Matrix Analysis and Operator Theory with Applications to Quantum Information Theory

by

Sarah Plosker

A Thesis
presented to
The University of Guelph

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

Guelph, Ontario, Canada

ABSTRACT

OPERATOR THEORY AND MATRIX ANALYSIS

APPLICATIONS TO

QUANTUM INFORMATION THEORY

Sarah Plosker

University of Guelph, 2013

Advisor:

Professor D. Kribs

We explore the connection between quantum error correction and quantum cryptography through the notion of conjugate (or complementary) channels. This connection is at the level of subspaces and operator subsystems; if we use a more general form of subsystem, the link between the two topics breaks down. We explore both the subspace and subsystem settings.

Error correction arises as a means of addressing the issue of the introduction of noise to a message being sent from one party to another. Noise also plays a role in quantum measurement theory: If one wishes to measure a system that is in a

particular state via a measurement apparatus, one can first act upon the system by a quantum channel, which can be thought of as a noise source, and then measure the resulting system using a different measurement apparatus. Such a setup amounts to the introduction of noise to the measurement process, yet has the advantage of preserving the measurement statistics. Preprocessing by a quantum channel leads to the partial order "cleaner than" on quantum probability measures. Other meaningful partial orders on quantum probability measures exist, and we shall investigate that of cleanness as well as that of absolute continuity.

Lastly, we investigate partial orders on vectors corresponding to quantum states; such partial orders, namely majorization and trumping, have been linked to entanglement theory. We characterize trumping first by means of yet another partial order, power majorization, which gives rise to a family of examples. We then characterize trumping through the complete monotonicity of certain Dirichlet polynomials corresponding to the states in question. This not only generalizes a recent characterization of trumping, but the use of such mathematical objects simplifies the derivation of the result.

This thesis is dedicated to my wonderful boyfriend, Tom Gustin. Thank you for your love, devotion, and understanding through my years of academia.

Acknowledgements

The author wishes to express her sincere gratitude to her advisor, Dr. David W. Kribs, for his guidance and support, Dr. Rajesh Pereira, who was always available for mathematical discussions, and Dr. Doug Farenick, a collaborator, informal mentor, and friend.

The author would also like to acknowledge the financial assistance received from the Natural Sciences and Engineering Research Council of Canada (NSERC), the department of Mathematics and Statistics at the University of Guelph, and Dr. David W. Kribs.

Table of Contents

List of Figures

Chapter 1

Introduction

Information theory was firmly established mathematically by Claude Shannon in his seminal article published in 1948 [Sha48]. His work focused on mathematical descriptions of communication systems, where an information source sends a message through a transmitter, which then sends a signal to a receiver via a noise source, and the receiver obtains an output message.

Quantum information theory (QIT) is the quantum analogue of what is now called classical information theory. The study concerns the use of quantum properties to store, transmit, and process information in an efficient, accurate, and secure way. QIT is a fascinating area of science where mathematical theory can be used to understand and characterize the subtle details that physicists observe in the lab. The basic objects of study are quantum states (unit vectors in $\mathbb{C}^{2^n}$), density matrices (positive

semi-definite, trace-one complex matrices), and quantum channels (completely positive, trace-preserving, linear maps).

Two vital questions in the development of the theory of quantum computing stem from quantum communication system descriptions analogous to the classical communication systems described by Shannon; namely:

1. How can we protect our quantum information against sources of quantum noise?

   Quantum error correction is used to recover information from errors introduced by noise that arises when sending quantum information through a quantum channel.

2. How can we protect information stored in a quantum state from an external observer?

   Quantum cryptography is used to hide information when sending quantum information through a quantum channel so that the original message cannot be recovered by a third party.

In any communication system, communicating parties wish to protect their messages from disturbances arising from noise, which lead to unwanted perturbations in their messages, as well as protect their messages from potentially malicious eavesdroppers. Thus, being able to answer these two questions is of utmost importance. In [KKS08], the authors establish a connection between these two concepts. We discuss quantum

error correction, quantum cryptography, and the intimate connection between the two subjects in chapter 2.

Quantum mechanics consists of four key components:

1. the physical systems to be studied;
2. the apparatuses by which properties of a physical system are measured (which includes all preparation and registration devices);
3. the ambient environment; and
4. the observers.

All four components can be described mathematically, as we shall see.

Unlike in the classical setting, a measurement of a quantum system, prepared in a certain state, may alter the state of the system. Furthermore, experimentation involves repeated measurements (using a fixed apparatus) and the measurement events can only be proclaimed with certain levels of probability, depending on the state in which the system was prepared (and the choice of apparatus).

In classical probability theory, one considers the space of probability measures, which correspond physically to the measurement apparatuses of the system, and the natural partial order of absolute continuity. Quantum mechanics requires a more general notion of probability to account for nonclassical randomness. Quantum measurements are represented by positive-operator valued measures (POVMs) taking values

in $\mathfrak{B}(\mathcal{H})$. That is, given a quantum system $\mathcal{H}$, a POVM is the mathematical formalism of a measuring apparatus of the quantum system. One can still define absolute continuity in this more general setting. The Radon-Nikodým theorem is used in classical probability theory to characterize absolute continuity, and one naturally would like to find a quantum analogue; that is, a Radon-Nikodým theorem for POVMs; this is done in chapter 3.

In the more general setting of quantum probability theory, we can partially order POVMs by partial orders other than that of absolute continuity. In [Hei05], the author describes a number of partial orderings of POVMs which characterize the idea of a POVM being "less than" another. In [BKD+05], the authors define "cleanness" as a partial order on the space of POVMs; that is, we can partially order POVMs by how clean they are. This is done by introducing a preprocessing step to a given quantum measurement, whereby the state of a quantum system is first manipulated by a quantum channel and is then measured by a different apparatus, which amounts to the introduction of noise to the measurement process. We investigate the mathematical consequences of this partial order in chapter 3.

Entanglement is the crux of quantum information theory; being able to manipulate entanglement is what makes quantum information such a powerful tool. A major area of research in quantum information theory is the problem of entanglement transformations: that is, can we manipulate a pure state of a composite system via local operations and classical communication (LOCC) and have it transform into another

particular state? Nielsen [Nie99] answered this question using majorization theory, thus giving majorization an important role in quantum information theory.

Trumping is a partial order on real vectors that generalizes the more familiar concept of majorization, and, in light of Nielsen's result, has recently been used as a tool in entanglement theory. Obtaining efficient criteria characterizing when trumping occurs has been identified as an open problem in quantum information [Wer05, Problem 4]. We add to the partial results on this subject in chapter 4.

1.1 Organization of the Thesis

The bulk of this chapter reviews preliminary definitions and theorems needed for the remainder of the thesis. This includes precise, mathematical definitions of physical concepts such as quantum states, quantum channels, and local operations and classical communication (LOCC), as well as statements of several fundamental theorems that help form a basis for QIT, and a discussion on the duality of the Heisenberg and Schrödinger pictures in quantum mechanics. We begin by introducing notation, which we have done our best to make consistent with the literature.

Chapters 2, 3, and 4 are self-contained and require only the basic knowledge presented in the remainder of this chapter.

Chapter 2 is dedicated to private quantum channels. We first consider the one-qubit setting with respect to the Bloch sphere, and then consider how conditional expectations, a mathematical object that arises in operator theory, can be seen as

private quantum channels through the use of trace vectors, a mathematical object from the field of matrix analysis. These results were published in [CKPP11]. We then delve into to the connection of private quantum channels and quantum error correction, and the ramifications thereof. Notably, we analyze an example of the more subtle notion of *private quantum subsystem*, which leads to the breakdown of the complementarity that exists between quantum error correcting codes and private quantum codes. This example is featured in [JOKLP13]. We set out algebraic conditions that are both necessary and sufficient for the existence of a private quantum subsystem [KP12] (and private quantum subspace, as a corollary [JOKLP13]), and apply these conditions to our example. We show that, for a certain class of channels, no private subspace will exist [JOKLP13]. Finally, we examine the analogue of private quantum subsystem in the quantum error correction setting, which we call a *generalized operator quantum error correcting code*.

Chapter 3 is dedicated to the ordering of positive, operator-valued measures. Motivated by classical probability theory and the space of probability measures, we derive a Radon-Nikodým theorem for POVMs. These results were published in [FPS11]. Next, building on work in [BKD+05], where the authors define "cleanness" as a partial order on the space of POVMs, we investigate the mathematical consequences of this definition. We give an analytic description of what it means for one quantum probability measure to be cleaner than another, and we determine the structure of clean

quantum probability measures. These results are with respect to finite dimensions herein; for an infinite-dimensional approach, see [FFP13].

Finally, chapter 4 is dedicated to trumping. More specifically, we explore connections between trumping and power majorization, a generalization of majorization that has been studied in the theory of inequalities. We prove an analogue of Rado's theorem for power majorization and consider a number of examples. This work is published in [KPP12]. Next, we approach trumping from a different point of view: we characterize trumping via the use of general Dirichlet polynomials, Mellin transforms, and completely monotone functions. This work is published in [PP13]. We believe that our new descriptions of trumping will benefit readers and further the study and understanding of trumping. Moreover, we generalize Turgut's result on trumping [Tur07] to convex sequences of higher order. The proof simplifies considerably through the use of the aforementioned mathematical objects.

1.2 Quantum Information Theory: A Brief Mathematical Introduction

In quantum mechanics, a quantum system is described by a Hilbert space, often taken to be finite-dimensional for simplicity. We will consider only finite-dimensional Hilbert spaces throughout this thesis and denote such spaces by $\mathcal{H}$ and $\mathcal{K}$. One often looks at tensor product quantum systems, for which different users (Alice, Bob,

Charlie, ...) have access to the individual tensored Hilbert spaces. In this setting, we will use subscripts to emphasize when a particular user has access to a particular Hilbert space; that is, if we are considering a tensor product space $\mathcal{H}_A \otimes \mathcal{H}_B$, then we are implying that Alice has access to her system $\mathcal{H}_A$ and Bob has access to his system $\mathcal{H}_B$. We may occasionally use a subscript $d, m, n, \ldots$ to denote the dimension of the space.

Let $\mathfrak{B}(\mathcal{H}, \mathcal{K})$ be the set of all (bounded) linear operators from $\mathcal{H}$ to $\mathcal{K}$, where we write $\mathfrak{B}(\mathcal{H}) \equiv \mathfrak{B}(\mathcal{H}, \mathcal{H})$ as the set of all linear operators acting on $\mathcal{H}$. The cone of all positive $Q \in \mathfrak{B}(\mathcal{H})$ is denoted by $\mathfrak{B}(\mathcal{H})_+$. If $\mathcal{H}$ is d-dimensional, then we can identify $\mathfrak{B}(\mathcal{H})$ with the matrix algebra $\mathbb{M}_d(\mathbb{C})$ of $d \times d$ complex-valued matrices, often simply denoted $\mathbb{M}_d$. Let $\mathfrak{B}(\mathcal{H})_t$ be the set of all trace class operators on $\mathcal{H}$ (trace-class operators are compact operators whose trace can be defined). In finite dimensions the sets $\mathfrak{B}(\mathcal{H})$ and $\mathfrak{B}(\mathcal{H})_t$ coincide and so, unless we wish to draw specific attention to the particular space that we are considering (as in section 1.2.3), we will simply write $\mathcal{L}(\mathcal{H})$ for the set of linear operators on $\mathcal{H}$. The identity element of an operator space $\mathcal{L}(\mathcal{H}_d)$ will be denoted by $\mathrm{I}_{\mathcal{L}(\mathcal{H}_d)}, \mathrm{I}_d$, or simply by I if the space is implied by the context. The identity map acting on $\mathcal{L}(\mathcal{H})$ will be denoted by $\mathrm{id} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$, which is given by $\mathrm{id}(X) = X$ for all $X \in \mathcal{L}(\mathcal{H})$.

We use † to represent the complex conjugate transpose of vectors and matrices, and * to represent the complex conjugate of scalars. We will use Dirac (bra-ket) notation from quantum mechanics: A column vector is called a "ket" and written as $|\psi\rangle \in \mathbb{C}^d$.

Every ket $|\psi\rangle$ has a dual "bra" $\langle\psi| \equiv |\psi\rangle^\dagger$. Explicitly, if $|\psi\rangle = (c_1 \; c_2 \; \ldots \; c_d)^T$ (a column vector), then $\langle\psi| = (c_1{}^*, \; c_2{}^*, \; \ldots, \; c_d{}^*)$ (a row vector). The canonical basis of $\mathbb{C}^d$ will be represented by $\big\{|i\rangle\big\}_{i=1}^d$, although the notation $\{|0\rangle, |1\rangle\}$ is used for $\mathbb{C}^2$. Unless otherwise noted, we will restrict our attention to unit vectors (that is, vectors $|\psi\rangle$ satisfying $\langle\psi|\psi\rangle = 1$). The use of bra-ket notation leads us to the convention that inner products are conjugate linear in the first variable.

A matrix $X \in \mathbb{M}_d$ is said to be *Hermitian* if $X^\dagger = X$ and it is said to be *unitary* if $X^\dagger X = \mathrm{I}$. A Hermitian matrix X is *positive*, denoted $X \geq 0$, if $\langle v|X|v\rangle \geq 0$ for all $|v\rangle \in \mathbb{C}^d$ (that is, if X is positive semi-definite).

1.2.1 Quantum States and Density Matrices

We now introduce some basic definitions from quantum information theory. For a more detailed version of these definitions and basic facts, see [NC00].

In the Schrödinger picture of quantum mechanics, quantum information is contained in quantum states, which can be either *pure* or *mixed*. Mathematically, pure states are described by unit vectors $|\psi\rangle \in \mathbb{C}^d$, which can be identified with the corresponding rank-one projection $|\psi\rangle\langle\psi|$, while mixed states are described by *density matrices*—matrices of the form $\rho \equiv \sum_i p_i |\psi_i\rangle\langle\psi_i|$, where $\{p_i\}$ forms a *probability distribution* (that is, $0 \leq p_i \leq 1$ for all i and $\sum_i p_i = 1$). Density matrices are precisely the trace-one, positive matrices.

We are somewhat loose with our notation: often, we write "state" when we are in fact referring to a density matrix. This is typical of the literature. Perhaps the most important density matrix is that corresponding to the maximally mixed state: the state $\frac{1}{d}\mathrm{I} \in \mathbb{M}_d$ described by the uniform probability distribution over $\{|i\rangle\}_{i=1}^{d}$.

In a tensor product quantum system, a pure state $|\psi\rangle \in \mathbb{C}^m \otimes \mathbb{C}^n$ is called *separable* if it can be written as an elementary tensor: $|\psi\rangle = |\varphi\rangle \otimes |\phi\rangle$ for some $|\varphi\rangle \in \mathbb{C}^m$ and $|\phi\rangle \in \mathbb{C}^n$. Otherwise, $|\psi\rangle$ is said to be *entangled*. In the case of mixed states, we say that $\rho \in \mathbb{M}_m \otimes \mathbb{M}_n$ is separable if it can be written as a convex combination of separable pure states [Wer89]:

$$\rho = \sum_i p_i |\varphi_i\rangle\langle\varphi_i| \otimes |\phi_i\rangle\langle\phi_i|,$$

where $\{p_i\}$ forms a probability distribution. Otherwise, ρ is said to be entangled.

Example 1.2.1. *The* maximally entangled state *in* $\mathcal{H}_m \otimes \mathcal{H}_n$ *is represented by the unit vector* $|\varphi_e\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |e_i\rangle \otimes |f_i\rangle$, *where* $\{|e_i\rangle\}$ *and* $\{|f_i\rangle\}$ *form orthonormal sets in* $\mathcal{H}_m$ *and* $\mathcal{H}_n$ *respectively, and* $d = \min\{m, n\}$.

Entangled states violate the classical principle of locality—the idea that an object is directly influenced only by its immediate surroundings. The easiest way to see evidence of entanglement is to measure one component of an entangled state. This measurement fixes the value of the other component of the state, regardless of the spacial separation of the two states, implying non-local communication between the two parts. In the famous EPR paper [EPR35], nonlocal behaviour, described as *spooky*

action at a distance, was deemed to illustrate how the description of reality given by quantum mechanics must be incomplete. The authors came to this conclusion using classical reasoning and classical intuition, which does not translate to quantum systems. Observation of entanglement via modern day physical experiments in quantum mechanics provides much insight into entanglement, most notably how it can be used as a resource in order to implement quantum teleportation [NC00], a technique for moving quantum states even in the absence of a quantum communication channel.

As an example of entanglement in action, consider electron spin, which, when measured, can be either of two states: spin up $\uparrow$ or spin down $\downarrow$. In the absence of measurement, electron spin is in a *superposition* of the two states $\uparrow$ / $\downarrow$ (even without entanglement). Now, consider an entangled electron pair; suppose the state of one of the electrons is measured by Alice to be "spin up". This measurement fixes the state of the other electron—if Bob uses the same measurement basis to measure the second electron, it will be "spin down". In this sense, even when spatially separated, entangled electron pairs behave as a single quantum object.

Schmidt's decomposition theorem allows us to decompose any pure bipartite state via orthonormal bases in the respective systems.

Theorem 1.2.2. *(Schmidt's decomposition theorem) Let* $d = \min\{\dim \mathcal{H}_A, \dim \mathcal{H}_B\}$. *Let* $|\varphi\rangle$ *be a pure state in the composite system* $\mathcal{H}_A \otimes \mathcal{H}_B$. *Then there exists orthonormal sets of vectors* $\{|i_A\rangle\}_{i=1}^d \subset \mathcal{H}_A$ *and* $\{|i_B\rangle\}_{i=1}^d \subset \mathcal{H}_B$ *such that* $|\varphi\rangle$ *may be written*

in the form

$$|\varphi\rangle = \sum_{i=1}^{d} \sqrt{\lambda_i} |i_A\rangle \otimes |i_B\rangle,$$

where $\{\lambda_i\}_{i=1}^d$ *are non-negative scalars such that* $\sum_{i=1}^d \lambda_i = 1$.

Note: We obtain the identity $\sum_{i=1}^d \lambda_i = 1$ from the fact that $|\varphi\rangle$ is a unit vector.

1.2.2 Completely Positive Maps

In the Schrödinger picture of quantum mechanics, a quantum channel represents evolution of quantum states [NC00]. In other words, a quantum channel is a map that sends density matrices to density matrices. The details of quantum mechanics further impose linearity. Additionally, one requires that, if a channel Φ is coupled with an ancilla system, the induced map $\mathrm{id} \otimes \Phi$ that acts as the identity on the ancilla and as the original channel on the (original) system must still be positive. This leads to a more formal, mathematical definition of quantum channel:

Definition 1.2.3. *We say that a linear map* $\Phi : \mathbb{M}_m \to \mathbb{M}_n$ *is* completely positive *(CP) if the induced mappings* $\Phi_d : \mathbb{M}_d \otimes \mathbb{M}_m \to \mathbb{M}_d \otimes \mathbb{M}_n$, *defined by* $\Phi_d = \mathrm{id}_d \otimes \Phi$, *are positive for all* d.

