle = \langle r_n^t e_1, f r_n^t e_1 \rangle - 1 < 0 \tag{5.91}$$

due to the fact that $r_n^t e_1$ is the fixed point of g_n and $\max\limits_{x \in S^2} \langle x, fx \rangle = 1$ is obtained only when $x = e_1$. It then follows that, $f \not\sqsubseteq g_n$, and therefore, $g_n \notin \uparrow\!\!\uparrow f$ for all n. On the other hand, because $q < p$, we have that

$$\begin{aligned} f - g &= pI + (1-p)s_1 - qI - (1-q)s_1 \\ &= I(p-q) + (q-p)s_1 \\ &= (p-q)(I - s_1) \geq 0, \end{aligned} \tag{5.92}$$

where the last inequality follows from the fact that I is the least element. Consequently, we have that $f \sqsubseteq g$ and $f \neq g$. In other words, the sequence $g_n \in \mathcal{U}_s \backslash \uparrow\!\!\uparrow f$, converges to $g \in \uparrow\!\!\uparrow f$. Thus the set $\mathcal{U}_s \backslash \uparrow\!\!\uparrow f$ does not contain all its limit points, and is therefore not closed. It then immediately follows that $\uparrow\!\!\uparrow f$ is not open in τ.

Similarly, let $f_n = p_n I + (1-p_n)s_1$ where $p_n < p$ and $p_n \to p$. It then follows that $\left(f_n\right) \to f$

and

$$\begin{aligned} f - f_n &= pI + (1-p)s_1 - p_n I - (1-p_n)s_1 \\ &= (p - p_n)(I - s_1) \geq 0, \end{aligned} \tag{5.93}$$

since I is the least element. Consequently, $f_n \in \uparrow\!\!\uparrow f$ and $\lim\left(f_n\right) = f$. However, since f is not the identity we have from Corollary 5.3.10 that $f \notin \uparrow\!\!\uparrow f$. Thus, $\uparrow\!\!\uparrow f$ does not contain all its limit points and we are left to conclude that $\uparrow\!\!\uparrow f$ is also not closed in τ. Therefore, we have shown that if $1 \in \sigma(f)$, and f is not the identity or a maximal element, then $\uparrow\!\!\uparrow f$ is neither open or closed in τ.

□

5.5 Summary

While this chapter is far more abstract than the rest of this dissertation, the verification of mathematical structure has proven to be invaluable in the exploration of new science. For instance, this dissertation began with practical applications in qubit communication. Specifically, in the context of Adaptive Quantum Information Processing, we have shown that the set of symmetric unital channels is sufficient when calculating error rates. Furthermore, we obtained explicit bounds on the possible reduction of error rates for qubit communication protocols. All of this was accomplished due to previous discoveries in the mathematical foundations of quantum information. In hopes of furthering this foundation, we have shown that not only does this same set of quantum channels form a domain, but there exists a physically meaningful measurement. This result gives us a plethora of rich structure to explore and provides the potential for further applications.

APPENDIX: SUPPLEMENTARY THEOREMS

The purpose of this appendix is to prove generalizations for some of the original theorems that were referenced to throughout this dissertation.

Theorem 1: Let $E_{q,f} \in \mathbb{E}$. Assume f is a symmetric unital channel with eigenvalues $\lambda_3 \leq \lambda_2 \leq \lambda_1$ that is diagonalized by $r \in SO(3)$, and the functions $p_j : X \to [0,1]$ are given by

$$p_j(x) = \sum_{i=1}^{n} q_i (r x_i)_j^2. \tag{1}$$

If $\max_{x \in X}(p_j) = \max_{x \in X}(p_k)$ for all $j \neq k$, then

$$\mathrm{Im}\big(E_{q,f}\big) = \left[\tfrac{1}{2}\big(1 - a\lambda_1 - b\lambda_2 - c\lambda_3\big), \tfrac{1}{2}\big(1 - c\lambda_1 - b\lambda_2 - a\lambda_3\big) \right] \tag{2}$$

where $a = \max_{x \in X}(p_i)$,

$$b = \begin{cases} 1 - a, & \text{if } a \geq \frac{1}{2} \\ a, & \text{otherwise,} \end{cases} \tag{3}$$

and $c = 1 - a - b$.