A linear map $\Phi : \mathbb{M}_m \to \mathbb{M}_n$ is completely positive if and only if there exist operators $\{A_k\}_{k=1}^{mn}$ such that $\Phi(\cdot) = \sum_{k=1}^{mn} A_k(\cdot)A_k^\dagger$ [Cho75, Kra71]. The A_k in this equivalence are typically called the *Choi* or *Kraus operators* of Φ. We will

abuse notation slightly and say $\Phi = \{A_k\}$ to indicate that $\{A_k\}_{k=1}^{mn}$ form a set of Kraus operators for the completely positive map Φ.

Remark 1.2.4. *Kraus operators are not in general unique; however, any two sets of Kraus operators $\{A_i\}_{i=1}^{d_1}$ and $\{B_j\}_{j=1}^{d_2}$, with $d_1 \leq d_2$, correspond to the same completely positive map if and only if there exists an isometry $U = (u_{i,j})_{i,j}$ (unitary when $d_1 = d_2$) such that $A_i = \sum_{j=1}^{d_2} u_{i,j} B_j$ for each $i = 1, \ldots, d_1$. [NC00, Theorem 8.2]. Thus we often say "the" Kraus operators of a completely positive map with this unitary freedom in mind.*

A map is called trace preserving if $\mathrm{Tr}(\Phi(X)) = \mathrm{Tr}(X)$ for all X. Unless stated otherwise, all maps considered herein are assumed linear. For more details on completely positive maps from a matrix-theoretic point of view, we refer the reader to [Bha07].

Observe that Φ is trace-preserving if and only if

$$\mathrm{Tr}(X) = \mathrm{Tr}\big(\Phi(X)\big) = \mathrm{Tr}\Big(\sum_{k=1}^{mn} A_k X A_k^\dagger\Big) = \mathrm{Tr}\Big(X \sum_{k=1}^{mn} A_k^\dagger A_k\Big) \quad \forall\, X \in \mathbb{M}_m.$$

In other words, Φ is trace-preserving if and only if $\sum_{k=1}^{mn} A_k^\dagger A_k = \mathrm{I}$. We further note that the condition $\sum_{k=1}^{mn} A_k A_k^\dagger = \mathrm{I}$ corresponds to Φ being *unital* (i.e., $\Phi(\mathrm{I}_m) = \mathrm{I}_n$).

Definition 1.2.5. *A* (quantum) channel Φ *is a linear, CP, trace preserving (CPTP) map.*

Example 1.2.6. *Perhaps the simplest non-trivial example of a quantum channel $\Phi : \mathbb{M}_m \to \mathbb{M}_n$ is that which has a single (unitary) Kraus operator $U \in \mathbb{M}_{n,m}$. Such*

a channel sends any density matrix ρ to $U\rho U^\dagger$. Trace-preservation forces U to be unitary. As a natural generalization of this example, a channel Φ is called a random unitary channel *if it admits a decomposition*

$$\Phi(\rho) = \sum_i p_i U_i \rho U_i^\dagger \qquad \forall \rho,$$

where p_i form a probability distribution and U_i are unitary operators.

The following three matrices are often used as Kraus operators of unitary quantum channels as they describe physically important operations; they are known as the *Pauli matrices*:

$$X \equiv \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; Y \equiv \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \; Z \equiv \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Note that X, Y, Z are Hermitian, unitary operators, and that the set $\{\mathrm{I}, X, Y, Z\}$ forms a basis for the real vector space of Hermitian matrices in $\mathbb{M}_2$. In physics, X represents bit flip, in the sense that $X|0\rangle = |1\rangle$ and $X|1\rangle = |0\rangle$; Z represents a phase flip, as $Z|0\rangle = |0\rangle$ and $Z|1\rangle = -|1\rangle$ — that is, $|1\rangle$ is rotated by $\frac{\pi}{2}$. Finally, Y can be seen as a combination of both bit and phase flip: $iY = XZ$.

Example 1.2.7. *A ubiquitous class of examples of quantum channels is the class of depolarizing channels: a map $\Phi : \mathbb{M}_n \to \mathbb{M}_n$ is called a* depolarizing channel *if, for some $0 < p \leq 1$,*

$$\Phi(X) = \frac{p}{n} \operatorname{Tr}(X)\, \mathrm{I} + (1-p)X \qquad \forall X \in \mathbb{M}_n.$$

We obtain the completely depolarizing channel *when $p = 1$.*

The evolution of an isolated system $\mathcal{H}$ from one state to another can be described mathematically by a unitary operator, which information theorists call a *quantum gate*. Since unitary operators are invertible, the dynamics of an isolated system are completely reversible in the sense that we can simply undo the action of the gate by applying its inverse. An important class of gates in quantum mechanics are unitary matrices known as *control* gates, for reasons we will see momentarily. We will define 2-qubit control gates, as that is all that is necessary for our purposes.

Let $U \in \mathbb{M}_2$ be a unitary matrix. Then

$$CU_{12} \equiv |0\rangle\langle 0| \otimes \mathrm{I}_2 + |1\rangle\langle 1| \otimes U,$$

which acts on two qubits, and mathematically can be described by $|\psi_1\rangle|\psi_2\rangle \mapsto CU_{12}|\psi_1\rangle|\psi_2\rangle$ for any $|\psi_1\rangle|\psi_2\rangle \in \mathbb{C}^4$. Physically, we can describe this action as follows: CU_{12} acts as the identity on the first qubit, and, if the first qubit is $|0\rangle$, acts as the identity on the second qubit; however, if the first qubit is $|1\rangle$, then the second qubit is acted on by U. In this way, the first qubit is said to be the *control*, as its value determines the action of the gate on the second qubit. Similarly, we define

$$CU_{21} \equiv \mathrm{I}_2 \otimes |0\rangle\langle 0| + U \otimes |1\rangle\langle 1|,$$

which mathematically can be described by $|\psi_1\rangle|\psi_2\rangle \mapsto CU_{21}|\psi_1\rangle|\psi_2\rangle$ for any $|\psi_1\rangle|\psi_2\rangle \in \mathbb{C}^4$. Physically, we can describe this action as: CU_{21} acts as the identity on the *second* qubit, and, if the second qubit is $|0\rangle$, acts as the identity on the first qubit; however, if the second qubit is $|1\rangle$, then the first qubit is acted on by U. In this way, the second

qubit is said to be the control, as its value determines the action of the gate on the first qubit.

1.2.3 Stinespring Dilation Theorem

Let us first define the partial trace, which is a generalization of the trace function. The partial trace is an operator-valued function acting on a tensor product of Hilbert spaces.

Definition 1.2.8. *The map* $\mathrm{Tr}_{\mathcal{K}} : \mathcal{L}(\mathcal{H}) \otimes \mathcal{L}(\mathcal{K}) \to \mathcal{L}(\mathcal{H})$ *defined by*

$$\mathrm{Tr}_{\mathcal{K}} = \mathrm{id}_{\mathcal{L}(\mathcal{H})} \otimes \mathrm{Tr}$$

is called the partial trace *(with respect to* $\mathcal{L}(\mathcal{K})$*) and we say that we "trace out" the system* $\mathcal{L}(\mathcal{K})$*.*

We could similarly define the partial trace with respect to $\mathcal{L}(\mathcal{H})$. The partial trace is in fact a linear, trace preserving, completely positive map, and as such it is a quantum channel.

We have been discussing the Schrödinger picture of quantum mechanics, where quantum channels are described by completely positive, trace preserving linear maps acting on trace-class operators. There exists an equivalent view, called the Heisenberg picture, in which mathematicians typically work. In this picture, one works with unital, completely positive maps acting on $\mathfrak{B}(\mathcal{H})$. One can pass freely from one picture to the other by viewing the set $\mathbb{M}_n$ (to which both $\mathfrak{B}(\mathcal{H}_n)$ and $\mathfrak{B}(\mathcal{H}_n)_t$ are

isomorphic as finite-dimensional sets) as a Hilbert space when endowed with the *Hilbert–Schmidt inner product* defined by $\langle X|Y\rangle \equiv \mathrm{Tr}(X^\dagger Y)$ for all $X, Y \in \mathbb{M}_n$. With this inner product in mind, we can define the *dual* map $\Phi^\dagger$ of any linear map Φ:

Definition 1.2.9. *Let $\Phi : \mathbb{M}_m \to \mathbb{M}_n$ be a linear map. The unique map $\Phi^\dagger : \mathbb{M}_n \to \mathbb{M}_m$ (note the reversal of subscripts) satisfying $\mathrm{Tr}\big((\Phi(X))^\dagger Y\big) = \mathrm{Tr}\big(X^\dagger \Phi^\dagger(Y)\big)$ for all $X \in \mathbb{M}_m$, $Y \in \mathbb{M}_n$ is called the* dual map of Φ.

In particular, given a quantum channel $\Phi : \mathfrak{B}(\mathcal{H}_m)_t \to \mathfrak{B}(\mathcal{H}_n)_t$ defined by its Choi-Kraus representation $\Phi(\rho) = \sum_i V_i(\rho)V_i^\dagger$, we have

$$\mathrm{Tr}\left(\Phi(X)^\dagger \rho\right) = \mathrm{Tr}\left(\left(\sum_i V_i(X)V_i^\dagger\right)^\dagger \rho\right) = \mathrm{Tr}\left(X^\dagger\left(\sum_i V_i^\dagger(\rho)V_i\right)\right) = \mathrm{Tr}\left(X^\dagger \Phi^\dagger(\rho)\right).$$

It follows that the dual map $\Phi^\dagger : \mathfrak{B}(\mathcal{H}_n) \to \mathfrak{B}(\mathcal{H}_m)$ has the Choi–Kraus representation $\Phi^\dagger(X) = \sum_i V_i^\dagger X V_i$.

The properties of being trace-preserving and being unital are dual in the sense that a map Φ is trace-preserving if and only if the dual map $\Phi^\dagger$ is unital.

We begin our discussion of Stinespring's theorem in the Heisenberg picture:

Theorem 1.2.10. *[Sti55] Suppose that $\Phi^\dagger : \mathfrak{B}(\mathcal{H}_B) \to \mathfrak{B}(\mathcal{H}_A)$ is a completely positive unital linear map. Then there is a Hilbert space $\mathcal{K}$ (of dimension at most $\dim(\mathcal{H}_A)\dim(\mathcal{H}_B)$) and an isometry $V \in \mathfrak{B}(\mathcal{H}_A, \mathcal{H}_B \otimes \mathcal{K})$ such that*

$$\Phi^\dagger(X) = V^\dagger(X \otimes \mathrm{I}_\mathcal{K})V \qquad \forall X.$$

The pair $(V, \mathcal{K})$ is called a Stinespring representation of $\Phi^\dagger$. We typically consider only *minimal* $\mathcal{K}$: a Hilbert space $\mathcal{K}'$ is minimal provided $\mathcal{H}_B \otimes \mathcal{K}' \subseteq \mathcal{H}_B \otimes \mathcal{K}$ is the closed linear span of $(X \otimes \mathrm{I}_{\mathcal{K}})V\mathcal{H}_A$. Note that V is unique up to a unitary on (minimal) $\mathcal{K}$.

Here the Kraus operators for $\Phi^\dagger$ can be read off as the "coordinate operators" of $V^\dagger$: $V^\dagger = (V_1^\dagger \ \cdots \ V_K^\dagger)$ where the Choi/Kraus decomposition of $\Phi^\dagger$ is

$$\Phi^\dagger(X) = \sum_{i=1}^{K} V_i^\dagger X V_i \qquad \forall X, \text{ where } V_i \in \mathfrak{B}(\mathcal{H}_A, \mathcal{H}_B).$$

Now, we pass to the Schrödinger picture:

Theorem 1.2.11. *Suppose that $\Phi : \mathfrak{B}(\mathcal{H}_A)_t \to \mathfrak{B}(\mathcal{H}_B)_t$ is a completely positive trace preserving linear map. Then there is a Hilbert space $\mathcal{K}$ (of dimension at most $\dim(A)\dim(B)$), a partial isometry $U \in \mathfrak{B}(\mathcal{H}_A \otimes \mathcal{K}, \mathcal{H}_B \otimes \mathcal{K})$, and a pure state $|\psi\rangle \in \mathcal{K}$ such that*

$$\Phi(\rho) = \mathrm{Tr}_{\mathcal{K}}(U(\rho \otimes |\psi\rangle\langle\psi|)U^\dagger) \qquad \forall \rho. \tag{1.1}$$

This is the Schrödinger picture for the (discrete) time evolution of quantum states. If we define $V|\phi\rangle := U(|\phi\rangle \otimes |\psi\rangle)$ for all pure states $|\phi\rangle \in \mathcal{H}_A$, for some fixed pure state $|\psi\rangle \in \mathcal{K}$, then we can write equation (1.1) more succinctly as $\Phi(\rho) = \mathrm{Tr}_{\mathcal{K}}(V\rho V^\dagger)$. Note that this V is the same V that arises in the Heisenberg picture, thus connecting the two viewpoints. In this way, the general form for U is $U = (\, V \mid * \,)$.

1.2.4 Purification of Mixed States

Fix a density operator $\rho_0 \in \mathfrak{B}(\mathcal{H})_t$, and consider the CPTP map $\Phi : \mathbb{C} \to \mathfrak{B}(\mathcal{H})_t$ defined by

$$\Phi(c \cdot 1) = c\,\rho_0 \qquad \forall c \in \mathbb{C}.$$

Then, with $\mathcal{K} = \mathbb{C} \otimes \mathcal{H} = \mathcal{H}$ and $U \in \mathfrak{B}(\mathcal{H}, \mathcal{H} \otimes \mathcal{H})_t$, Stinespring's theorem gives

$$\rho_0 = \Phi(1) = \mathrm{Tr}_{\mathcal{K}}(U(1 \otimes |\psi\rangle\langle\psi|)U^\dagger) = \mathrm{Tr}_{\mathcal{K}}(|\psi'\rangle\langle\psi'|),$$

where $|\psi'\rangle \in \mathcal{H} \otimes \mathcal{H}$ is a purification of ρ_0, and the freedom in U (and thus in $|\psi\rangle$) yields all possible purifications $|\psi'\rangle\langle\psi'|$ for ρ_0.

This idea of dilation constructions of channels and states is known as "going to the church of the larger Hilbert space", a phrase credited to John Smolin.

1.2.5 Conjugate/Complementary Channels

Definition 1.2.12. *[Hol07, KMNR07] Given a quantum channel map $\Phi : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$, consider the Stinespring representation given by $V \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B \otimes \mathcal{K})$ and $\mathcal{K}$ for which*

$$\Phi(\rho) = \mathrm{Tr}_{\mathcal{K}}(V \rho V^\dagger).$$

*Then the corresponding **conjugate** (or **complementary**) channel is the CPTP map $\Phi^\sharp : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{K})$ given by*

$$\Phi^\sharp(\rho) = \mathrm{Tr}_{\mathcal{H}_B}(V \rho V^\dagger).$$

Fact: Any two conjugates $\Phi^\sharp$, Φ' obtained in this way are related by a partial isometry W such that $\Phi^\sharp(\cdot) = W\Phi'(\cdot)W^\dagger$. We talk of "the" conjugate channel for Φ with this understanding.

Now that we have defined the conjugate channel, the next natural step is to compute the Kraus operators for the conjugate channel, knowing the Kraus operators of the original channel. This was done in both [Hol07, KMNR07]; we sketch the idea from the latter.

Suppose that $V_i \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ are the Kraus operators for $\Phi : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$. Then we can obtain Kraus operators $\{R_\mu\}$ for $\Phi^\sharp$ as follows.

Fix an orthonormal basis $\{|e_i\rangle\}$ for $\mathcal{K}$ and define for $\rho \in \mathcal{L}(\mathcal{H}_A)$,

$$F(\rho) = \sum_{i,j} |e_i\rangle\langle e_j| \otimes V_i \rho V_j^\dagger \in \mathcal{L}(\mathcal{K} \otimes \mathcal{H}_B).$$

Then $\Phi(\rho) = \mathrm{Tr}_{\mathcal{K}} F(\rho)$ and

$$\Phi^\sharp(\rho) = \mathrm{Tr}_{\mathcal{H}_B} F(\rho) = \sum_{i,j} \mathrm{Tr}(V_i \rho V_j^\dagger)|e_i\rangle\langle e_j| = \sum_\mu R_\mu \rho R_\mu^\dagger,$$

where $R_\mu^\dagger = [V_1^\dagger|f_\mu\rangle\, V_2^\dagger|f_\mu\rangle \cdots]$ and $\{|f_\mu\rangle\}$ is an orthonormal basis for $\mathcal{H}_B$.

Note that there are $\dim \mathcal{K}$ Kraus operators $V_i \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ for the original channel Φ and $\dim \mathcal{H}_B$ Kraus operators $R_\mu \in \mathcal{L}(\mathcal{H}_A, \mathcal{K})$ for the conjugate channel. The number of Kraus operators is determined by the dimension of the environment, which conceptually is the space one "traces out".

The calculations for determining Kraus operators of a complementary channel are best illustrated through example.

Example 1.2.13. *Consider the 2-qubit swap channel* $\Phi : \mathbb{M}_4 \to \mathbb{M}_4$ *given by* $\Phi(\sigma \otimes \rho) = \rho \otimes \sigma$*, which has a single Kraus operator, the swap unitary*

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

The conjugate map $\Phi^\sharp : \mathbb{M}_4 \to \mathbb{C}$ *is implemented with four Kraus operators:*

$$\begin{aligned} R_1 &= \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix} \quad R_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix} \\ R_3 &= \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} \quad R_4 = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Note that Φ *has domain and range* $\mathbb{M}_4$ *and is implemented by* one *Kraus operator, whereas* $\Phi^\sharp$ *has domain* $\mathbb{M}_4$ *and range* $\mathbb{C}$ *(which is* one*-dimensional) and is implemented by* four *Kraus operators.*

Example 1.2.14. *Consider the 2-qubit phase flip channel* $\Phi : \mathbb{M}_4 \to \mathbb{M}_4$ *with (equally weighted) Kraus operators* $\{\mathrm{I}, Z\,\mathrm{I}\}$*, where* $Z\,\mathrm{I}$ *is shorthand for* $Z \otimes \mathrm{I}_2$*. The conjugate channel* $\Phi^\sharp : \mathbb{M}_4 \to \mathbb{M}_2$ *is implemented with the following Kraus operators:*

$$\begin{aligned} R_1 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad R_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \\ R_3 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix} \quad R_4 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \end{aligned}$$

Note that Φ *has domain and range* $\mathbb{M}_4$ *and is implemented by* two *Kraus operators, whereas* $\Phi^\sharp$ *has domain* $\mathbb{M}_4$*, range* $\mathbb{M}_2$*, and is implemented by* four *Kraus operators.*

1.2.6 Local operations and classical communication (LOCC)

If two parties, Alice and Bob, can only carry out operations on their local systems and have a classical communication channel to transmit bits, it is called *local operations and classical communication (LOCC)* [BBPS96]. The most general implementation of LOCC is a potentially unlimited back-and-forth scenario: (1) Alice applies a quantum channel to her system and communicates classical information to Bob, and (2) Bob applies a quantum channel on his system and communicates classical information to Alice. However, theorem 4.1.3 [LP01] states that all LOCC protocols can be simplified into a one-way communication setup (thus, only communication from Alice to Bob is necessary).

A local operation is mathematically described as a trace decreasing CP map,

$$\Phi_A \otimes \Phi_B : \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathcal{L}(\mathcal{H}'_A \otimes \mathcal{H}'_B)$$

that acts separately on each component of the tensor product:

$$\Phi_A \otimes \Phi_B = (\Phi_A \otimes \mathrm{id}_B) \circ (\mathrm{id}_A \otimes \Phi_B).$$

Let $\mathrm{diag}(\mathbb{M}_d) \subset \mathbb{M}_d$ denote the classical algebra of $d \times d$ diagonal matrices for some d. Classical communication is mathematically represented by

$$\Phi_A : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}'_A) \otimes \mathrm{diag}(\mathbb{M}_d)$$

and/or

$$\Phi_B : \mathcal{L}(\mathcal{H}_B) \otimes \mathrm{diag}(\mathbb{M}_d) \to \mathcal{L}(\mathcal{H}'_B)$$

That is, Alice can tensor her output (or Bob can tensor his input) with a diagonal matrix, representing classical information. We can think of Φ_B as $\oplus\Phi_{B_i}$, a direct sum of trace-decreasing, CP maps Φ_{B_i} each acting on one block of $\mathcal{L}(\mathcal{H}_B) \otimes \mathrm{diag}(\mathbb{M}_d)$.

1.2.7 Operator Systems

Our study of ordering POVMs in chapter 3 naturally gives rise to the use of operator systems.

Definition 1.2.15. ([CE77, Pau02]) *An* operator system in $\mathfrak{B}(\mathcal{H})$ *is a linear (not necessarily norm-closed) subspace* $\mathcal{S} \subset \mathfrak{B}(\mathcal{H})$ *with the properties that* $\mathrm{I} \in \mathcal{S}$ *and* $S^\dagger \in \mathcal{S}$ *for every* $S \in \mathcal{S}$.

The operator algebra $\mathfrak{B}(\mathcal{H})$ is an operator system, as is every unital C*-subalgebra $\mathcal{A} \subset \mathfrak{B}(\mathcal{H})$. We can define the notions of *completely positive* and *unital* for linear maps between operator systems acting on Hilbert spaces, in the same fashion as the definitions of completely positive and unital linear maps between the operator algebras of bounded linear maps acting on Hilbert spaces:

Definition 1.2.16. *Assume that* $\mathcal{S} \subset \mathfrak{B}(\mathcal{H})$ *and* $\mathcal{T} \subset \mathfrak{B}(\mathcal{K})$ *are operator systems acting on Hilbert space* $\mathcal{H}$ *and* $\mathcal{K}$. *A linear map* $\phi : \mathcal{S} \to \mathcal{T}$ *is* completely positive *if, for every* $p \in \mathbb{N}$, *the linear map* $\phi^{(p)} : \mathbb{M}_p(\mathcal{S}) \to \mathbb{M}_p(\mathcal{T})$ *in which*

$$\phi^{(p)}\left[S_{ij}\right]_{i,j=1}^{p} \;=\; \left[\phi(S_{ij})\right]_{i,j=1}^{p}$$

has the property of mapping the positive cone $\mathbb{M}_p(\mathcal{S})_+$ *of* $\mathbb{M}_p(\mathcal{S})$ *into the positive cone* $\mathbb{M}_p(\mathcal{T})_+$ *of* $\mathbb{M}_p(\mathcal{T})$*, where a* $p \times p$ *matrix* X *of operators is positive if* X *is positive as an operator acting on the Hilbert space* $\mathcal{H}^{(p)} = \mathcal{H} \oplus \cdots \oplus \mathcal{H}$*. Furthermore, if a completely positive linear map* $\phi : \mathcal{S} \to \mathcal{T}$ *is such that* $\phi(\mathrm{I}) = \mathrm{I}$*, then* ϕ *is* unital *and* ϕ *is called a* ucp map*.*

Completely positive linear maps of $\mathfrak{B}(\mathcal{H})$ or, more generally, of C*-algebras admit a Stinespring decomposition [Pau02, Chapter 4], which is an extremely important tool by which one studies complete positivity. In contrast, there is no Stinespring decomposition for completely positive maps on operator systems that are not C*-algebras, which adds a degree of difficulty in working with such structures.

Definition 1.2.17. ([Pau02]) *Two operator systems* $\mathcal{S}$ *and* $\mathcal{T}$ *are* completely order isomorphic *if there is a linear bijection* $\phi : \mathcal{S} \to \mathcal{T}$ *such that* ϕ *and* ϕ^{-1} *are completely positive. If, in addition,* $\phi(\mathrm{I}) = \mathrm{I}$*, then* $\mathcal{S}$ *and* $\mathcal{T}$ *are said to be* unitally completely order isomorphic*.*

Chapter 2

Private Quantum Channels and Quantum Error Correction: An Operator-Theoretic Approach

2.1 Private Quantum Channels

First introduced in [AMTdW00, BR03], private quantum channels are at the heart of quantum cryptography; they are the quantum analogue to the classical one-time pad. The oft described situation is as follows: Alice wishes to send a message $|\psi\rangle$ to Bob without an eavesdropper, Eve, being able to learn the message. Depending on the situation, Alice can first append an ancilla state ρ_a to her message: $|\psi\rangle \to |\psi\rangle\langle\psi| \otimes \rho_a$. There is a set of unitaries $\{U_i\}_{i=1}^n$ and corresponding probabilities $\{p_i\}$ that are known

publicly, available for Alice to use. Alice and Bob share a key (privately), i_0, and with probability $p_{i_0} \in \{p_i\}$, Alice applies a unitary $U_{i_0} \in \{U_i\}$ to her message: $|\psi\rangle\langle\psi| \otimes \rho_a \to U_{i_0}(|\psi\rangle\langle\psi| \otimes \rho_a)U_{i_0}^\dagger$. Bob receives the output message, and, knowing i_0, can undo Alice's operation, discard the ancilla (i.e. "trace out"), and recover the original message.

This is secure against Eve, provided Alice and Bob can ensure that Eve always "sees" a fixed output ρ_0, regardless of Alice's message $|\psi\rangle$. If Eve measures the state $U_{i_0}(|\psi\rangle\langle\psi| \otimes \rho_a)U_{i_0}^\dagger$, she will disrupt it (Heisenberg Uncertainty Principle), so her best description of the situation is knowing each U_i is used with probability p_i. Thus, Eve's best model of the system, without disrupting it, is

$$\Phi(|\psi\rangle\langle\psi| \otimes \rho_a) = \sum_i p_i U_i(|\psi\rangle\langle\psi| \otimes \rho_a)U_i^\dagger = \rho_0, \tag{2.1}$$

regardless of what state Alice sends. That is, Eve will obtain *no* information about the original message, as the output is independent of the input $|\psi\rangle$. A channel Φ satisfying equation (2.1) is then said to be private.

The completely depolarizing channel ($\Phi(\rho) = \frac{1}{\dim \mathcal{H}} \mathrm{I}$ for all ρ) is perhaps the simplest example of a quantum channel that is private. In this case, an ancilla ρ_a is not used. The entire Hilbert space $\mathcal{H}$ acts as a private code for the channel, and so in order to implement such a private channel for one qubit ($\mathcal{H} \cong \mathbb{C}^2$) the full set of Pauli matrices (including the identity) is used; for more general dimension 2^n, the set of all possible tensor products of the four Pauli matrices is used. However, in some situations it may be the case that this full set is unavailable for use. For this reason,

it is often desirable to consider channels with fewer physical operations such that we can still encode qubits from a smaller set $\mathcal{S} \subset \mathcal{H}$ for privacy.

In the literature, examples of private quantum channels have been limited to channels formed using tensor products of Pauli matrices. Using the machinery of *trace vectors* and *conditional expectations*, we can give new examples.

Based on the Alice & Bob paradigm discussed above, a mathematical definition of private quantum channel can be given.

Definition 2.1.1. *[AMTdW00] Let $\mathcal{S} \subseteq \mathcal{H}_A$ be a set of pure states and let $\Phi : \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_a) \to \mathcal{L}(\mathcal{H}_B)$ be a channel. Let $\rho_a \in \mathcal{L}(\mathcal{H}_a)$ and $\rho_0 \in \mathcal{L}(\mathcal{H}_B)$ be density operators. Then $[\mathcal{S}, \Phi, \rho_a, \rho_0]$ is called a* private quantum channel *(PQC) if for any unit vector $|\phi\rangle \in \mathcal{S}$, we have*

$$\Phi(|\phi\rangle\langle\phi| \otimes \rho_a) = \rho_0.$$

We note that the ancilla ρ_a is often dropped in the literature; in this case, we simply consider the triple $[\mathcal{S}, \Phi, \rho_0]$.

2.2 Conditional Expectations and Trace Vectors

We recall a basic definition from operator algebras. Suppose there exists an orthogonal direct sum decomposition of a Hilbert space as $\mathcal{H} = \bigoplus_i (M_i \otimes N_i) \oplus K$. Let $\mathcal{A}$ be an algebra of operators in $\mathcal{L}(\mathcal{H})$ consisting of all operators that belong to the set $\mathcal{A} = \bigoplus_i (\mathrm{I}_{M_i} \otimes \mathcal{L}(N_i)) \oplus 0_K$, where 0_K is the zero operator on K. We call $\mathcal{A}$ a

(concrete finite dimensional) C-algebra.* $\mathcal{A}$ is unital if $\mathrm{I}_{\mathcal{H}} \in \mathcal{A}$; i.e., $\mathcal{K}$ is the zero subspace. A $*$-subalgebra $\mathcal{B}$ of $\mathcal{A}$ is a subset that is also a C*-algebra. See [Dav96] for basic C*-algebra theory.

Definition 2.2.1. *Let $\mathcal{A}$ be a C*-algebra and let $\mathcal{B} \subseteq \mathcal{A}$ be a unital *-subalgebra. We call a linear map $\mathcal{E}_{\mathcal{B}} : \mathcal{A} \to \mathcal{B}$ a* conditional expectation *of $\mathcal{A}$ onto $\mathcal{B}$ if*

(i) $\mathcal{E}_{\mathcal{B}}(b) = b$ *for all* $b \in \mathcal{B}$*;*

(ii) $\mathcal{E}_{\mathcal{B}}(b_1 a b_2) = b_1 \mathcal{E}_{\mathcal{B}}(a) b_2$ *for all* $b_1, b_2 \in \mathcal{B}$ *and for all* $a \in \mathcal{A}$*;*

(iii) $a \in \mathcal{A}$, $a \geq 0$ *implies* $\mathcal{E}_{\mathcal{B}}(a) \geq 0$.

Conditional expectations were first considered in [Ume54] and further explored in [Tom57]. Note that all conditional expectations are completely positive [Stø73]. We are interested in conditional expectations from $\mathbb{M}_n$ onto a subalgebra that are also quantum channels. We will therefore restrict ourselves to trace preserving conditional expectations.

We will call trace-preserving conditional expectations *conditional expectation channels*.

More examples of conditional expectation channels will be discussed below; we note here that the n-qubit completely depolarizing channel $\mathcal{E}_{\mathbb{C}}$ is the conditional expectation onto the trivial scalar algebra $\mathbb{C} \cdot \mathrm{I}_{2^n}$. One way to see how conditional expectations inevitably arise in the theory is through trace inner products.

Definition 2.2.2. *A linear functional* $\tau : \mathcal{A} \to \mathbb{C}$ *is a* faithful trace *if*

(i) $\tau(a_1 a_2) = \tau(a_2 a_1)$

(ii) $\tau(a^\dagger a) > 0$ *for all* $a \in \mathcal{A}$ *with* $a \neq 0$.

Given a faithful trace τ on $\mathcal{A}$ we can define an inner product $\langle a_1 \,|\, a_2 \rangle = \tau(a_1^\dagger a_2)$. We note that if $\mathcal{A}$ has a faithful trace τ, the orthogonal projection onto $\mathcal{B}$ with respect to this inner product is the unique τ-preserving conditional expectation from $\mathcal{A}$ to $\mathcal{B}$. The essential structure of this argument can be found in [Ume54]. The most well-known example of a faithful trace is the Hilbert-Schmidt inner product $\langle A \,|\, B \rangle = \mathrm{Tr}(A^\dagger B)$ on $\mathbb{M}_n$.

We now consider trace vectors, a notion that initially arose in work of Murray and von Neumann [MvN37], and has more recently been studied in the field of matrix theory.

Definition 2.2.3. *Let* $\mathcal{A}$ *be a* $*$*-subalgebra of* $\mathcal{L}(\mathcal{H}_n)$. *A vector* $|v\rangle$ *is a* trace vector *of* $\mathcal{A}$ *if*

$$\langle v|a|v\rangle = \frac{1}{n} \operatorname{Tr} a \quad \forall a \in \mathcal{A}.$$

More generally, given a density operator ρ_0, *we say* $|v\rangle$ *is a* trace vector with respect to ρ_0 *of* $\mathcal{A}$ *if*

$$\langle v|a|v\rangle = \mathrm{Tr}(\rho_0 a) \quad \forall a \in \mathcal{A}. \tag{2.2}$$

Thus by "trace vector", we really mean "trace vector with respect to $\frac{1}{n} \mathrm{I}_n$*".*

By letting $a = \mathrm{I}$ in the definition of a trace vector, we find $\langle v|v\rangle = 1$; that is, a trace vector has unit length. It is easy to build a trace vector from other trace vectors in order to create a more general class of examples. Indeed, if $|v_i\rangle$ is a trace vector of the algebra $\mathcal{A}_i = (\mathrm{I}_{M_i} \otimes \mathcal{L}(N_i)) \oplus 0_K$ for $i \in \{1, \ldots, q\}$, then $|v\rangle = \bigoplus_{i=1}^{q} |v_i\rangle$ is a trace vector of the algebra $\mathcal{A} = \bigoplus_{i=1}^{q} \mathcal{A}_i$. In this way, trace vectors behave predictably. This also allows us to consider each summand separately, as we will do later.

Example 2.2.4. *As a fundamental example for quantum information, consider a maximally entangled state $|\varphi_e\rangle \in \mathcal{H}_m \otimes \mathcal{H}_n$. If $m \geq n$, then we find for any $a \in \mathcal{L}(\mathcal{H}_n)$,*

$$\begin{aligned}
\langle \varphi_e | \, \mathrm{I}_m \otimes a | \varphi_e \rangle &= \frac{1}{n} \sum_{i=1}^{n} \langle e_i | \otimes \langle f_i | (\mathrm{I}_m \otimes a) \sum_{j=1}^{n} |e_j\rangle \otimes |f_j\rangle \\
&= \frac{1}{n} \sum_{i,j} \langle e_i | \, \mathrm{I}_m \, |e_j\rangle \langle f_i | a | f_j \rangle \\
&= \frac{1}{n} \sum_{i} 1 \cdot \langle f_i | a | f_i \rangle \\
&= \frac{1}{n} \operatorname{Tr}(\mathrm{I}_n \otimes a).
\end{aligned}$$

It follows by definition that $|\varphi_e\rangle$ is a trace vector for the algebra $\mathrm{I}_m \otimes \mathcal{L}(\mathcal{H}_n)$. If $m = n$ an analogous calculation works for $\mathcal{L}(\mathcal{H}_m) \otimes \mathrm{I}_n$.

The general case is clarified by the following theorem of [Per03], which we state below. A related infinite dimensional open problem dates back to von Neumann [Ge]. We recall that a vector $|v\rangle$ is a *separating vector* of an algebra $\mathcal{A}$ if $a|v\rangle = 0$ for some $a \in \mathcal{A}$ implies $a = 0$.

Theorem 2.2.5. *[Per03] If $\mathcal{A}$ is a unital $*$-subalgebra of $\mathbb{M}_n$, then the following conditions are equivalent:*

1. *$\mathcal{A}$ is unitarily equivalent to $\oplus_{i=1}^{q}\left(\mathrm{I}_{m_i}\otimes\mathbb{M}_{n_i}\right)$, where $m_i \geq n_i$ for all i and $\sum_{i=1}^{q} m_i n_i = n$.*

2. *$\mathcal{A}$ has a separating vector.*

3. *$\mathcal{A}$ has a trace vector.*

4. *There exists a set of trace vectors of $\mathcal{A}$ that form an orthonormal basis of $\mathbb{C}^n$.*

Example 2.2.6. *1. It is clear that $\mathbb{M}_n$ itself has no trace vectors—from theorem 2.2.5, if $\mathbb{M}_n$ did have a trace vector, then we could write $\mathbb{M}_n = U \oplus_{i=1}^{q} \left(\mathrm{I}_{m_i}\otimes\mathbb{M}_{n_i}\right) U^\dagger$ where $m_i \geq n_i$ for all i, $\sum_{i=1}^{q} m_i n_i = n$, and U is some unitary. However, this decomposition implies $m_i = 1$, so $m_i \geq n_i$ cannot be satisfied (other than for the case where $n = 1$). Also, it is clear that the equation from the definition of trace vectors cannot be satisfied by a single vector for* all *matrices in $\mathbb{M}_n$.*

2. *Let Δ_2 be the algebra of 2×2 diagonal matrices with respect to a basis $\{|0\rangle, |1\rangle\}$. A simple calculation shows that the trace vectors for Δ_2 are (up to complex phase) all vectors of the form $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\theta}|1\rangle)$, for $0 \leq \theta < 2\pi$; in other words, the set of all states that lie on the equator in the Bloch sphere representation for qubits (this point is further elucidated in the next section).*

2.3 Private Quantum Channels on the Bloch Sphere

In this section we give a geometric characterization of single qubit unital PQCs in terms of the Bloch sphere representation [NC00] for single qubit states. We also show how the private states for such PQCs are determined by trace vectors. An alternative description was discussed in [BZ07], where the entropy of sets of such private states was considered.

Every unital quantum channel is a random unitary channel in the single qubit case [LS93]. Thus, our private quantum channel $[\mathcal{S}, \mathcal{E}, \rho_0]$ in this case is given by a random unitary channel $\mathcal{E} : \mathbb{M}_2 \to \mathbb{M}_2$, a set of pure states $\mathcal{S}$, and an output density matrix ρ_0. We would like to allow for the possibility of orthonormal vectors in $\mathcal{S}$. As the channel is unital this can only occur if $\rho_0 = \frac{1}{2}\,\mathrm{I}$, and hence we shall focus on this case here.

We can write any pure state $|\psi\rangle \in \mathbb{C}^2$, up to a global phase shift, as

$$|\psi\rangle = \cos\left(\tfrac{\theta}{2}\right)|0\rangle \;+\; e^{i\phi}\sin\left(\tfrac{\theta}{2}\right)|1\rangle \tag{2.3}$$

where $0 \leq \theta \leq \pi$ and $0 \leq \phi < 2\pi$. The parameters θ and ϕ specify a point $\vec{r} = (\sin\theta\cos\phi,\ \sin\theta\sin\phi,\ \cos\theta)$ on the surface of the Bloch sphere. Alternatively, we can think of $\vec{r}$ as a vector satisfying $\|\vec{r}\| = 1$ for any pure state $|\psi\rangle$.