Proof: We begin by noting that since f is symmetric it can be diagonalized by some rotation r. That is,

$$E_{q,f}(x) = \tfrac{1}{2}\left[1 - \sum_{i=1}^{n} q_i \langle r x_i, \lambda r x_i \rangle \right] \tag{4}$$

where

$$\lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}. \tag{5}$$

Therefore, using the projection operators $\pi_i : X \to S^2$ where $\pi_i(x) = x_i$, we have that our maximum and minimum values are given by

$$\max_{x\in X}\left(E_{q,f}\right) = \frac{1}{2}\left[1 - \min_{x\in X}\left(\sum_{i=1}^{n} q_i\langle r\pi_i(x), \lambda r\pi_i(x)\rangle\right)\right] \tag{6}$$

and

$$\min_{x\in X}\left(E_{q,f}\right) = \frac{1}{2}\left[1 - \max_{x\in X}\left(\sum_{i=1}^{n} q_i\langle r\pi_i(x), \lambda r\pi_i(x)\rangle\right)\right]. \tag{7}$$

Note, since the inner product and every rotation r and projection operator π_i is a continuous function, we then have that the convex sum of inner products in Eq. (7) is also continuous. It then follows that the extrema of this sum are assumed by points in the compact set X and we may write

$$\max_{x\in X}\left(E_{q,f}\right) = \frac{1}{2}\left[1 - \min_{x\in X}\left(\sum_{i=1}^{n} q_i\langle rx_i, \lambda rx_i\rangle\right)\right] \tag{8}$$

and

$$\min_{x\in X}\left(E_{q,f}\right) = \frac{1}{2}\left[1 - \max_{x\in X}\left(\sum_{i=1}^{n} q_i\langle rx_i, \lambda rx_i\rangle\right)\right] \tag{9}$$

Lastly, noting that $r\pi_i(x) = rx_i \in S^2$, we have that

$$\begin{aligned}\sum_{i=1}^{n} q_i\langle rx_i, \lambda rx_i\rangle &= \sum_{i=1}^{n} q_i \sum_{j=1}^{3}(rx_i)_j^2\lambda_j \\ &= \sum_{j=1}^{3}\lambda_j\left(\sum_{i=1}^{n} q_i(rx_i)_j^2\right) \\ &= \sum_{j=1}^{3} p_j(x)\lambda_j,\end{aligned} \tag{10}$$

and furthermore,

$$\sum_{j=1}^{3} p_j(x) = \sum_{i=1}^{n} q_i \sum_{j=1}^{3}(rx_i)_j^2 = \sum_{i=1}^{n} q_i = 1. \tag{11}$$

Consequently, since $a = \max\limits_{x \in X} \left(p_j \right)$ for each j, it follows from Lemma 2.3.5 that

$$\max_{x \in X} \left(E_{q,f} \right) = \frac{1}{2} \Bigg[1 - c\lambda_1 - b\lambda_2 - a\lambda_3 \Bigg] \tag{12}$$

and

$$\min_{x \in X} \left(E_{q,f} \right) = \frac{1}{2} \Bigg[1 - a\lambda_1 - b\lambda_2 - c\lambda_3 \Bigg]. \tag{13}$$

□

Lemma 2: Let $S = \left\{ p \in [0,a] \times [0,b] \times [0,c] \;\middle|\; \sum\limits_{i=1}^{3} p_i = 1 \right\}$ be a non-empty subset of Δ^3. Then for real numbers $\lambda_3 \leq \lambda_2 \leq \lambda_1$ and the continuous function $\phi : S \to \mathbb{R}$ where

$$\phi(p) = \sum_{i=1}^{3} p_i \lambda_i, \tag{14}$$

the extrema of ϕ are given by:

$$\max_{p \in S} \left(\phi \right) = a\lambda_1 + \beta_+ \lambda_2 + \gamma\lambda_3 \quad \& \quad \min_{p \in S} \left(\phi \right) = \alpha\lambda_1 + \beta_- \lambda_2 + c\lambda_3, \tag{15}$$

where

$$\beta_+ = \begin{cases} 1 - a, & \text{if } b > 1 - a \\ b, & \text{otherwise,} \end{cases} \tag{16}$$

$$\beta_- = \begin{cases} 1 - c, & \text{if } b > 1 - c \\ b, & \text{otherwise,} \end{cases} \tag{17}$$

$\alpha = 1 - \beta_- - c$, and $\gamma = 1 - a - \beta_+$.