In general, we can associate to any density matrix $\rho \in \mathbb{M}_2$ a Bloch vector $\vec{r} \in \mathbb{R}^3$ satisfying $\|\vec{r}\| \leq 1$, where

$$\rho = \frac{\mathrm{I} + \vec{r}\cdot\vec{\sigma}}{2}. \tag{2.4}$$

We use $\vec{\sigma}$ to denote the *Pauli vector*, that is, $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)^T$. Indeed, if we were to write ρ as an arbitrary linear combination $\rho = k_0\,\mathrm{I} + k_1\sigma_x + k_2\sigma_y + k_3\sigma_z$ for $k_0, \ldots, k_3 \in \mathbb{C}$, we would find

1. $\rho = \rho^\dagger \Leftrightarrow k_0, \ldots, k_3 \in \mathbb{R}$;

2. $\mathrm{Tr}\rho = 1 \Leftrightarrow k_0 = 1/2$;

3. $\rho \geq 0 \Leftrightarrow k_1^2 + k_2^2 + k_3^2 \leq \left(\frac{1}{2}\right)^2$ (with $r_i/2 = k_i$ for $i = 1, 2, 3$, this is precisely $||\vec{r}|| \leq 1$).

A state is pure if and only if $\|\vec{r}\| = 1$; the maximally mixed state $\frac{1}{2}\,\mathrm{I}$ has Bloch vector $\vec{r} = \vec{0}$. That is, the closer the vector is to the centre of the sphere, the more "mixed" the state.

As discussed in [KR01], every linear map $\Phi : \mathbb{M}_2 \to \mathbb{M}_2$ can be represented in the basis $\{\mathrm{I}, \sigma_x, \sigma_y, \sigma_z\}$ by a 4×4 matrix $\mathbb{T}$, and Φ preserves the trace if and only if the first row of the matrix $\mathbb{T}$ satisfies $t_{1k} = \delta_{1k}$; i.e.,

$$\mathbb{T} = \begin{pmatrix} 1 & \mathbf{0} \\ \vec{t} & T \end{pmatrix}$$

where T is a 3×3 matrix, $\mathbf{0}$ is a row vector, and $\vec{t}$ is a column vector. Here, the submatrix T represents a deformation of the Bloch sphere, while the vector $\vec{t}$ represents a translation. The transformation Φ maps the subspace of Hermitian matrices into itself *iff* $\mathbb{T}$ is real; finally, the map Φ is unital *iff* $\vec{t} = \vec{0}$.

Thus, every unital qubit channel $\mathcal{E}$ can be represented as

$$\mathcal{E}\left(\frac{1}{2}\left[\mathrm{I}+\vec{r}\cdot\vec{\sigma}\right]\right)=\frac{1}{2}\left[\mathrm{I}+(T\vec{r})\cdot\vec{\sigma}\right], \tag{2.5}$$

where T is real, and we recall any density matrix can be written as in equation (2.4). These affine mappings of the Bloch sphere into itself are also discussed in section 8.3.2 of [NC00].

We are of course interested in cases where $\mathcal{S}$ is nonempty. This is easily seen to occur precisely when T in equation (2.5) has non-trivial nullspace. Thus we consider the subspace of vectors $\vec{r}$ such that $T\vec{r}=0$ is one, two, or three-dimensional.

In the single qubit case, the unital subalgebras of the algebra $\mathcal{A}=\mathbb{M}_2$ can be easily classified. They are $\mathbb{M}_2$, $\mathbb{C}\cdot\mathrm{I}_2$ (the two trivial cases), and, up to unitary conjugation, Δ_2, the subalgebra of all diagonal matrices in $\mathbb{M}_2$. To be precise, this third case refers to the subalgebras $\mathcal{B}$ of the form $U^\dagger\Delta_2 U$, where $U\in\mathcal{A}$ is unitary.

Theorem 2.3.1. *Let $\mathcal{E}:\mathbb{M}_2\to\mathbb{M}_2$ be a unital qubit channel, with T the mapping induced by $\mathcal{E}$ as in equation (2.5). Then there are three possibilities for a private quantum channel $[\mathcal{S},\mathcal{E},\frac{1}{2}\mathrm{I}]$ with $\mathcal{S}$ nonempty:*

1. *If the nullspace of T is 1-dimensional, then $\mathcal{S}$ consists of a pair of orthonormal states.*

2. *If the nullspace of T is 2-dimensional, then the set $\mathcal{S}$ is the set of all trace vectors of the subalgebra $U^\dagger\Delta_2 U$ of 2×2 diagonal matrices up to unitary equivalence (where U is arbitrary).*

3. *If the nullspace of T is 3-dimensional, then $\mathcal{E}$ is the completely depolarizing channel and $\mathcal{S}$ is the set of all unit vectors. In other words, $\mathcal{S}$ is the set of all trace vectors of $\mathbb{C} \cdot \mathrm{I}_2$.*

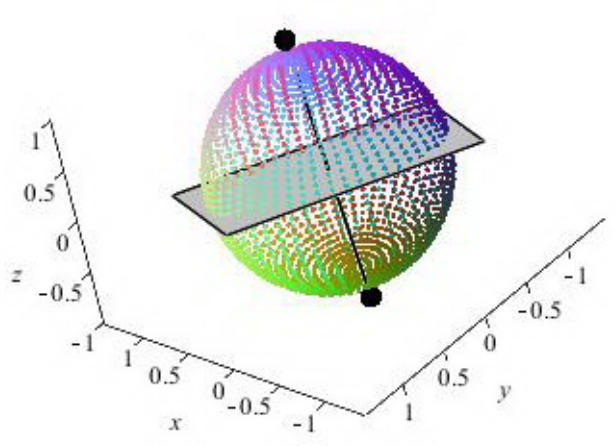

Figure 2.1: Case (1) of theorem 2.3.1

Proof. We shall write $\vec{r}_\phi$ for the Bloch sphere vector representation of a single qubit state $|\phi\rangle$. It is clear from equation (2.5) that $\mathcal{E}(|\phi\rangle\langle\phi|) = \frac{1}{2}\,\mathrm{I}$ if and only if $T\vec{r}_\phi = 0$. Hence the relevant set that yields private states here is the intersection of the nullspace of T and the surface of the Bloch sphere.

Case (1): The nullspace of T is 1-dimensional. In this case, the nullspace is a single line through the origin of the Bloch sphere and the range of T is a plane through the

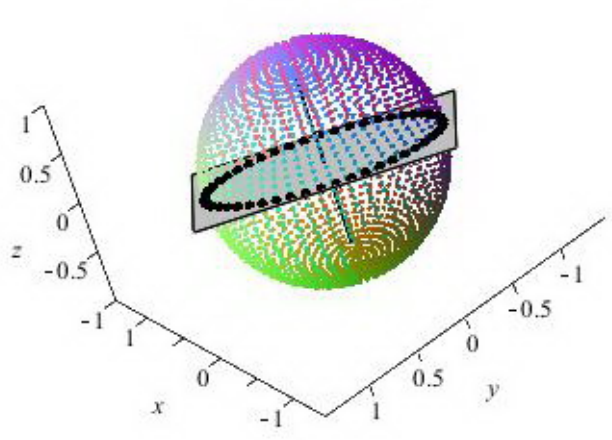

Figure 2.2: Case (2) of theorem 2.3.1

origin. Obviously this line meets the surface of the Bloch sphere in two antipodal points. These two antipodal points correspond to a pair of orthonormal single qubit states. Figure 1 gives an example.

Case (2): The nullspace of T is 2-dimensional. In this case, the nullspace is a plane through the origin of the Bloch sphere. This plane meets the surface of the sphere in a great circle. The pure states corresponding to the points on this circle are precisely the private states for the channel. See an illustration in Figure 2.

To see how these private states arise from the trace vector perspective, let us consider the action of the channel more directly. As the nullspace of T is 2-dimensional,

its range is a line through the origin. For simplicity we shall assume this line is the z-axis; other cases are unitarily equivalent to this case. Thus, the range of T intersects the sphere in the north and south poles, corresponding to the pure states $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ respectively. The action of T here will be a possible rotation of the Bloch sphere followed by a projection of the sphere onto the z-axis, followed by a possible contraction. By unitary equivalence, we only need consider the case where there is no initial rotation of the Bloch sphere. In terms of the Pauli matrices $\sigma_x, \sigma_y, \sigma_z$, this means the action of the channel is given by $\mathcal{E}(\sigma_x) = 0$, $\mathcal{E}(\sigma_y) = 0$ and $\mathcal{E}(\sigma_z) = p\sigma_z$ for some $0 < p \leq 1$.

Now Δ_2 is the algebra of all diagonal matrices with respect to the ordered basis $\{|0\rangle, |1\rangle\}$; explicitly, Δ_2 is the set of all operators of the form $a|0\rangle\langle 0| + b|1\rangle\langle 1|$ for arbitrary scalars a, b. Then the projection onto the z-axis is a conditional expectation onto the subalgebra Δ_2; call it $\mathcal{E}_\Delta$. Explicitly,

$$\mathcal{E}_\Delta\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}, \text{ for any matrix } \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

One can check directly that $\mathcal{E} = p\mathcal{E}_\Delta + (1-p)\mathcal{E}_{\mathbb{C}}$, where $\mathcal{E}_{\mathbb{C}}$ is the completely depolarizing channel.

As $\mathcal{E}_{\mathbb{C}}$ adds no restrictions to the private states for $\mathcal{E}$, it suffices to show that the trace vectors for Δ_2 are precisely the pure states that lie on the equator of the Bloch sphere. But the equator states are precisely the states that satisfy $|\langle\phi\,|\,0\rangle| = \frac{1}{\sqrt{2}} = |\langle\phi\,|\,1\rangle|$ (that is, θ in equation (2.3) is $\pi/2$). And it is easy to see that these are the states which do indeed satisfy the trace vector condition for the algebra Δ_2.

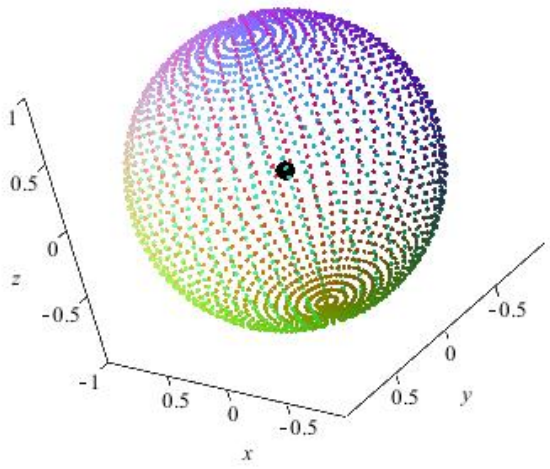

Figure 2.3: Case (3) of theorem 2.3.1

Case (3): The nullspace of T is 3-dimensional, in other words T is the zero operator. In this case T maps the entire Bloch sphere to its origin, which corresponds to the maximally mixed state $\frac{1}{2}\,\mathrm{I}$, as shown in Figure 3. It is clear in this case that $\mathcal{E}$ is the completely depolarizing channel $\mathcal{E}_{\mathbb{C}}$. Moreover, the set $\mathcal{S}$ has no restrictions; that is, $\mathcal{S}$ is the set of all unit vectors. In other words, $\mathcal{S}$ is the set of all trace vectors of $\mathbb{C} \cdot \mathrm{I}_2$. ■

Remark 2.3.2. *The term* private code *typically refers to the special case when the set $\mathcal{S}$ of private states corresponds to $\mathcal{L}(\mathcal{K})$ for some Hilbert space $\mathcal{K}$. This situation is ideal in that it allows for superpositions of private pure states to also be private. That is, if $|\psi_i\rangle \in \mathcal{S}$ for $i = 1, \ldots, m$ and $\mathcal{S}$ is a private code for some channel Φ, then*

$\sum_{i=1}^{m} a_i|\psi_i\rangle \in \mathcal{S}$ *for all* $a_i \in \mathbb{C}$ *satisfying* $\sum_{i=1}^{m} |a_i|^2 = 1$ *(such* a_i *are called* probability amplitudes*). Note that only case (3) of theorem 2.3.1 is a private code; both cases (1) and (2) do not allow for arbitrary superpositions of private states in* $\mathcal{S}$ *to also be elements of* $\mathcal{S}$*.*

2.4 Private States for Conditional Expectation Channels

The following result clarifies the general connection between conditional expectation channels, trace vectors and private states.

Theorem 2.4.1. *Let* $\mathcal{E} : \mathbb{M}_n \to \mathcal{A}$ *be a conditional expectation channel. Then* $[\mathcal{S}, \mathcal{E}, \rho_0]$ *is a private quantum channel if and only if* $\mathcal{S}$ *is a set of trace vectors of* $\mathcal{A}$ *with respect to* $\rho_0 \in \mathcal{A}$*.*

Proof. Let us first assume that $[\mathcal{S}, \mathcal{E}, \rho_0]$ is a PQC. Then $\mathcal{E}(|v\rangle\langle v|) = \rho_0$ for all $|v\rangle \in \mathcal{S}$, and in particular note that ρ_0 belongs to $\mathcal{A}$. Thus for all $|v\rangle \in \mathcal{S}$ and for all $a \in \mathcal{A}$, we have

$$\begin{aligned} \langle v|a|v\rangle &= \mathrm{Tr}(|v\rangle\langle v|a) \\ &= \mathrm{Tr}(\mathcal{E}(|v\rangle\langle v|a)) \\ &= \mathrm{Tr}(\mathcal{E}(|v\rangle\langle v|)a) = \mathrm{Tr}(\rho_0 a), \end{aligned}$$

where the second and third identities follow from the trace preservation and condi-

tional expectation properties of $\mathcal{E}$ respectively. It follows that the states of $\mathcal{S}$ are trace vectors of $\mathcal{A}$ with respect to ρ_0.

For the converse, observe that when the vector states of $\mathcal{S}$ are trace vectors of $\mathcal{A}$ with respect to ρ_0, a similar calculation shows for all $|v\rangle \in \mathcal{S}$ and for all $a \in \mathcal{A}$ that

$$\begin{aligned} \mathrm{Tr}(\rho_0 a) &= \langle v|a|v\rangle \\ &= \mathrm{Tr}(|v\rangle\langle v|a) \\ &= \mathrm{Tr}(\mathcal{E}(|v\rangle\langle v|a)) = \mathrm{Tr}(\mathcal{E}(|v\rangle\langle v|)a). \end{aligned}$$

As ρ_0 belongs to $\mathcal{A}$, it follows that $[\mathcal{S}, \mathcal{E}, \rho_0]$ forms a private quantum channel. ■

Example 2.4.2. *Of course the three cases of theorem 2.3.1 when applied to a unital single qubit conditional expectation channel $\mathcal{E}_{\mathcal{A}} : \mathbb{M}_2 \to \mathcal{A}$ are covered by this theorem. Indeed, applying theorem 2.4.1 to $\mathcal{E}_{\mathcal{A}}$ and letting $\rho_0 = \frac{1}{2}\mathrm{I}$ yields $[\mathcal{S}, \mathcal{E}_{\mathcal{A}}, \frac{1}{2}\mathrm{I}]$ is a PQC if and only if $\mathcal{A} = U^\dagger \Delta_2 U$ or $\mathcal{A} = \mathbb{C} \cdot \mathrm{I}_2$ and $\mathcal{S}$ is a set of trace vectors of $\mathcal{A}$. Case (1) of theorem 2.3.1 is an example of when $\mathcal{S}$ is a proper subset of the set of all trace vectors of $U^\dagger \Delta_2 U$, whereas Case (2) occurs when $\mathcal{S}$ is the entire set. Case (3) occurs when $\mathcal{S}$ is the set of all trace vectors of $\mathbb{C} \cdot \mathrm{I}_2$.*

Example 2.4.3. *Conditional expectations arise as the most basic non-trivial examples of private quantum communication using a private shared Cartesian frame [BHS05]. A private shared Cartesian frame is a non-local resource shared between Alice and Bob that allows for both private classical and private quantum communication. Let $\mathcal{H} = (\mathbb{C}^2)^{\otimes N}$, and for simplicity suppose N is even. Decompose the space*

as

$$(\mathbb{C}^2)^{\otimes N} = \bigoplus_{j=0}^{N/2} \mathbb{H}_j \otimes \mathbb{K}_j,$$

where the special unitary group $\mathrm{SU}(2)$ *acts irreducibly on* $\mathbb{H}_j$ *and trivially on* $\mathbb{K}_j$*. As formulated in [BHS05], if Alice and Bob share a reference frame to which Eve does not have access, and Alice prepares* N *qubits in a state* ρ *and sends them to Bob, Eve will see the resulting state simply as a mixture of all rotations* $\Omega \in \mathrm{SU}(2)$*. This situation can be summed up with the channel* $\mathcal{E}$*, defined by*

$$\mathcal{E}(\rho) = \sum_{j=0}^{N/2} (\mathcal{E}_{\mathbb{C}j} \otimes id_{\mathbb{K}_j})(\Pi_j \rho \Pi_j),$$

where $\mathcal{E}_{\mathbb{C}j}$ *is the completely depolarizing channel on* $\mathbb{H}_j$ *and* Π_j *is the projection onto* $\mathbb{H}_j$*. One can see immediately that* $\mathcal{E}$ *is in fact a conditional expectation channel that maps onto the algebra* $\oplus_j(\mathrm{I}_{\mathbb{H}_j} \otimes \mathcal{L}(\mathbb{K}_j))$*. Thus, as noted in theorem 2.4.1, private states for* $\mathcal{E}$ *can be found using trace vectors, which in this case can be constructed on the summands of the direct sum in a manner analogous to example 2.2.4.*

2.5 Quantum Error Correction

Definition 2.5.1. *We say that a subspace of states* $\mathcal{C}$ *is an* error correcting code *for a channel* $\mathcal{E}$ *if there exists an error correction operation (that is, another quantum channel)* $\mathcal{R}$ *such that* $\mathcal{R} \circ \mathcal{E}(\rho) = \rho$ *for all* ρ *supported on* $\mathcal{C}$*.*

The celebrated Knill-Laflamme theorem in quantum error correction [KL97] gives testable conditions in terms of Kraus operators for determining whether a given code

is correctable. In particular, it characterizes a correctable code $\mathcal{C}$ for a channel strictly in terms of its Kraus operators and the projection onto $\mathcal{C}$.

Theorem 2.5.2. *(Knill-Laflamme Theorem for QEC) [KL97] Let $\mathcal{C}$ be a quantum code and let P be the projector onto $\mathcal{C}$. Suppose $\mathcal{E}$ is a quantum channel with Kraus operators $\{E_i\}$. Then there exists an error-correction operation $\mathcal{R}$ correcting $\mathcal{E}$ on $\mathcal{C}$ if and only if $PE_i^\dagger E_j P = \alpha_{ij} P$ for some Hermitian matrix $\alpha = (\alpha_{ij})_{ij}$ of complex numbers.*

The generalization of these conditions to the case of operator error-correcting subsystems was established in [KLP05, KLPL06, NP07].

2.6 The Connection between QEC and Quantum Cryptography

As in quantum error correction, we wish to allow for arbitrary superpositions of our code states and this demands the set of states considered is linearly closed. Thus we restrict ourselves to consider private subspaces, rather than private subsets. This restriction will be relaxed slightly to subsystems later in this chapter. In what follows, we will let $P = P_{\mathcal{C}}$ be the projection onto the subspace $\mathcal{C}$.

Theorem 2.6.1. *[KKS08] Given a conjugate pair of CPTP maps Φ, $\Phi^\sharp$, a code is an error-correcting code for one if and only if it is a private code for the other.*

The extreme example of this phenomena is given by a unitary channel paired with the completely depolarizing channel—where the entire Hilbert space is the code. We illustrate this connection further with a pair of fairly simple examples.

Example 2.6.2. *Consider the 2-qubit swap channel $\Phi(\sigma \otimes \rho) = \rho \otimes \sigma$ from example 1.2.13. Since the channel is implemented by a single unitary Kraus operator U, it is easily seen to be correctable for all density operators $\rho \in \mathbb{M}_4$, since the recovery operator would simply be the channel having single Kraus operator $U^\dagger$. On the other hand, one can compute, using the Kraus operators determined in example 1.2.13, that $\Phi^\sharp(\rho) = \rho_{11} + \rho_{22} + \rho_{33} + \rho_{44} = \operatorname{Tr}\rho = 1$ for all $\rho \in \mathbb{M}_4$, and so the complementary channel is private for all $\rho \in \mathbb{M}_4$, as expected from theorem 2.6.1.*

Example 2.6.3. *Consider the 2-qubit phase flip channel Φ with (equally weighted) Kraus operators $\{\mathrm{I}, Z\,\mathrm{I}\}$ from example 1.2.14. We have $\Phi(\rho) = \frac{1}{2}(\rho + Z\,\mathrm{I}\,\rho Z\,\mathrm{I})$ and $\Phi^\sharp(\rho) = \sum_{i=1}^4 R_i \rho R_i^\dagger$ for all 2-qubit ρ. It is clear that the code $\{|00\rangle, |01\rangle\}$ is correctable for Φ; in fact the recovery operator would be the identity map, and the channel is said to be noiseless/decoherence-free. Thus we know this code is private for the conjugate channel $\Phi^\sharp$.*

The channel $\Phi^\sharp$ is private for any density matrix with support inside the codespace. Every density operator ρ supported on $\{|00\rangle, |01\rangle\}$ is of the form

$$\rho = \begin{pmatrix} \rho_{11} & \rho_{12} & 0 & 0 \\ \rho_{21} & \rho_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

and by direct calculation, we find

$$\begin{aligned}\Phi^\sharp(\rho) &= \frac{1}{2}\begin{pmatrix}\rho_{11}+\rho_{22}+\rho_{33}+\rho_{44} & \rho_{11}+\rho_{22}-\rho_{33}-\rho_{44}\\ \rho_{11}+\rho_{22}-\rho_{33}-\rho_{44} & \rho_{11}+\rho_{22}+\rho_{33}+\rho_{44}\end{pmatrix}\\ &= \frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}.\end{aligned}$$

Thus the channel is indeed private on $\mathcal{S}$ *with* $\rho_0 = \frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$.

2.6.1 Knill-Laflamme Analogue for Private Quantum Channels

Using the algebraic bridge given by the notion of conjugate channels, we investigate Knill-Laflamme type conditions for private codes. The following result gives an algebraic characterization of private quantum codes in terms of the dual map of a channel, and can be seen as a move toward a structure theory for private quantum codes. This theorem can be proved directly, or it can be proved by considering complementary channels and using theorem 2.6.1. We give details for both proofs, as we believe it is useful to see the power of the complementarity "in action", although the direct proof shows that it is not necessary in this case.

Theorem 2.6.4. *[KP12] Let* $\Phi : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ *be a quantum channel. Then a subspace* $\mathcal{C}$ *of* $\mathcal{H}_A$ *is private for* Φ *with output state* ρ_0*; i.e.,* $\Phi(\rho) = \rho_0$ *for all* $\rho \in \mathcal{L}(\mathcal{C})$*, if and only if for any* $X \in \mathcal{L}(\mathcal{H}_B)$ *there exists a* $\lambda_X \in \mathbb{C}$ *such that*

$$P_{\mathcal{C}}\Phi^\dagger(X)P_{\mathcal{C}} = \lambda_X P_{\mathcal{C}},$$

where $P_{\mathcal{C}}$ *is the projection onto* $\mathcal{C}$*. And in this case,* $\lambda_X = \mathrm{Tr}(\rho_0 X)$.

Proof. Choose an orthonormal set $\{|\phi_k\rangle\}$ in $\mathcal{H}_A$ such that $P_{\mathcal{C}} = \sum_{k=1}^{\mathrm{rank}P_{\mathcal{C}}} |\phi_k\rangle\langle\phi_k|$. Let us first assume that Φ is private on $\mathcal{C}$. By manipulating the bra-ket notation, and using the definition of $\Phi^\dagger$, we find

$$\begin{aligned}
P_{\mathcal{C}}\Phi^\dagger(X)P_{\mathcal{C}} &= \sum_{k,\ell=1}^{\mathrm{rank}P_{\mathcal{C}}} |\phi_k\rangle\langle\phi_k|\Phi^\dagger(X)|\phi_\ell\rangle\langle\phi_\ell| \\
&= \sum_{k,\ell} \langle\phi_k|\Phi^\dagger(X)|\phi_\ell\rangle|\phi_k\rangle\langle\phi_\ell| \\
&= \sum_{k,\ell} \mathrm{Tr}\left(\Phi^\dagger(X)|\phi_\ell\rangle\langle\phi_k|\right)|\phi_k\rangle\langle\phi_\ell| \\
&= \sum_{k,\ell} \mathrm{Tr}\left(X\,\Phi\left(|\phi_\ell\rangle\langle\phi_k|\right)\right)|\phi_k\rangle\langle\phi_\ell| \\
&= \sum_{k,\ell} \mathrm{Tr}\left(X\rho_0\right)\mathrm{Tr}(|\phi_\ell\rangle\langle\phi_k|)|\phi_k\rangle\langle\phi_\ell| \\
&= \sum_{k} \mathrm{Tr}\left(X\rho_0\right)|\phi_k\rangle\langle\phi_k| = \mathrm{Tr}\left(X\rho_0\right)P_{\mathcal{C}},
\end{aligned}$$

where the fifth equality follows because $|\phi_\ell\rangle\langle\phi_k| \in \mathcal{L}(\mathcal{C})$ and Φ is trace-preserving, so that $\Phi(|\phi_\ell\rangle\langle\phi_k|) = \mathrm{Tr}(|\phi_\ell\rangle\langle\phi_k|)\rho_0$. Defining $\lambda_X := \mathrm{Tr}\left(X\rho_0\right)$ completes this direction of the proof.

For the other direction, we assume $P_{\mathcal{C}}\Phi^\dagger(X)P_{\mathcal{C}} = \lambda_X P_{\mathcal{C}}$ for all $X \in \mathcal{L}(\mathcal{H}_B)$. Similar to the above calculation, we can write this as

$$\sum_{k,\ell} \mathrm{Tr}\left(X\Phi\left(|\phi_\ell\rangle\langle\phi_k|\right)\right)|\phi_k\rangle\langle\phi_\ell| = \lambda_X \sum_k |\phi_k\rangle\langle\phi_k|.$$

For each k, compressing this equation by the rank one projection $|\phi_k\rangle\langle\phi_k|$ yields

$$\lambda_X = \mathrm{Tr}\left(X\Phi\left(|\phi_k\rangle\langle\phi_k|\right)\right) \quad \forall X \in \mathcal{L}(\mathcal{H}_B).$$

The left hand side of the above equation does not depend on k, and so $\Phi\left(|\phi_k\rangle\langle\phi_k|\right)$ is a fixed density operator, say ρ_0, independent of k. And as the basis $\{|\phi_k\rangle\}$ for $\mathcal{C}$

was arbitrary it follows that $\Phi(\rho) = \rho_0$ for all $\rho \in \mathcal{L}(\mathcal{C})$. Thus $\mathcal{C}$ is private for Φ and this completes the proof.

Alternate Proof:

With both the conjugate channel machinery and Knill-Laflamme result in hand, we can arrive at the conclusion of theorem 2.6.4 from an alternate perspective. Indeed, letting $\{V_i\}_{i=1}^K$ be the Kraus operators of the channel Φ and $\{R_\mu\}_{\mu=1}^b$ be the Kraus operators of its conjugate channel $\Phi^\sharp$ (as before), and $\{|f_\mu\rangle\}$ any orthonormal basis for $\mathcal{H}_B$, we compute

$$R_\mu^* R_\nu = \begin{pmatrix} V_1^\dagger|f_\mu\rangle & V_2^\dagger|f_\mu\rangle & \dots & V_K^\dagger|f_\mu\rangle \end{pmatrix} \begin{pmatrix} \langle f_\nu|V_1 \\ \langle f_\nu|V_2 \\ \vdots \\ \langle f_\nu|V_K \end{pmatrix} = \sum_{i=1}^K V_i^\dagger(|f_\mu\rangle\langle f_\nu|)V_i. \tag{2.6}$$

We recall the fact that the code being private for Φ is equivalent to it being correctable for the conjugate channel $\Phi^\sharp$. Then, by the Knill-Laflamme conditions from theorem 2.5.2 and the calculation of equation (2.6), we have for all $\mu, \nu \in \{1, \dots, b\}$ that $\mathcal{C}$ is private for Φ if and only if

$$\lambda_{\mu\nu} P_\mathcal{C} \quad = \quad P_\mathcal{C} R_\mu^\dagger R_\nu P_\mathcal{C} = P \sum_{i=1}^K V_i^\dagger(|f_\mu\rangle\langle f_\nu|)V_i P = P\Phi^\dagger(|f_\mu\rangle\langle f_\nu|)P.$$

And as the identity holds for arbitrary matrix units $|f_\mu\rangle\langle f_\nu|$, by linearity it extends to all operators X.

■

Note that if $\{|e_i\rangle\}$ is an orthonormal basis for $\mathcal{H}_B$, then the scalars λ_X obtained in the theorem from the use of the set of matrices $\{|e_i\rangle\langle e_j|\}$ form a Hermitian matrix;

in particular the matrix coefficients for the output state ρ_0 in this basis. We also note that in many cases of interest, such as the class of Pauli channels for instance, the channel Φ is self-dual; that is, $\Phi = \Phi^\dagger$.

Corollary 2.6.5. *Suppose we use the spectral theorem to write $P_\mathcal{C}$ and ρ_0 in their diagonal forms: $P_\mathcal{C} = \sum_{\ell=1}^{d} |\phi_\ell\rangle\langle\phi_\ell|$, where $d = \text{rank}P_\mathcal{C} \leq \dim \mathcal{H}^A$, and $\rho_0 = \sum_{k=1}^{m} c_k|\psi_k\rangle\langle\psi_k|$ where $m = \dim \mathcal{H}^B$. Then the subspace $\mathcal{C}$ is private for $\Phi = \{V_i\}$ with output ρ_0 for all $\rho \in \mathcal{L}(\mathcal{C})$ if and only if for each i, we have*

$$P_\mathcal{C} V_i^\dagger = \sum_{k,\ell} u_{i,k\ell}\sqrt{c_k}|\phi_\ell\rangle\langle\psi_k|,$$

where $u_{i,k\ell}$ form an isometry $U \equiv (u_{i,k\ell})_{i,k\ell}$ (unitary when U is a square matrix).

Proof. We know that the subspace $\mathcal{C}$ is private for $\Phi = \{V_i\}$ with output ρ_0 for all $\rho \in \mathcal{L}(\mathcal{C})$ if and only if

$$P_\mathcal{C}\Phi^\dagger(X)P_\mathcal{C} = \text{Tr}(\rho_0 X)P_\mathcal{C}.$$

Let us consider both the left-hand side and right-hand side of this equality as completely positive maps. The left-hand side is the composition $X \mapsto \Phi^\dagger(X) \mapsto P_\mathcal{C}\Phi^\dagger(X)P_\mathcal{C}$, and so its Kraus operators are $\{P_\mathcal{C}V_i^\dagger\}$. The right-hand side is the map $X \mapsto \text{Tr}(\rho_0 X)P_\mathcal{C}$ given by the Kraus operators $\{\sqrt{c_k}|\phi_\ell\rangle\langle\psi_k|\}_{k,\ell=1}^{m,d}$.

From remark 1.2.4, it follows that there exists U as described if and only if $\mathcal{C}$ is private for Φ with output ρ_0. ■

Corollary 2.6.6. *Theorem 2.4.1 follows as a direct result of theorem 2.6.4.*

Proof. Let $\mathcal{E} : \mathbb{M}_n \to \mathcal{A}$ be a conditional expectation channel. Let $b \in \mathcal{A}^\perp$. Then $\mathrm{Tr}(ab) = 0$ for all $a \in \mathcal{A}$. So

$$\begin{aligned} 0 &= \mathrm{Tr}(ab) = \mathrm{Tr}(\mathcal{E}(ab)) \text{ since } \mathcal{E} \text{ is trace preserving} \\ &= \mathrm{Tr}(a\mathcal{E}(b)) \text{ by the properties of conditional expectations.} \end{aligned}$$

Note $\mathcal{E}(b) \in \mathcal{A}$. Choose $a = \mathcal{E}(b)^\dagger$. Then, using the Frobenius norm, we have $||\mathcal{E}(b)||_F^2 = \mathrm{Tr}(\mathcal{E}(b)^\dagger \mathcal{E}(b)) = 0$. It follows that $\mathcal{E}(b) = 0$ for all $b \in A^\perp$. Thus a conditional expectation is an orthogonal projection. Since orthogonal projections are self-adjoint, we conclude $\mathcal{E} = \mathcal{E}^\dagger$.

Now, consider the equation

$$P_\mathcal{C} \Phi^\dagger(X) P_\mathcal{C} = \mathrm{Tr}(\rho_0 X) P_\mathcal{C}$$

from theorem 2.6.4. Here, $\Phi^\dagger = \mathcal{E}^\dagger = \mathcal{E}$. Let $|v\rangle \in \mathcal{C}$ be an arbitrary unit vector and note that $P_\mathcal{C}|v\rangle = |v\rangle$. Then our equation is equivalent to

$$\langle v|\mathcal{E}(X)|v\rangle = \mathrm{Tr}(\rho_0 X)\langle v|v\rangle = \mathrm{Tr}(\rho_0 X).$$

Setting $X = a \in \mathcal{A}$, we have $\langle v|a|v\rangle = \langle v|\mathcal{E}(a)|v\rangle = \mathrm{Tr}(\rho_0 a)$ for all $a \in \mathcal{A}$. In other words, $[\mathcal{C}, \mathcal{E}, \rho_0]$ is a private quantum channel if and only if $\mathcal{C}$ is a set of trace vectors of $\mathcal{A}$ with respect to $\rho_0 \in \mathcal{A}$. ■

Example 2.6.7. *Again, consider the 2-qubit swap channel* $\Phi(\sigma \otimes \rho) = \rho \otimes \sigma$ *from example 1.2.13. The complementary channel* $\Phi^\sharp$ *has Kraus operators* $\{R_i\}$ *given in the example. Since* $\mathcal{C} = \mathbb{M}_4$*, the projection onto* $\mathcal{C}$ *is the identity:* $P_\mathcal{C} = \mathrm{I}_4$ *and the*

equation in theorem 2.6.4 simplifies to $(\Phi^\sharp)^\dagger(X) = \mathrm{Tr}(\rho_0 X)\,\mathrm{I}_4$ *where* $(\Phi^\sharp)^\dagger : \mathbb{C} \to \mathbb{M}_4$ *has Kraus operators* $\{R_i^\dagger\}$, $X \in \mathbb{C}$, *and one can compute* $(\Phi^\sharp)^\dagger(X) = \sum_i R_i^\dagger X R_i = \mathrm{diag}(X, X, X, X) = X\,\mathrm{I}_4$. *On the other hand,* $\mathrm{Tr}(\rho_0 X)\,\mathrm{I}_4$ *is equal to* $X\,\mathrm{I}_4$ *precisely when* $\rho_0 = 1 \in \mathbb{C}$. *Note that we obtain the same result as in example 2.6.2; namely,* $[\mathbb{M}_4, \Phi, 1]$ *is a private quantum channel.*

Example 2.6.8. *Consider again the 2-qubit phase flip channel* Φ *with (equally weighted) Kraus operators* $\{\mathrm{I}, Z\,\mathrm{I}\}$ *from example 1.2.13. We have* $(\Phi^\sharp)^\dagger : \mathbb{M}_2 \to \mathbb{M}_4$. *Using the projection*

$$P_{\mathcal{C}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

the equation $P_{\mathcal{C}}(\Phi^\sharp)^\dagger(X)P_{\mathcal{C}} = \mathrm{Tr}(\rho_0 X)P_{\mathcal{C}}$ *in theorem 2.6.4 can be verified by setting* X *to be each matrix unit* $|i\rangle\langle j|$. *Doing so yields* $[\rho_0]_{ij} = \frac{1}{2}$ *for all* $1 \leq i, j \leq 2$. *Thus the channel* $\Phi^\sharp$ *is indeed private on the code* $\{|00\rangle, |01\rangle\}$ *(which corresponds to the projection* $P_{\mathcal{C}}$ *given above) with* $\rho_0 = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, *as we saw in example 2.6.3.*

2.6.2 Private Quantum Subsystems

We now consider the most general notion of a private quantum code, which involves the encoding of quantum bits into subsystems. Given a Hilbert space $\mathcal{H}$ representing our quantum system, if we can write $\mathcal{H} = (\mathcal{H}^A \otimes \mathcal{H}^B) \oplus (\mathcal{H}^A \otimes \mathcal{H}^B)^\perp$, then the Hilbert spaces $\mathcal{H}^A$ and $\mathcal{H}^B$ represent *subsystems* of the system $\mathcal{H}$. The subspaces of $\mathcal{H}$ can be viewed as subsystems $\mathcal{H}^B$ for which $\mathcal{H}^A$ is one-dimensional. A subscript

such as σ_B means the operator belongs to $\mathcal{L}(\mathcal{H}^B)$.

The question of privacy then becomes: Is there a subsystem $\mathcal{H}^A$ that is private for Φ? Private quantum channels, private subspaces, and what we refer to as "operator" private subsystems—are captured as special cases of this general phenomena. Formally, we have the following definition:

Definition 2.6.9. *A subsystem $\mathcal{H}^B$ is a* private subsystem *for Φ if there is a $\rho_0 \in \mathcal{L}(\mathcal{H})$ and $\sigma_A \in \mathcal{L}(\mathcal{H}^A)$ such that*

$$\Phi(\sigma_A \otimes \sigma_B) = \rho_0 \quad \forall \sigma_B \in \mathcal{L}(\mathcal{H}^B). \tag{2.7}$$

A subsystem $\mathcal{H}^B$ is an operator private subsystem *for Φ if there is a $\rho_0 \in \mathcal{L}(\mathcal{H})$ and $\sigma_A \in \mathcal{L}(\mathcal{H}^A)$ such that*

$$\Phi(\sigma_A \otimes \sigma_B) = \rho_0 \quad \forall \sigma_A \in \mathcal{L}(\mathcal{H}^A), \forall \sigma_B \in \mathcal{L}(\mathcal{H}^B). \tag{2.8}$$

Operator private subsystems are precisely the private subsystems that are complementary to operator quantum error-correcting subsystems discussed in [KKS08]. If a channel Φ has an operator private subsystem, one can regard Φ as a product of channels on the individual subsystems $\mathcal{H}^A$ and $\mathcal{H}^B$ when the channel is restricted to the combined product subspace $\mathcal{H}^A \otimes \mathcal{H}^B$. That is, since the output is independent of the inputs of both subsystems, the channel itself can be thought to act as $\Phi(\sigma_A \otimes \sigma_B) = \Phi_A(\sigma_A) \otimes \Phi_B(\sigma_B) \ \ \forall \sigma_A \in \mathcal{L}(\mathcal{H}^A), \forall \sigma_B \in \mathcal{L}(\mathcal{H}^B)$. Such private subsystems cannot exist without the existence of private subspaces; indeed, if equation (2.8)

holds, it follows that every sub*space* $|\psi\rangle \otimes \mathcal{H}^B$ is private for Φ for any fixed pure state $|\psi\rangle$ on $\mathcal{H}^A$.