Proof: We begin by noting that the function ϕ is continuous. Furthermore, since S is the intersection of Δ^3 and $[0,a] \times [0,b] \times [0,c]$, it is a closed, bounded subset of $\mathbb{R}^3$. Therefore, S is compact and ϕ assumes its extrema. Before proving the values of our extrema, we first show that both (a, β_+, γ) and (α, β_-, c) are contained in S. Starting with (a, β_+, γ), we have two cases.

(i) $\underline{b > 1 - a}$:

Then by Eq. (16) $\beta_+ = 1 - a$ and $\gamma = 0$. It then follows that $a + \beta_+ + \gamma = 1$ where $a \in [0,a]$, $\beta_+ \in [0,b]$, and $\gamma \in [0,c]$. Therefore $(a, \beta_+, \gamma) \in S$.

(ii) $\underline{b \leq 1 - a}$:

Then $\beta_+ = b$ and $\gamma = 1 - a - b$. Therefore, we have that $a + \beta_+ + \gamma = 1$ where $a \in [0,a]$, and $\beta_+ \in [0,b]$. Then all that is left to show is that $\gamma \in [0,c]$. Noting that by assumption $\beta_+ \leq 1 - a$, we then have that $\gamma = 1 - a - \beta_+ \geq \beta_+ - \beta_+ = 0$. Furthermore, if $\gamma > c$, then we would have that $1 = a + \beta_+ + \gamma > a + b + c$, and therefore it follows that $S = \emptyset$. This is of course a contradiction and we are left to conclude that $\gamma \in [0,c]$, and thus $(a, \beta_+, \gamma) \in S$.

Similarly for (α, β_-, c), we have the following two cases:

(i) $\underline{b > 1 - c}$:

Then by Eq. (17) $\beta_- = 1 - c$ and $\alpha = 0$. This implies that $c + \beta_- + \alpha = 1$ where $\alpha \in [0,a]$, $\beta_- \in [0,b]$, and $c \in [0,c]$. Therefore $(\alpha, \beta_-, c) \in S$.

(ii) $\underline{b \leq 1 - c}$:

Then $\beta_- = b$ and $\alpha = 1 - b - c$. It then follows that $\alpha + \beta_- + c = 1$ where $\beta_- \in [0,b]$, and $c \in [0,c]$. All that is left to show is that $\alpha \in [0,a]$. Noting that by assumption $\beta_- \leq 1 - c$, we then have that $\alpha = 1 - \beta_- - c \geq \beta_- - \beta_- = 0$. Lastly, if $\alpha > a$, then we would have that $1 = \alpha + \beta_- + c > a + b + c$, and therefore $S = \emptyset$. In conclusion we have that $(\alpha, \beta_-, c) \in S$.

Next we verify the values of our extrema by contradiction. Assuming there exists a $p \in S$ such that

$$\phi(p) = p_1\lambda_1 + p_2\lambda_2 + p_3\lambda_3 > a\lambda_1 + \beta_+\lambda_2 + \gamma\lambda_3. \tag{18}$$

Then since the sum of our p_i's is 1, it follows that

$$p_2(\lambda_2 - \lambda_3) + \beta_+(\lambda_3 - \lambda_2) > p_1(\lambda_3 - \lambda_1) + a(\lambda_1 - \lambda_3), \tag{19}$$

and therefore,

$$(p_2 - \beta_+)(\lambda_2 - \lambda_3) > (a - p_1)(\lambda_1 - \lambda_3) \geq 0. \tag{20}$$

Upon inspection, we see that both $p_2 > \beta_+$ and $\lambda_2 > \lambda_3$. Furthermore, since $\lambda_1 - \lambda_3 \geq \lambda_2 - \lambda_3$, we have that

$$(p_2 - \beta_+)(\lambda_1 - \lambda_3) > (a - p_1)(\lambda_1 - \lambda_3). \tag{21}$$

Assuming that $\lambda_1 - \lambda_3$ is positive (otherwise we have the trivial case where $\phi(p) = \lambda_1$ for all p in S), we finally arrive at

$$p_1 + p_2 > a + \beta_+. \tag{22}$$

Once again we are left with two possible cases:

(i) $\underline{b \leq 1 - a}$: $\beta_+ = b$, and $p_1 + p_2 > a + b$. Therefore, $p \notin S$.