We note that while the established definition of private quantum channel [AMTdW00] does include an ancilla as above, the extra ancilla state (which corresponds to σ_A in equation (2.7)) and subsystem structure does not figure centrally into the results of the paper; indeed, all examples of private channels provided therein have either a private set or a private subspace of states, rather than a private subsystem. Further examples in the literature [BR03, BRS04, BHS05] thus far have either been of operator type, or are already subspaces.

We start our investigation into private subsystems by considering a very basic type of channel that arises naturally in the physical setting: a class of phase damping channels. We show that certain classes of channels can only be private in the subtle subsystem sense; thus establishing that private subsystems can exist in the absence of private subspaces by giving a bona fide example of a private subsystem.

We ask: is it possible to start with a single qubit channel that is not private, but for which a certain number of copies of the channel has a single qubit private code? And if so, what is the minimal number of qubits required?

Consider the family of phase damping channels that can be applied to any qubit of a larger Hilbert space $\mathcal{H}$ of n qubits,

$$\Lambda_i(\rho) = \frac{1}{2}(\rho + Z_i \rho Z_i), \qquad \forall \rho \in \mathcal{L}(\mathcal{H}),$$

where we recall that Z_i is notational shorthand for the operator Z being applied to

qubit i, and I_2 being applied to all other qubits. A single qubit phase damping channel is not private. Indeed, we find for any input density matrix ρ, Λ_1 will decohere all off-diagonal terms (and preserve all diagonal terms) in the computational basis; as such, the resulting output density matrix will be diagonal. Since the diagonal terms are preserved, different superpositions of private states lead to different outputs, meaning Eve would be able to learn some information about the original message based on what diagonal terms she observes—namely, the strength of the diagonal terms of Alice's original message (the coordinate of the Z operator in the Bloch sphere). In this way, the channel is not private.

Yet we can ask: can composing the phase damping channel on multiple qubits yield a private subspace $\mathcal{S} \subseteq \mathcal{H}$? Such a question is analogous to the sort of questions that have been asked in quantum error correction for some time; for example, given a set of errors that are uncorrectable on a single qubit, does there exist a larger Hilbert space such that the action of the error on the encoded Hilbert space is correctable? The answer to such a question in quantum error correction is yes, as demonstrated by the five-qubit code which corrects for arbitrary single-qubit errors, an error that would be uncorrectable if one did not have access to a larger Hilbert space to encode the quantum information into a quantum code.

Let Λ be the composition of the application of the maps Λ_i on each of the n qubits of the state $\rho \in \mathcal{L}(\mathcal{H})$,

$$\begin{aligned}\Lambda(\rho) &= \Lambda_n \circ \Lambda_{n-1} \circ \cdots \circ \Lambda_1(\rho) \\ &= \Lambda_n\Big(\Lambda_{n-1}\Big(\ldots \Lambda_1(\rho)\Big)\Big).\end{aligned}$$

Again, on any input state ρ, this channel will decohere all off-diagonal terms in the computational basis; as such, the resulting output density matrix will be diagonal. Equivalently one could consider the n-product map $\Lambda_1^{\otimes n}$ of the single qubit channel Λ_1.

Consider the case when $n = 2$. Since the output state of the channel Λ must be diagonal in the computational basis, the most general form for the density matrices corresponding to the output states of a private quantum channel on any subspace $\mathcal{S}$ of $\mathcal{H}$ is

$$\rho_0 = \frac{1}{4}\Big(\mathrm{I}\,\mathrm{I} + \alpha\,\mathrm{I}\,Z + \beta Z\,\mathrm{I} + \gamma ZZ\Big).$$

The goal is to find a subspace $\mathcal{S}$ of dimension 2 and a state $\rho_0 \in \mathcal{L}(\mathcal{H})$ such that $\Lambda(\rho) = \rho_0$ for all $\rho \in \mathcal{L}(\mathcal{S})$. This would show that Λ has a private qubit subspace, defined by a pair of orthogonal basis states in $\mathcal{S}$, which, because they act in an identical fashion as $|0\rangle$ and $|1\rangle$ in $\mathbb{C}^2$, they are referred to as *logical states* $|0_L\rangle, |1_L\rangle$ in $\mathcal{S}$.

However, one can show that such a subspace *does not* exist. In fact we can prove the following more general result, which gives a necessary condition on the Kraus

operators of an arbitrary random unitary channel in order for there *not* to exist a private subspace.

Theorem 2.6.10. *Let Φ be a random unitary channel with mutually commuting Kraus operators. Then Φ has no private subspaces.*

Proof. Let Φ be a random unitary channel with mutually commuting Kraus operators. Then the action of the channel on any input state $\rho \in \mathcal{L}(\mathcal{H})$ is

$$\Phi(\rho) = \sum_i p_i U_i \rho U_i^\dagger.$$

Since the set of unitaries $\{U_i\}$ commute, there exists a common eigenbasis $|e_j\rangle_{j=1}^d$ for all of the unitaries such that,

$$U_i|e_j\rangle = \alpha_{ij}|e_j\rangle \text{ with } \quad |\alpha_{ij}| = 1.$$

Suppose a non-trivial private subspace exists. Then there must exist at least two pure states $|0_L\rangle$, $|1_L\rangle$ such that $\Lambda(|0_L\rangle\langle 0_L|) = \Lambda(|1_L\rangle\langle 1_L|) = \rho_0$, where ρ_0 is some fixed density matrix. Then for some scalars $\beta_j, \gamma_j \in \mathbb{C}$, we can write

$$|0_L\rangle = \sum_{j=1}^d \beta_j |e_j\rangle, \quad |1_L\rangle = \sum_{j=1}^d \gamma_j |e_j\rangle.$$

Consider the action of the channel on these states:

$$\begin{aligned}
\Lambda(|0_L\rangle\langle 0_L|) &= \textstyle\sum_i p_i U_i \Big(\sum_{j,k=1}^d \beta_j \beta_k^* |e_j\rangle\langle e_k| \Big) U_i^\dagger \\
&= \textstyle\sum_i p_i \sum_{j,k=1}^d \alpha_{ij} \alpha_{ik}^* \beta_j \beta_k^* |e_j\rangle\langle e_k| \\
&= \textstyle\sum_{j,k=1}^d \Big(\sum_i p_i \alpha_{ij} \alpha_{ik}^* \Big) \beta_j \beta_k^* |e_j\rangle\langle e_k|
\end{aligned}$$

Similarly,

$$\Lambda(|1_L\rangle\langle 1_L|) = \sum_{j,k=1}^{d} \Big(\sum_i p_i \alpha_{ij} \alpha_{ik}^* \Big) \gamma_j \gamma_k^* |e_j\rangle\langle e_k|.$$

Comparing the diagonal terms, where $j = k$, the inside sum over i is always equal to 1 since the modulus of the eigenvalues is 1, thus the respective coefficients are $|\beta_j|^2$ and $|\gamma_j|^2$. Therefore, if the output of the channel is the same in both cases, we must have $|\beta_j| = |\gamma_j|$, $\forall j$. However, we prove below that no such $|0_L\rangle$ and $|1_L\rangle$ can form a subspace.

We proceed by contradiction: assume there is a subspace $\mathcal{S}$ of dimension greater or equal to 2 that encodes a logical qubit and a state $\rho_0 \in \mathcal{H}$ such that $\forall\ \rho \in \mathcal{L}(\mathcal{S})$, $\Lambda(\rho) = \rho_0$. Let us write $|0_L\rangle$ and $|1_L\rangle$ in terms of the basis $\{|e_j\rangle\}$ given above,

$$\begin{aligned} |0_L\rangle &= \textstyle\sum_{j=1}^{d} \beta_j |e_j\rangle = |a_1||e_1\rangle + \sum_{j=2}^{d} |a_j| e^{i\theta_j} |e_j\rangle \\ |1_L\rangle &= \textstyle\sum_{j=1}^{d} \gamma_j |e_j\rangle = |a_1||e_1\rangle + \sum_{j=2}^{d} |a_j| e^{i\phi_j} |e_j\rangle, \end{aligned}$$

where we have, without loss of generality, performed a global phase shift on the two vectors so that the coefficient of $|e_j\rangle$ is real *for both vectors* (under a global phase shift, the vectors remain orthogonal). We have relabelled the coefficients to reflect the fact that $|\beta_j| = |\gamma_j|$.

Any linear combination of the basis states must additionally be in $\mathcal{S}$ by the closure of the subspace under scalar addition. With this in mind, consider the normalized state,

$$\frac{|0_L\rangle + |1_L\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}} \Big(2|a_1||e_1\rangle + \textstyle\sum_{j=2}^{d} |a_j| (e^{i\theta_j} + e^{i\phi_j}) \Big).$$

Since such a state must be an element of $\mathcal{S}$, it must satisfy the conditions on the moduli of its coefficients; namely, the jth coefficient must be equal in modulus to $|a_j|$. However, one can clearly see that the modulus of the coefficient of the $|e_1\rangle$ term is equal to $\sqrt{2}|a_1|$ which is not equal to $|a_1|$ unless $|a_1| = 0$. Therefore, we have reduced the basis states to have the form,

$$\begin{aligned} |0_L\rangle &= |a_2||e_1\rangle + \textstyle\sum_{j=3}^{d} |a_j| e^{i\theta'_j} |e_j\rangle \\ |1_L\rangle &= |a_2||e_1\rangle + \textstyle\sum_{j=3}^{d} |a_j| e^{i\phi'_j} |e_j\rangle, \end{aligned}$$

where we have performed a global phase shift on both states and redefined the phase on the components $|e_j\rangle, 3 \leq j \leq d$. By the same argument as above, we can show that all coefficients must be equal to zero in order for the channel Λ to be private while $\mathcal{S}$ remains a subspace. As such, there does not exist two orthonormal basis vectors satisfying the requirements for the channel to be private, implying that no non-trivial subspace $\mathcal{S} \subset \mathcal{H}$ exists.

■

Corollary 2.6.11. *Let $\mathcal{H}$ be a n-qubit Hilbert space. Then there exists no subspace $\mathcal{S} \subset \mathcal{H}$ where $dim(\mathcal{S}) \geq 2$ such that the channel $\Lambda = \Lambda_n \circ \Lambda_{n-1} \circ \cdots \circ \Lambda_1$ is private on $\mathcal{S}$.*

Proof. All Kraus operators are tensor products of I_2 and Z, and are easily seen to commute. Theorem 2.6.10 therefore applies. ■

Somewhat surprisingly, we can find private subsystems for these channels, as we

show below.

Indeed, consider the following logically encoded qubits in two-qubit Hilbert space:

$$\rho_L = \frac{1}{4}(\mathrm{I}\,\mathrm{I} + \alpha XX + \beta Y\,\mathrm{I} + \gamma ZX). \tag{2.9}$$

This describes a single qubit encoding, as equation (2.9) describes the coordinates for a logical Bloch sphere in two-qubit Hilbert space with logical Pauli operators given by $X_L = XX, Y_L = Y\,\mathrm{I}, Z_L = ZX$. Now, observe that the dephasing map $\Lambda = \Lambda_2 \circ \Lambda_1$ acting on each density operator ρ_L produces an output state that is maximally mixed; that is, $\Lambda(\rho_L) = \frac{1}{4}\,\mathrm{I}\,\mathrm{I}$ for all ρ_L. Thus, we see that equation (2.9) yields a private set of two-qubit density operators for the dephasing map Λ. However, we know from theorem 2.6.10 that the input set $\{\rho_L\}$ cannot be a subspace, and we have already noted this implies it also cannot be an operator subsystem. It is however still a private subsystem in the sense of equation (2.7). Let us discuss the encoding in more detail.

We claim that this private code can be viewed as a single qubit subsystem embedded inside two qubit space, where the ancilla operator σ_A, from equation (2.7), in this case is the single qubit identity operator I_2; that is, up to a unitary equivalence the set of operators ρ_L can be seen to generate the two qubit operator algebra $\mathrm{I}_2 \otimes \mathbb{M}_2$. To see this, it is enough to show that all two-qubit states ρ of the form $\rho = \frac{1}{4}(\mathrm{I}\,\mathrm{I} + \alpha XX + \beta Y\,\mathrm{I} + \gamma ZX)$ can be sent through appropriate unitary gates to obtain ρ' of the form $\rho' = \mathrm{I}_2 \otimes \frac{1}{4}(\mathrm{I}_2 + \alpha' X + \beta' Y + \gamma' Z)$. Since I, X, Y, Z form a basis for $\mathbb{M}_2$, the claim will follow.

We find that an application of the inverse of the T-gate,

$$T = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} = \frac{1}{\sqrt{2}}(|0\rangle(\langle 0| + \langle 1|) + i|1\rangle(\langle 0| - \langle 1|)),$$

on the first qubit, and applications of $CNOT_{21} \equiv CX_{21}$ and $CNOT_{12} \equiv CX_{12}$, yields the desired transformation. Indeed, the composition $CNOT_{1,2}CNOT_{2,1}\big((T^\dagger \otimes \mathrm{I}_2)(\cdot)(T \otimes \mathrm{I}_2)\big)CNOT_{2,1}CNOT_{1,2}$ acts as

$$\begin{aligned} XX &\longmapsto \quad YX \quad \longmapsto ZY \longmapsto \mathrm{I}\,Y \\ Y\,\mathrm{I} &\longmapsto \quad Z\,\mathrm{I} \quad \longmapsto ZZ \longmapsto \mathrm{I}\,Z \\ ZX &\longmapsto \quad XX \quad \longmapsto \mathrm{I}\,X \longmapsto \mathrm{I}\,X. \end{aligned}$$

Thus, we obtain $\rho' = \frac{1}{4}(\mathrm{I}_4 + \gamma\,\mathrm{I}\,X + \alpha\,\mathrm{I}\,Y + \beta\,\mathrm{I}\,Z)$. In particular, by defining the unitary

$$U = CNOT_{1,2} \circ CNOT_{2,1} \circ (T^\dagger \otimes \mathrm{I}_2) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & -i & 0 \\ 0 & 1 & 0 & i \\ 0 & 1 & 0 & -i \\ 1 & 0 & i & 0 \end{pmatrix},$$

we see the set of operators $U\rho_L U^\dagger$ generate the algebra $\mathrm{I}_2 \otimes \mathbb{M}_2$.

This subsystem encoding fits into the framework of the definition of private quantum subsystem; namely, $\mathcal{S}$ is a private subsystem for the channel Λ. We show below how this example does not fit into the previously considered notion of operator private subsystems.

Now, we have shown explicitly that $\mathcal{L}(\mathcal{S})$ is isomorphic to $\mathrm{I}_2 \otimes \mathbb{M}_2$, where $\mathcal{S}$ is the set of all private states for our channel Λ, via

$$CNOT_{1,2}CNOT_{2,1}\big((T^\dagger \otimes \mathrm{I}_2)(\mathcal{L}(\mathcal{S}))(T \otimes \mathrm{I}_2)\big)CNOT_{2,1}CNOT_{1,2} = \mathrm{I}_2 \otimes \mathbb{M}_2.$$

Thus, if we wish to say something akin to "Λ is private for $\mathrm{I}_2 \otimes \mathbb{M}_2$", we must be careful; it is in fact the new channel

$$\Lambda' := CNOT_{1,2}CNOT_{2,1}\big((T^\dagger \otimes \mathrm{I}_2)(\Lambda)(T \otimes \mathrm{I}_2)\big)CNOT_{2,1}CNOT_{1,2}$$

defined by applying

$$CNOT_{1,2}CNOT_{2,1}\big((T^\dagger \otimes \mathrm{I}_2)(\cdot)(T \otimes \mathrm{I}_2)\big)CNOT_{2,1}CNOT_{1,2}$$

to each of the four Kraus operators of Λ, that is private for $\mathrm{I}_2 \otimes \mathbb{M}_2$. We find the new Kraus operators to be $\Lambda' = \{\frac{1}{2}\,\mathrm{II}, \frac{1}{2}ZZ, \frac{1}{2}XX, -\frac{1}{2}YY\}$.

For any

$$\sigma_A = \begin{pmatrix} a_A & b_A \\ c_A & d_A \end{pmatrix} \in \mathcal{L}(\mathcal{H}_A) \quad \sigma_B = \begin{pmatrix} a_B & b_B \\ c_B & d_B \end{pmatrix} \in \mathcal{L}(\mathcal{H}_B),$$

we compute

$$\Lambda'(\sigma_A \otimes \sigma_B) = \frac{1}{2}\begin{pmatrix} a_Aa_B + d_Ad_B & 0 & 0 & b_Ab_B + c_Ac_B \\ 0 & a_Aa_B + d_Ad_B & b_Ac_B + c_Ab_B & 0 \\ 0 & b_Ac_B + c_Ab_B & a_Aa_B + d_Ad_B & 0 \\ b_Ab_B + c_Ac_B & 0 & 0 & a_Aa_B + d_Ad_B \end{pmatrix}. \tag{2.10}$$

We note that this output is symmetric in the subsystems $\mathcal{H}^A$ and $\mathcal{H}^B$. In particular, $\Lambda'(\sigma_A \otimes \frac{1}{2}\,\mathrm{I}_2) = \frac{1}{4}\,\mathrm{diag}(a_A+d_A, a_A+d_A, a_A+d_A, a_A+d_A)$ and $\Lambda'(\frac{1}{2}\,\mathrm{I}_2 \otimes \sigma_B) = \frac{1}{4}\,\mathrm{diag}(a_B+d_B, a_B+d_B, a_B+d_B, a_B+d_B)$. In order for σ_A (resp., σ_B) to be a valid density matrix,

its trace must be 1. Thus $\Lambda'(\sigma_A \otimes \frac{1}{2}\,\mathrm{I}_2) = \Lambda'(\frac{1}{2}\,\mathrm{I}_2 \otimes \sigma_B) = \frac{1}{4}\,\mathrm{I}_4$. So Λ' is in fact private for either subsystems $\mathrm{I}_2 \otimes \mathbb{M}_2$ and $\mathbb{M}_2 \otimes \mathrm{I}_2$.

Using the operator private subsystem definition, as well as the complementarity result between QECC and PQCs, one would expect our channel Λ' to split up into two distinct actions; i.e. the output can be expresses as the tensor product of two channels, acting on systems $\mathcal{H}^A$ and $\mathcal{H}^B$, respectively. Thus, we ask whether there exists density matrices τ_A, τ_B such that $\Lambda'(\sigma_A \otimes \sigma_B) = \tau_A \otimes \tau_B$. Equating this equation with equation (2.10), we find the system has no solution for general σ_A, σ_B (equating components forces τ_B to be the all-zeros matrix, which then forces $\Lambda'(\sigma_A \otimes \sigma_B)$ to be the all-zeros matrix). As a specific example, let $\sigma_{A_0} = \frac{1}{2}\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$ and let σ_B be arbitrary. Then

$$\Lambda'(\sigma_{A_0} \otimes \sigma_B) = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & 2\,\mathrm{Im}(b_B) \\ 0 & 1 & -2\,\mathrm{Im}(b_B) & 0 \\ 0 & -2\,\mathrm{Im}(b_B) & 1 & 0 \\ 2\,\mathrm{Im}(b_B) & 0 & 0 & 1 \end{pmatrix},$$

where $2\,\mathrm{Im}(b_B)$ is 2 times the imaginary part of b_B. Thus, $\Lambda'(\sigma_{A_0} \otimes \sigma_B)$ can be written as $\tau_A \otimes \tau_B$ if and only if the entry b_B is real. The private subsystem $\mathrm{I}_2 \otimes \mathbb{M}_2$ therefore *does not* extend to a private subspace, as in the operator private subsystem setting. The channel Λ' (and hence, Λ) is private for a subsystem that *does not* fit the original definition of an operator private subsystem.

More generally, a logical qubit encoding into a subsystem of a n-qubit Hilbert space can be constructed to privatize the n-qubit phase damping channel

$\Lambda = \Lambda_n \circ \cdots \circ \Lambda_2 \circ \Lambda_1$, which by theorem 2.6.10 cannot have a private subspace. Hence we have the following result:

Proposition 2.6.12. *For any n-qubit Hilbert space $\mathcal{H}$, there exist quantum channels Φ for which a private quantum subsystem $\mathcal{H}^B$ of $\mathcal{H}$ can be constructed in the absence of the existence of any private quantum subspace $\mathcal{S} \subseteq \mathcal{H}$.*

2.7 Testable Conditions For Private Quantum Codes

If we are given a quantum channel $\Phi(\rho) = \sum_i V_i \rho V_i^\dagger$ and a subsystem $\mathcal{H}^B$, we can ask if it is possible to decide whether $\mathcal{H}^B$ is private for Φ; and more to the point, we can ask if this can be answered in terms of the Kraus operators V_i for the channel. These conditions can be thought of as Knill-Laflamme conditions in that they are a set of algebraic constraints giving necessary and sufficient conditions for the existence of private quantum subsystems.

The following result answers this question for private quantum subsystems. In addition to Kraus operators, we would expect the algebra to include the fixed $\mathcal{H}^A$ state σ_A and output state ρ_0—observe that this information is indeed included in the conditions.

Theorem 2.7.1. *A subsystem $\mathcal{H}^B$ is private for a channel $\Phi(\rho) = \sum_i V_i \rho V_i^\dagger$ with fixed state $\sigma_A \in \mathcal{L}(\mathcal{H}^A)$ and output state $\rho_0 \in \mathcal{L}(\mathcal{H})$ if and only if there are complex scalars $\lambda_{ijk\ell}$ forming an isometry matrix $\lambda = (\lambda_{ijk\ell})$ such that $\sqrt{p_k} V_j |\psi_{A,k}\rangle =$*

$\sum_{i,\ell} \lambda_{ijk\ell}\sqrt{q_\ell}|\phi_\ell\rangle\langle\psi_{B,i}|$, *where* $|\psi_{A,k}\rangle$ *(*p_k*) and* $|\phi_\ell\rangle$ *(*q_ℓ*) are eigenstates (eigenvalues) of* σ_A *and* ρ_0 *respectively,* $|\psi_{B,i}\rangle$ *is an orthonormal basis for* $\mathcal{H}^B$*, and where* $|\psi_{A,k}\rangle$ *is viewed as a channel from* $\mathcal{H}^B$ *into* $\mathcal{H}$*.*

The key observation in establishing this result is that the left and right hand sides of equation (2.7) each define channels from $\mathcal{H}^B$ to $\mathcal{H}$ which are in fact the same. One can then use basic results from the theory of completely positive maps to obtain the equations spelled out in the theorem.

Note: We regard $|\psi_{A,k}\rangle$ as operators from the $\mathcal{H}^A$ subsystem to the full tensor product $\mathcal{H}^A \otimes \mathcal{H}^B$ subspace. That is, $|\psi_{A,k}\rangle : |\psi_B\rangle \mapsto |\psi_{A,k}\rangle \otimes |\psi_B\rangle$ for all states $|\psi_B\rangle$ in the subsystem $\mathcal{H}^B$. In this way, $V_j|\psi_{A,k}\rangle$ is well-defined.

Proof. Consider first the left-hand side of the equation (2.7) of the definition of private quantum subsystem. Let $\Phi : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ be a quantum channel satisfying this definition. Let $\{V_j\}$ be the Kraus operators of Φ. Since Hermitian matrices are diagonalizable, without loss of generality we can write $\sigma_A = \sum_k p_k|\psi_{A,k}\rangle\langle\psi_{A,k}|$, where $|\psi_{A,k}\rangle$ and p_k are the eigenstates and eigenvalues, respectively, of σ_A. We can consider the action of Φ on $\mathcal{L}(\mathcal{H})$ as the composition of maps $\Phi \circ \Psi(\sigma_B)$, where, for fixed σ_A, $\Psi : \mathcal{L}(\mathcal{H}^B) \to \mathcal{L}(\mathcal{H})$ is the map $\sigma_B \mapsto \sigma_A \otimes \sigma_B$. The Kraus operators of Ψ are $\{\sqrt{p_k}|\psi_{A,k}\rangle\}$, where again we note that we are viewing $|\psi_{A,k}\rangle$ as an operator. It follows that the Kraus operators of the composition $\Phi \circ \Psi(\sigma_B)$ are $\{\sqrt{p_k}V_j|\psi_{A,k}\rangle\}$.

The right-hand side of equation (2.7) can be viewed as a quantum channel

$$\sigma_B \mapsto \mathrm{Tr}(\sigma_B) \sum_\ell q_\ell |\phi_\ell\rangle\langle\phi_\ell| = \sum_{i,\ell} q_\ell |\phi_\ell\rangle\langle\psi_{B,i}|\sigma_B|\psi_{B,i}\rangle\langle\phi_\ell|,$$

where $\{|\psi_{B,i}\rangle\}$ is an orthonormal basis for the subsystem B, and we have used the fact that we can diagonalize ρ_0, with $|\phi_\ell\rangle$ and q_ℓ its eigenstates and eigenvalues, respectively. The Kraus operators of this map are $\{\sqrt{q_\ell}|\phi_\ell\rangle\langle\psi_{B,i}|\}$.

The quantum channels described by the left- and right-hand sides of equation (2.7) are equal in that, given an arbitrary input σ_B, their outputs are equal. It follows immediately from remark 1.2.4 that $\sqrt{p_k}V_j|\psi_{A,k}\rangle = \sum_{i,\ell} \lambda_{ijk\ell}\sqrt{q_\ell}|\phi_\ell\rangle\langle\psi_{B,i}|$, for some isometry (or, appropriately, unitary) λ, as desired. ■

It is important to note that this result is new even for private sub*spaces*. In the notation of the theorem for that case, $\mathcal{H}^A$ is one-dimensional and $\mathcal{H}^B$ is the subspace. If we let P_B be the projector of $\mathcal{H}$ onto $\mathcal{H}^B$, then we see that the characterization of privacy is given by the conditions: $V_j P_B = \sum_{i,\ell} \lambda_{ij\ell}\sqrt{q_\ell}|\phi_\ell\rangle\langle\psi_{B,i}|$ for all j. Taking the dual of both sides, this is equivalent to corollary 2.6.5. As a simple illustration, in the case of the completely depolarizing channel on N-dimensional Hilbert space, P_B is the identity operator and these conditions reduce to the Kraus operators satisfying $\sqrt{N}\,V_j = \sum_{i_1,i_2} \lambda_{i_1 i_2 j}|i_1\rangle\langle i_2|$, for some choice of orthonormal bases $\{|i_1\rangle\}$ and $\{|i_2\rangle\}$ and unitary matrix $(\lambda_{i_1 i_2 j})_{i_1,i_2}$.

Here we point out how the 2-qubit phase damping channel Λ can be assembled from this result. We must use the Kraus operators $\{V_j\} = \{\frac{1}{2}\,\mathrm{I}\,\mathrm{I}, \frac{1}{2}XX, \frac{1}{2}ZZ, -\frac{1}{2}YY\}$

of Λ'. The eigenstates of $\rho_0 = \frac{1}{4}\,\mathrm{I}_4$ are $|00\rangle, |01\rangle, |10\rangle, |11\rangle$, each having eigenvalue $\frac{1}{4}$. For simplicity, we will use the standard orthonormal basis on the subsystem $\mathcal{H}^B$: $\{\langle\psi_{B,i}|\} = \{\langle 0|, \langle 1|\}$. In our example, $\sigma_A = \frac{1}{2}\,\mathrm{I}_2$, hence its eigenstates are $\{|\psi_{A,k}\rangle\} = \{|0\rangle, |1\rangle\}$, with corresponding eigenvalues $\frac{1}{2}$.

We compute $V_j|\psi_{A,k}\rangle$ as follows:

$$
\begin{aligned}
V_1|\psi_{A,1}\rangle &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}\left(|00\rangle\langle 0| + |01\rangle\langle 1|\right) \\
V_1|\psi_{A,2}\rangle &= \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}\left(|10\rangle\langle 0| + |11\rangle\langle 1|\right) \\
V_2|\psi_{A,1}\rangle &= \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} = \frac{1}{2}\left(|10\rangle\langle 1| + |11\rangle\langle 0|\right) \\
V_2|\psi_{A,2}\rangle &= \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}\left(|00\rangle\langle 1| + |01\rangle\langle 0|\right) \\
V_3|\psi_{A,1}\rangle &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}\left(|00\rangle\langle 0| - |01\rangle\langle 1|\right) \\
V_3|\psi_{A,2}\rangle &= \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}\left(-|10\rangle\langle 0| + |11\rangle\langle 1|\right)
\end{aligned}
$$

$$V_4|\psi_{A,1}\rangle = \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & -1 \\ 1 & 0 \end{pmatrix} = \frac{1}{2}\left(-|10\rangle\langle 1| + |11\rangle\langle 0|\right)$$

$$V_4|\psi_{A,2}\rangle = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}\left(|00\rangle\langle 1| + |01\rangle\langle 0|\right).$$

Note that the V_j are 4×2 matrices formed with 2×2 Pauli operators and zero blocks.

Note that $\sqrt{p_k} = \frac{1}{\sqrt{2}}$ for all k, and each V_j has a factor of $\frac{1}{2}$, so the coefficient of $\sqrt{p_k}V_j|\psi_{A,k}\rangle$ is always $\frac{1}{2\sqrt{2}}$. The coefficient of $\lambda_{ijk\ell}\sqrt{q_\ell}|\phi_\ell\rangle\langle\psi_{B,i}|$ is $\lambda_{ijk\ell}\sqrt{q_\ell} = \frac{1}{\sqrt{2}} \cdot \frac{1}{2}$ for all i, j, k, ℓ.

Thus in our example, we find that λ is the following matrix

$$\lambda = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The scalar-valued matrix $\lambda = (\lambda_{ijk\ell})$ is indeed an isometry. Furthermore, because the number of operators $V_j|\psi_{A,k}\rangle$ agrees with the number of operators $|\phi_\ell\rangle\langle\psi_{B,i}|$ (namely, 8), the matrix λ is in fact unitary.

Does our result translate into a quantum error correcting code through the use of the complementarity between private quantum codes and quantum error correction?

For our phase damping channel Λ, we can easily compute the Kraus operators

of the complementary channel $\Lambda^\sharp$ by "stacking" the j-th column of each of the eight Kraus operators V_i of Λ one below the next, to obtain the j-th Kraus operator of $\Lambda^\sharp$:

$$\begin{aligned} A_1 &= \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad A_2 = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \\ A_3 &= \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix} \quad A_4 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

By the bridge between QEC and quantum cryptography, if we had an operator private subsystem $\mathcal{S}$ for Λ, we should have a correctable subsystem code for $\Lambda^\sharp$. Again, we have to be careful: the subsystem $\mathrm{I}\otimes\mathbb{M}_2$ is private for Λ', so the subsystem $\mathrm{I}\otimes\mathbb{M}_2$ should be correctable for $(\Lambda^\sharp)'$, where we obtain $(\Lambda^\sharp)'$ by applying the unitary transformation $U(\cdot)U^\dagger$, where $U = CNOT_{1,2}CNOT_{2,1}(T^\dagger \otimes \mathrm{I}_2)$, as before.

We compute the Kraus operators of $(\Lambda^\sharp)'$ to be $\{B_i = UA_iU^\dagger\}$, where

$$\begin{aligned} B_1 &= \frac{1}{4}\begin{pmatrix} 1-i & 0 & 0 & 1-i \\ 1+i & 0 & 0 & 1+i \\ 1-i & 0 & 0 & 1-i \\ 1+i & 0 & 0 & 1+i \end{pmatrix} \quad B_2 = \frac{1}{4}\begin{pmatrix} 0 & 1-i & 1-i & 0 \\ 0 & -1-i & -1-i & 0 \\ 0 & -1+i & -1+i & 0 \\ 0 & 1+i & 1+i & 0 \end{pmatrix} \\ B_3 &= \frac{1}{4}\begin{pmatrix} -1+i & 0 & 0 & 1-i \\ 1+i & 0 & 0 & -1-i \\ -1+i & 0 & 0 & 1-i \\ 1+i & 0 & 0 & -1-i \end{pmatrix} \quad B_4 = \frac{1}{4}\begin{pmatrix} 0 & 1-i & 1+i & 0 \\ 0 & 1+i & -1-i & 0 \\ 0 & -1+i & 1-i & 0 \\ 0 & -1-i & 1+i & 0 \end{pmatrix}. \end{aligned}$$

Now, for any $\sigma_B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{L}(\mathcal{H}_B)$, we find

$$(\Lambda^\sharp)'(\frac{1}{2}\,\mathrm{I}_A \otimes \sigma_B) = \sum_i B_i(\frac{1}{2}\,\mathrm{I}_A \otimes \sigma_B)B_i^\dagger = \frac{1}{4}\,\mathrm{I}_4\,.$$

Instead of being correctable on the private subsystem $\mathrm{I}_2 \otimes \mathbb{M}_2$, the complementary channel $(\Lambda^\sharp)'$ (with the proper unitary transformation) is completely depolarizing! All information is lost, so there is no possibility of the channel being quantum error correctable. Thus, not only does the complementarity result fail, it fails dramatically.

Note that the Kraus operators of the complementary map $\Lambda^\sharp$ are four orthogonal rank-one projectors in two-qubit Hilbert space, and in particular the map determines a von Neumann measurement. No error-correcting subsystem can be extracted in such a setting.

2.7.1 Quantum Error Correction Revisited

The following definition makes precise the notion of error correctable subsystem.

Definition 2.7.2. *Let $\mathcal{H} = (\mathcal{H}^A \otimes \mathcal{H}^B) \oplus (\mathcal{H}^A \otimes \mathcal{H}^B)^\perp$ and let $\mathcal{E}$ be a channel acting on $\mathcal{L}(\mathcal{H})$. Then $\mathcal{H}^B$ is an* operator quantum error correcting code (OQECC) *for $\mathcal{E}$ if there exists a quantum channel $\mathcal{R}$ such that for all σ_A, for all σ_B, there exists some fixed state $\tau_A = \tau_A(\sigma_A)$ (dependent on σ_A) such that*

$$\mathcal{R} \circ \mathcal{E}(\sigma_A \otimes \sigma_B) = \tau_A \otimes \sigma_B.$$

The discussion of private subsystem versus operator private subsystem motivates the following observation: The notion of an operator quantum error-correcting subsystem can be expanded to mimic the general definition of a private quantum subsystem. Indeed, a revised definition analogous to that of equation (2.7) could be proposed as follows:

Definition 2.7.3. *Let* $\mathcal{H} = (\mathcal{H}^A \otimes \mathcal{H}^B) \oplus (\mathcal{H}^A \otimes \mathcal{H}^B)^\perp$ *and let* $\mathcal{E}$ *be a channel acting on* $\mathcal{L}(\mathcal{H})$. *Then* $\mathcal{H}^B$ *is a* generalized operator quantum error correcting code (GenOQECC) *for* $\mathcal{E}$ *if there exists a quantum channel* $\mathcal{R}$ *such that there exists a fixed state* σ_A *and a state* $\tau_A = \tau_A(\sigma_A)$ *(dependent on* σ_A*) such that for all* σ_B, *we have*

$$\mathcal{R} \circ \mathcal{E}(\sigma_A \otimes \sigma_B) = \tau_A \otimes \sigma_B.$$

It is clear that the existence of an operator quantum error correcting code implies the existence of a generalized operator quantum error correcting code for any fixed σ_A. The following theorem states that if $\mathcal{H}^B$ is a generalized operator quantum error correcting code for a channel $\mathcal{E}$ as in definition 2.7.3, then $\mathcal{H}^B$ is an operator quantum error correcting code for $\mathcal{E}$ with $\mathcal{H}^{A'}$ a subspace of $\mathcal{H}^A$.

Theorem 2.7.4. *Given a fixed decomposition* $\mathcal{H} = (\mathcal{H}^A \otimes \mathcal{H}^B) \oplus (\mathcal{H}^A \otimes \mathcal{H}^B)^\perp$ *and a map* $\mathcal{E}$ *on* $\mathcal{L}(\mathcal{H})$, *assume there exists density operators* σ_A *and* $\tau_A = \tau_A(\sigma_A)$ *(dependent on* σ_A*) such that* $\forall \sigma_B$,

$$\mathcal{R} \circ \mathcal{E}(\sigma_A \otimes \sigma_B) = \tau_A \otimes \sigma_B.$$

Then there exists a Hilbert space $\mathcal{H}^{A'}$, *which is a subspace of* $\mathcal{H}^A$, *such that*

$$\forall \sigma_{A'} \, \forall \sigma_B, \ \exists \tau_{A'} \ : \ \mathcal{R} \circ \mathcal{E}(\sigma_{A'} \otimes \sigma_B) = \tau_{A'} \otimes \sigma_B.$$

In words, if a subsystem $\mathcal{H}^B$ is correctable for a channel $\mathcal{E}$ for a particular $\sigma_A \in \mathcal{L}(\mathcal{H}^A)$, then $\mathcal{H}^B$ is correctable for a channel $\mathcal{E}$ for any $\sigma_{A'} \in \mathcal{L}(\mathcal{H}^{A'}) \subseteq \mathcal{L}(\mathcal{H}^A)$. We shall see in the proof of this theorem that $\mathcal{H}^{A'}$ is in fact the span of the eigenspaces of σ_A.