(ii) $\underline{b > 1 - a}$: $\beta_+ = 1 - a$, then $p_1 + p_2 > 1$. Which again implies, $p \notin S$.

Therefore, we are left to conclude that

$$\max_{p \in S}(\phi) = a\lambda_1 + \beta_+\lambda_2 + \gamma\lambda_3. \tag{23}$$

The minimum value is proven by similar arguments. Assuming there exists a $p \in S$ such that

$$\phi(p) = p_1\lambda_1 + p_2\lambda_2 + p_3\lambda_3 < \alpha\lambda_1 + \beta_-\lambda_2 + c\lambda_3, \tag{24}$$

and again noting that the sum of the p_i's is 1, we have that

$$p_2(\lambda_2 - \lambda_1) + \beta_-(\lambda_1 - \lambda_2) < p_3(\lambda_1 - \lambda_3) + c(\lambda_3 - \lambda_1), \tag{25}$$

and therefore,

$$(\beta_- - p_2)(\lambda_1 - \lambda_2) < (p_3 - c)(\lambda_1 - \lambda_3) \leq 0. \tag{26}$$

It then follows that we must have that $p_2 \geq \beta_-$ and $\lambda_1 > \lambda_2$. Furthermore, since $(\lambda_1 - \lambda_3) \geq (\lambda_2 - \lambda_3)$,

$$(\beta_- - p_2)(\lambda_1 - \lambda_3) < (p_3 - c)(\lambda_1 - \lambda_3). \tag{27}$$

Again, assuming that $\lambda_1 - \lambda_3$ is positive, we finally arrive at

$$p_2 + p_3 > \beta_- + c. \tag{28}$$

Lastly, we have the following two cases:

(i) $\underline{b \leq 1 - c}$: $\beta_- = b$, and $p_2 + p_3 > b + c$. Therefore, $p \notin S$.

(ii) $\underline{b > 1 - c}$: $\beta_- = 1 - c$, then $p_2 + p_3 > 1$. Which again implies, $p \notin S$.

Consequently,

$$\max_{p \in S}(\phi) = a\lambda_1 + \beta_+\lambda_2 + \gamma\lambda_3, \tag{29}$$

and the proof is complete.

□

Theorem 3: Let f be a symmetric unital channel with eigenvalues $\lambda_3 \leq \lambda_2 \leq \lambda_1$ that is diagonalized by $r \in SO(3)$. If $E_{q,f}$ is an error function such that for the functions $p_j : X \to \mathbb{R}$ where $p_j(x) = \sum_{i=1}^{n} q_i (r x_i)_j^2$ we have that $a = \max_{x \in X}(p_1),\ b = \max_{x \in X}(p_2),$ and $c = \max_{x \in X}(p_3)$, then

$$\mathrm{Im}\big(E_{q,f}\big) = \left[\tfrac{1}{2}\big(1 - a\lambda_1 - \beta_+\lambda_2 - \gamma\lambda_3\big), \tfrac{1}{2}\big(1 - \alpha\lambda_1 - \beta_-\lambda_2 - c\lambda_3\big)\right], \tag{30}$$

where $a,\ c,\ \beta_+,\ \beta_-,\ \alpha$, and γ are as defined in Lemma 1.

Proof: Since f is symmetric and diagonalized by the rotation r, it follows that

$$E_{q,f}(x) = \tfrac{1}{2}\left[1 - \sum_{i=1}^{n} q_i \langle r x_i, \lambda r x_i \rangle\right] \tag{31}$$

where

$$\lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}. \tag{32}$$

Therefore, the extrema of $E_{q,f}$ are given by:

$$\min_{x \in X}\big(E_{q,f}\big) = \tfrac{1}{2}\Big[1 - \max_{x \in X}\big(\sum_{i=1}^{n} q_i \langle r\pi_i(x), \lambda r \pi_i(x)\big)\Big], \tag{33}$$

and

$$\max_{x \in X}\big(E_{q,f}\big) = \tfrac{1}{2}\Big[1 - \min_{x \in X}\big(\sum_{i=1}^{n} q_i \langle r\pi_i(x), \lambda r \pi_i(x)\big)\Big], \tag{34}$$

where $\pi_i : X \to S^2$ is given by $\pi_i(x) = x_i$. Then since the composition of each projection operator with a rotation $r\pi_i(x) = r x_i$ is continuous we still have that the extrema are assumed by points in

the compact set X. Noting that $r\pi_i(x) = rx_i \in S^2$, then

$$\begin{aligned}\sum_{i=1}^{n} q_i \langle rx_i, \lambda rx_i \rangle &= \sum_{i=1}^{n} q_i \sum_{j=1}^{3} (rx_i)_j^2 \lambda_j \\ &= \sum_{j=1}^{3} \lambda_j \left(\sum_{i=1}^{n} q_i (rx_i)_j^2 \right) \\ &= \sum_{j=1}^{3} p_j \lambda_j,\end{aligned} \tag{35}$$

where we have defined $p_j = \sum_{i=1}^{n} q_i (rx_i)_j^2$ and

$$\sum_{j=1}^{3} p_j (rx_i) = \sum_{i=1}^{n} q_i \sum_{j=1}^{3} (rx_i)_j^2 = \sum_{i=1}^{n} q_i = 1, \tag{36}$$

then since $p_1 \in [0, a]$, $p_1 \in [0, b]$, and $p_1 \in [0, c]$, we may apply Lemma 2. That is, we have that

$$\max_{x \in X} \left(E_{q,f} \right) = \frac{1}{2} \left[1 - \alpha\lambda_1 - \beta_- \lambda_2 - c\lambda_3 \right] \tag{37}$$

and

$$\min_{x \in X} \left(E_{q,f} \right) = \frac{1}{2} \left[1 - a\lambda_1 - \beta_+ \lambda_2 - \gamma\lambda_3 \right], \tag{38}$$

and the proof is complete.

□

LIST OF REFERENCES

[1] Bennett C.H., Brassard G., Breidbart S., Wiesner S., *Quantum Cryptography, or Unforgeable Subway Tokens*. In: Chaum D., Rivest R.L., Sherman A.T. (eds) Advances in Cryptology. Springer, Boston, MA, 1983.

[2] C. H. Bennett, G. Brassard, *Quantum cryptography: Public key distribution and coin tossing*, Proceedings of IEEE International Conference on Computers Systems and Signal Processing Bangalore India, pp. 175-179, 1984.

[3] Bonato C., Aspelmeyer M., and Jennewein T., *Feasibility of Satellite Quantum Key Distribution*. *Opt. Express*, **14**, 10050, 2006.

[4] Bonior, D., *Relativistic Effects on Orbiting Spin Entangled Electrons*. BSc Thesis Middle Tennessee State University, 2013.

[5] Bonior, D., *Limits of Adaptive Quantum Processing*. Proceedings of 2018 IEEE Conference on Antenna Measurements & Applications (CAMA). (In Press)

[6] Bonior, D. and Martin K., *The Domain of Unital Channels*. Manuscript in preparation.

[7] Caves C. M., "Symmetric informationally complete POVMs." http://info.phys.unm.edu/ caves/reports/infopovm.pdf, 2002.

[8] Choi M.D., *Completely positive linear maps on complex matrices*, Linear Algebra Appl. **10** (1975), 285-290.

[9] Chuang I. L. and Nielsen M. A., *Quantum Computation and Quantum Information*. Cambridge(UK): Cambridge University Press, 2008.

[10] Cover T.M. and Thomas J.A., *Elements of information theory*, 2nd Ed., Wiley Hoboken, NJ, 2006.

[11] Crowder T., *Representations of Quantum Channels*. Ph.D. Thesis Howard University, 2013.

[12] Feng, J., *Domain Theoretic Structures in Quantum Information Theory*. Ph.D. Thesis Tulane University, 2011.

[13] Hobson M.P., Efstathious G., and Lasenby A. N., *General Relativity*. (MA): Addison-Welsey, 2006.

[14] Hughes R. J., Morgan G. L., and Nordholt J. E., *Background Noise of Satellite-to-Ground Quantum Key Distribution*, *Proc. SPIE*, **4635**, 116, 2002.

[15] Jones G. A., Jones J. M., *Information and Coding Theory*. University of Southampton (UK): Springer Publish. Pages (42-54), 2000.

[16] Lanzagorta, M., *Quantum Information in Gravitational Fields*. San Rafael (CA): Morgan & Claypool Publishers, Pages (3.8-3.13), (4.1-4.3), (5.29-5.31) and (5.42-5.44), 2013.

[17] Martin K. *A Foundation for Computation*. PhD thesis, Department of Mathematics, Tulane University, 2000.

[18] Martin K. and Panangaden P., *A Domain of Spacetime Intervals in General Relativity*. Springer-Verlag 267:563, 2006.

[19] Martin K., Moskowitz I.S., *Noisy Timing Channels with Binary Inputs and Outputs*. In: Camenisch J.L., Collberg C.S., Johnson N.F., Sallee P. (eds) Information Hiding. IH 2006. Lecture Notes in Computer Science, vol 4437. Springer, Berlin, Heidelberg (2007)

[20] Martin, K. and Panangaden, P., *A Technique for Verifying Measurements*. Electronic Notes in Theoretical Computer Science, Vol. 218, pages 261-273, 2008.

[21] Martin K., *The scope of a quantum channel*. Mathematical Structures in Computer Science, Cambridge(UK): Cambridge University Press, 2008.

[22] Martin, K., *Topology in information theory in topology*. Theoretical Computer Science, Vol. 405, Issues 1-2, pages 75-87, 2008.

[23] Martin, K., *Domain Theory and Measurement*. In: Coecke B. (eds) New Structures for Physics. Lecture Notes in Physics, vol 813. Springer, Berlin, Heidelberg, 2010.

[24] Martin, K., Feng, J., Krishnan, S., *A Free Object in Quantum Information Theory*. Electronic Notes in Theoretical Computer Science, Volume 265, 6 September 2010, Pages 35-47.

[25] Martin, K. and Feng, J., *A new fixed point theorem in Domain theory*. Nat Comput, Springer Netherlands, https://doi.org/10/1007/s11047-018-9672-3, 2018.

[26] Matsushima, Y., *Differentiable Manifolds* Marcel Dekker, Inc., 1972

[27] Munkres, J. R., *Topology, 2nd Edition* Prentice-Hall, Inc., 2000.

[28] Noether E. *Invariante Variationsprobleme*. Nachrichten von der Gesellschaft der Wissenschaften zu Gttingen, Mathematisch-Physikalische Klasse, pp. 235-257, 1918.

[29] Ohnuki Y., *Unitary Representations of the Poincare Group and Relativistic Wave Equations*. (Singapore: World Scientific), 1988.

[30] Pechukas P., *Reduced dynamics need not be completely positive*, Phys. Rev. Lett. **73** (8) (1994), 1060-1062.

[31] Scott D., *Outline of a Mathematical Theory of Computation*. Technical Monograph PRG-2, November 1970.

[32] Shannon C. E., *A mathematical theory of communication*. Bell Systems Technical Journal 27, 379-423 and 623-656, 1948.

[33] Terashima H. and Ueda M., *Einstein-Rosen Correlation in Gravitational Field*. *Phys. Rev.* A **69** 032113, 2004.

[34] Villoresi P., Jennewein T., Tamburini F., Aspelmeyer M., Bonato C., Ursin R., Pernechele C., Luceri V., Bianco G., Zeilinger A., and Barbieri C., *Experimental Verification of the Feasibility of a Quantum Channel Between Space and Earth*. New J. Phys. **10** 033038, 2008.

[35] Weinberg S., *The Quantum Theory of Fields*. Vol 1, Cambridge(UK): Cambridge University Press, 1995.

[36] Whiteney H., *Differentiable manifolds*. Ann. of Math. 37, 645-680, 1936.

[37] Wiesner S., *Conjugate Coding*. written circa 1970 and belatedly published in *Sigact News* 15(1), pp. 78-88, 1983.

[38] Zhang H. *Dualities of Domains*. PhD thesis, Department of Mathematics, Tulane University, 1993.