Proof. First let $|\psi\rangle \in \mathcal{H}_B$ and $P = |\psi\rangle\langle\psi|$. Let $\{|\alpha_k\rangle\}$ be the normalized eigenvectors of σ_A so that $\sigma_A = \sum_{k=1}^{m} p_k |\alpha_k\rangle\langle\alpha_k|$ where $0 < p_k \leq 1$. By assumption and using the positivity of $\mathcal{E}$ we have for all k,

$$\begin{aligned} 0 &\leq \mathcal{E}(p_k|\alpha_k\rangle\langle\alpha_k| \otimes P) = p_k \mathcal{E}(|\alpha_k\rangle\langle\alpha_k| \otimes P) \leq \mathcal{E}(\sigma_A \otimes P) \\ &= \tau_A \otimes P \\ &= (\mathrm{I}_A \otimes P)(\tau_A \otimes P)(\mathrm{I}_A \otimes P). \end{aligned}$$

It follows that there are positive operators $\sigma_{\psi,k} \in \mathcal{L}(\mathcal{H}_A)$ such that $\mathcal{E}(p_k|\alpha_k\rangle\langle\alpha_k| \otimes P) = \sigma_{\psi,k} \otimes P$ for all k. Trace-preservation forces $\operatorname{Tr} \sigma_{\psi,k} = p_k$, thus dividing by p_k, we find $\mathcal{E}(|\alpha_k\rangle\langle\alpha_k| \otimes P) = \sigma_{\psi,k} \otimes P$ for all k, where $\sigma_{\psi,k}$ are now *valid density operators*.

The density operators τ_A may be in the $\mathcal{H}^A$ subsystem, and so the recovery operation is needed in order to map $\tau_A \mapsto \tau_{A'}$, where $\tau_{A'}$ lives in the $\mathcal{H}^{A'}$ subsystem.

The following argument relates solely to the subsystem $\mathcal{H}^B$, so is identical to the proof of implication $2 \Rightarrow 1$ of lemma 3.3 in [KLPL05]. We reproduce the argument here for completeness.

In fact, the operators $\sigma_{\psi,k}$ do not depend on $|\psi\rangle$. To verify this claim, for clarity we shall suppose that $\dim \mathcal{H}^B = 2$. The case of general $\mathcal{H}^B$ easily follows. So let $|\psi_i\rangle$, $i = 1, 2$, be an orthonormal basis for $\mathcal{H}^B$. Let $P_i = |\psi_i\rangle\langle\psi_i|$, $i = 1, 2$, and put $P_\pm = |\pm\rangle\langle\pm|$ where $|\pm\rangle = \frac{1}{\sqrt{2}}(|\psi_1\rangle \pm |\psi_2\rangle)$. Fix $\alpha = \alpha_k$. By the above argument, there are operators $\sigma_{\pm,\alpha}$ and $\sigma_{i,\alpha}$ on $\mathcal{H}^A$ such that

$$\mathcal{E}(|\alpha\rangle\langle\alpha| \otimes P_\pm) = \sigma_{\pm,\alpha} \otimes P_\pm \text{ and } \mathcal{E}(|\alpha\rangle\langle\alpha| \otimes P_i) = \sigma_{i,\alpha} \otimes P_i.$$

In particular, as $\mathrm{I}_B = P_+ + P_- = P_1 + P_2$, we have

$$\begin{aligned}\mathcal{E}(|\alpha\rangle\langle\alpha| \otimes \mathrm{I}_B) &= \sigma_{1,\alpha} \otimes P_1 + \sigma_{2,\alpha} \otimes P_2 \\ &= \sigma_{+,\alpha} \otimes P_+ + \sigma_{-,\alpha} \otimes P_-.\end{aligned}$$

If we compress this equation by the projection $\mathrm{I}_A \otimes P_1$, we obtain

$$\begin{aligned}(\mathrm{I}_A \otimes P_1)\mathcal{E}(|\alpha\rangle\langle\alpha| \otimes \mathrm{I}_B)(\mathrm{I}_A \otimes P_1) &= \sigma_{1,\alpha} \otimes P_1 \\ &= \frac{1}{2}(\sigma_{+,\alpha} + \sigma_{-,\alpha}) \otimes P_1.\end{aligned}$$

Thus, $\sigma_{1,\alpha} = \frac{1}{2}(\sigma_{+,\alpha} + \sigma_{-,\alpha})$ and since the same identity holds for $\sigma_{2,\alpha}$ when we compress by $\mathrm{I}_A \otimes P_2$, we obtain $\sigma_{1,\alpha} = \sigma_{2,\alpha}$. As $|\alpha\rangle$ and $|\psi_i\rangle$, $i = 1, 2$, were chosen arbitrarily, the claim holds.

Let $\mathcal{H}^{A'} = \mathrm{span}\{|\alpha_k\rangle\}$. By linearity of $\mathcal{E}$, we conclude

$$\forall \sigma_{A'} \in \mathcal{L}(\mathcal{H}_{A'}),\ \forall \sigma_B,\ \exists \tau_{A'}\ :\ \mathcal{R} \circ \mathcal{E}(\sigma_{A'} \otimes \sigma_B) = \tau_{A'} \otimes \sigma_B.$$

■

It follows from the proof of the theorem that if the fixed state σ_A in definition 2.7.3 is a pure state, then $\mathcal{H}^{A'} \cong \mathbb{C}$ and $\mathcal{H}^B$ is therefore a correctable subspace; if σ_A is full-rank, then $\mathcal{H}^{A'} = \mathcal{H}^A$. The case when σ_A is a mixed state that is not full rank is therefore the case when GenOQECC is of interest.

Chapter 3

Ordering Quantum Probability Measures

Motivated by classical measure theory, where the natural partial order on probability measures is that of absolute continuity, we consider the notion of absolute continuity with regards to quantum probability measures (see definition 3.1.7). Absolute continuity in the classical setting is characterized by the Radon-Nikodým theorem; analogously, we characterize the quantum version of absolute continuity with a quantum Radon-Nikodým theorem. This result is theorem 3.2.9, and gives rise to a non-principal, or quantum, Radon-Nikodým derivative.

In the quantum setting, partial orders other than that of absolute continuity arise. Motivated by the desire to characterize what it means to be "less than", we next investigate the partial order arising from the notion of *cleanness* for positive

operator valued measures, which was originally introduced and studied by Buscemi *et al* [BKD+05] in connection with the addition of a preprocessing step to the measurement of a quantum system. A related work of Heinonen [Hei05], which appeared at roughly the same time as [BKD+05], addressed the very general issue of optimality in quantum measurements, and analysed several additional partial orders on POVMs other than that of cleanness. Subsequent studies of clean quantum measurements were undertaken by Kahn [Kah07] and Pellonpää [Pel11]. Pellonpää in fact uses a slightly weaker definition than the original definition of clean quantum measurement put forward in [BKD+05]. To explain this, we first recall some of the standard nomenclature in quantum mechanics.

If a quantum system is represented by a Hilbert space $\mathcal{H}$, then a measurement of the system is represented by a positive operator valued probability measure $\nu : \mathcal{O}(X) \to \mathfrak{B}(\mathcal{H})$, where X represents a set of measurement outcomes for the system and $\mathcal{O}(X)$ is a σ-algebra of measurement events. The corresponding measurement statistics, which capture the probability that event $E \in \mathcal{O}(X)$ is measured by the apparatus ν when the system $\mathcal{H}$ is in state ρ, are given by the real numbers $\mathrm{Tr}_{\mathcal{H}}(\rho\nu(E))$.

To describe the idea introduced in [BKD+05], assume that ν is a measurement of a quantum system $\mathcal{H}$. Suppose that a preprocessing step is introduced through the use an irreversible quantum channel Φ that maps the states of $\mathcal{H}$ to states in some other system $\mathcal{H}'$ which is to be measured by ν'. This preprocessing step is represented mathematically in the Heisenberg picture by $\nu = \Phi^\dagger \circ \nu'$, where $\Phi^\dagger$ is the

dual of Φ, and it transforms the measurement statistics according to the equation $\mathrm{Tr}_{\mathcal{H}}(\rho\,\nu(E)) = \mathrm{Tr}_{\mathcal{H}'}(\Phi(\rho)\nu'(E))$. The measurement apparatus ν' is thought to be cleaner than ν because of the quantum noise introduced by the quantum channel Φ. Put differently, ν is obtained from a cleaner quantum probability measure ν' by irreversible preprocessing, a relation that is denoted by $\nu \ll_{\mathrm{cl}} \nu'$. A *clean quantum measurement* ν is one in which there is no ν' from which ν is obtained irreversibly by preprocessing.

The main results of this chapter are theorem 3.3.2, which gives an analytic description of the order relation $\nu \ll_{\mathrm{cl}} \nu'$ for quantum probability measures, and theorem 3.4.3, which determines the structure of clean quantum probability measures.

Because we are using Pellonpää's definition of clean quantum probability measure, which is more stringent than the definition used in [BKD+05] or [Kah07], our main theorem (theorem 3.4.3) neither implies nor is implied by the results of [BKD+05] or [Kah07].

Assumption: Throughout this chapter every Hilbert space is assumed to have finite dimension. This assumption has a profound effect on our results; we refer the reader to [FFP13] for the infinite-dimensional generalization of the main results herein.

3.1 Quantum Probability Measures (POVMs) and Measurement Spaces

Throughout X shall denote a nonempty set and $\mathcal{O}(X)$ will denote a σ-algebra of subsets of X. In the language of probability, X is a sample space of measurement outcomes and $\mathcal{O}(X)$ is a σ-algebra of measurement events. If X is a locally compact Hausdorff space, then $\mathcal{O}(X)$ is assumed to be the σ-algebra of Borel sets of X. In particular, if X is a finite set (endowed the discrete topology), then $\mathcal{O}(X)$ is assumed to be the power set of X.

Definition 3.1.1. *A map $\nu : \mathcal{O}(X) \to \mathfrak{B}(\mathcal{H})$ is a* positive operator valued probability measure (POVM)*, or a* quantum probability measure*, if*

(i) $\nu(E) \in \mathfrak{B}(\mathcal{H})_+$ *for every* $E \in \mathcal{O}(X)$*,*

(ii) $\nu(X) = \mathrm{I} \in \mathfrak{B}(\mathcal{H})$*, and*

(iii) for every countable collection $\{E_k\}_{k\in\mathbb{N}} \subseteq \mathcal{O}(X)$ *with* $E_j \cap E_k = \emptyset$ *for* $j \neq k$ *we have*

$$\nu\left(\bigcup_{k\in\mathbb{N}} E_k\right) = \sum_{k\in\mathbb{N}} \nu(E_k)\,.$$

In addition:

(iv) if $\nu(E \cap F) = \nu(E)\nu(F)$ *for all* $E, F \in \mathcal{O}(X)$*, then* ν *is a* projective quantum measure*.*

With regards to convergence of the sum $\sum_{k\in\mathbb{N}} \nu(E_k)$, we mean that $\sum_{k\in\mathbb{N}} \nu(E_k)$ is the unique operator A on $\mathcal{H}$ for which

$$\lim_{n\to\infty} \left\| A|\xi\rangle - \sum_{k=1}^{n} \nu(E_k)|\xi\rangle \right\| = 0$$

for every $|\xi\rangle \in \mathcal{H}$.

Definition 3.1.2. *The* support *of a quantum probability measure ν is the smallest closed subset $K_\nu \subset X$ for which $\nu(X \setminus K_\nu) = 0$.*

If the support of ν consists of a single point, say $K_\nu = \{x_0\}$, then ν is a *Dirac measure* and is necessarily of the form $\nu = \delta_{x_0}\,\mathrm{I}$, where δ_{x_0} is a (classical) probability measure on X satisfying, for $E \in \mathcal{O}(X)$, $\delta_{x_0}(E) = 1$ if $x_0 \in E$ and $\delta_{x_0}(E) = 0$ otherwise.

If ν has finite support $K_\nu = \{x_1, \ldots, x_m\}$, then each $H_j = \nu(\{x_j\}) \neq 0$, $\sum_{j=1}^{m} H_j = \mathrm{I}$, and

$$\nu = \sum_{j=1}^{m} \delta_{x_j} H_j\,.$$

One new idea introduced in this thesis (and appearing in [FFP13]) is to study POVMs by means of linear spaces of operators associated to them.

Definition 3.1.3. *If ν is a quantum probability measure on $(X, \mathcal{O}(X))$, then the* range *of ν is the set*

$$\mathcal{R}_\nu = \{\nu(E) : E \in \mathcal{O}(X)\} \subset \mathfrak{B}(\mathcal{H})_+,$$

and the measurement space of ν *is the vector space*

$$\mathcal{T}_\nu = \mathrm{Span}_{\mathbb{C}}\, \mathcal{R}_\nu \subset \mathfrak{B}(\mathcal{H})\,.$$

We shall make extensive use of the fact that $\mathcal{T}_\nu$ is an operator system (see definition 1.2.15).

Measurement bases first appeared in [FFP13] as a powerful tool when working with quantum probability measures, and seem to be of independent interest. The following proposition is a finite-dimensional version of [FFP13, Prop. II.4].

Definition 3.1.4. *A* measurement basis *for a quantum probability measure ν is a finite or countably infinite set $\mathcal{B}_\nu$ of positive operators such that*

(i) $\mathcal{B}_\nu = \{\nu(E) : E \in \mathcal{F}_\nu\}$ *for some finite or countable family $\mathcal{F}_\nu \subset \mathcal{O}(X)$ of pairwise disjoint sets; and*

(ii) *the sets $A_j = \nu(E_j)$ $j = 1, \ldots, m$ form a basis for $\mathcal{T}_\nu$.*

If $E_0 = X \setminus \left(\bigcup_{E\in\mathcal{F}_\nu} E\right)$, then the operator $A_0 = \nu(E_0)$ is called the basis residual *for $\mathcal{B}_\nu$; if $A_0 = 0$, then $\mathcal{B}_\nu$ is said to admit a* trivial basis residual.

Note that $\mathrm{I} = A_0 + \sum_{A\in\mathcal{B}_\nu} A$, if $\mathcal{B}_\nu$ is measurement basis for ν.

Example 3.1.5. *There exist quantum probability measures having non-trivial basis residual.*

Proof. Let $X = \{x_1, x_2, x_3\}$ and $\nu : \mathcal{O}(X) \to \mathbb{M}_2(\mathbb{C}) \equiv \mathfrak{B}(\mathbb{C}^2)$ be defined by $\nu(\emptyset) = 0$, $\nu(X) = \mathrm{I}$, and

$$\begin{aligned}
\nu(\{x_1\}) &= \begin{pmatrix} 1/4 & 1/16 \\ 1/16 & 1/4 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 & 1/4 \\ 1/4 & 1 \end{pmatrix} \\
\nu(\{x_2\}) &= \begin{pmatrix} 1/6 & -1/24 \\ -1/24 & 1/6 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 1 & -1/4 \\ -1/4 & 1 \end{pmatrix} \\
\nu(\{x_3\}) &= \begin{pmatrix} 7/12 & -1/48 \\ -1/48 & 7/12 \end{pmatrix} = \frac{1}{12}\begin{pmatrix} 7 & -1/4 \\ -1/4 & 7 \end{pmatrix}.
\end{aligned}$$

Then

$$\nu(\{x_1, x_2\}) = \nu(\{x_1\}) + \nu(\{x_2\}) = \frac{1}{12}\begin{pmatrix} 5 & 1/4 \\ 1/4 & 5 \end{pmatrix}.$$

One can readily check that $\nu(\{x_1\})+2\nu(\{x_2\}) = \nu(\{x_3\})$, and so $\{A_1 \equiv \nu(\{x_1\}), A_2 \equiv \nu(\{x_2\})\}$ forms a measurement basis for ν, with basis residual

$$A_0 = \mathrm{I} - (A_1 + A_2) = \frac{1}{12}\begin{pmatrix} 7 & -1/4 \\ -1/4 & 7 \end{pmatrix} = \nu(\{x_3\}) \neq 0.$$

■

Proposition 3.1.6. *If ω is a projective measurement then $\mathcal{T}_\omega$ has a measurement basis and every measurement basis for $\mathcal{T}_\omega$ has trivial residual.*

Proof. Assume that $\omega : \mathcal{O}(X) \to \mathfrak{B}(\mathcal{K})$ is a projective measurement and that $\{\omega(F_1), \ldots, \omega(F_m)\}$ is a linear basis for $\mathcal{T}_\omega$. Set $Q_j = \omega(F_j)$ for each j; thus, $Q_1, \ldots, Q_m$ are linearly independent (pairwise commuting) projections. Let $E_1 = F_1$ and define, iteratively,

$$E_j \;=\; F_j \setminus \left(\bigcup_{i=1}^{j-1} F_i\right), \quad j = 2, \ldots, m\,.$$

The sets $E_1, \ldots, E_m$ are pairwise disjoint and nonempty, and therefore the projections $P_1, \ldots, P_m$, where each $P_j = \omega(E_j)$, are nonzero and pairwise orthogonal. Thus, $\{P_1, \ldots, P_m\}$ is a set of m linearly independent operators whose linear span is a subspace of the m-dimensional space $\mathcal{T}_\omega$. Hence, $\{P_1, \ldots, P_m\}$ is a measurement basis for $\mathcal{T}_\omega$. Note that $P_0 = \mathrm{I} - \sum_{j=1}^m P_j$ is orthogonal to each P_j, and thus P_0 is either zero or is linearly independent of $P_1, \ldots, P_m$. But $m = \dim\mathcal{T}_\omega$ yields $P_0 = 0$, showing that the measurement basis $\mathcal{B}_\omega = \{P_1, \ldots, P_m\}$ for ω has trivial residual.

Indeed this latter argument shows that every measurement basis for $\mathcal{T}_\omega$ has trivial residual. ■

The decomposition of X into a finite disjoint-union of measurable sets $E_0, \ldots, E_m$ in the proof of proposition 3.1.6 is an idea that appears in other works on quantum measurement—for example, as in the decomposition of phase space into cells given in [BŻ06, (10.19)].

The following result describes the structure of POVMs in terms of measurement bases and classical real-valued measures. To state the result we require an operator version of absolute continuity.

Definition 3.1.7. *If $\upsilon_j : \mathcal{O}(X) \to \mathfrak{B}(\mathcal{H}_j)$ are functions, for $j = 1, 2$, then υ_2 is* absolutely continuous *with respect to υ_1, denoted by $\upsilon_2 \ll_{\text{ac}} \upsilon_1$, if $\upsilon_2(E) = 0$ for every $E \in \mathcal{O}(X)$ for which $\upsilon_1(E) = 0$.*

Proposition 3.1.8. *[FFP13, Prop. IV.3] If $\{A_1, \ldots, A_m\}$ is a measurement basis for a quantum probability measure ν, then there exist finite signed measures $\upsilon_1, \ldots, \upsilon_m$ on $(X, \mathcal{O}(X))$ such that each $\upsilon_j \ll_{\text{ac}} \nu$ and*

$$\nu(E) = \sum_{j=1}^{m} \upsilon_j(E) A_j \,, \text{ for all } E \in \mathcal{O}(X) \,.$$

Proof. For every $E \in \mathcal{O}(X)$ there exist unique $\alpha_1^E, \ldots, \alpha_m^E \in \mathbb{R}$ such that $\nu(E) = \sum_{j=1}^m \alpha_j^E A_j$. Define $\upsilon_j : \mathcal{O}(X) \to \mathbb{R}$ by $\upsilon_j(E) = \alpha_j^E$, $j = 1, \ldots, m$. By the linear independence of $A_1, \ldots, A_m$, we have $0 = \nu(\emptyset) = \sum_{j=1}^m \alpha_j^E A_j \Rightarrow \alpha_j^E = 0$ for each j; that is, $\upsilon_j(\emptyset) = 0$ for each j. To show that υ_j is countably additive, suppose

that $\{E_k\}_{k\in\mathbb{N}}$ is a countable collection of pairwise disjoint measurable sets. By the countable additivity of ν,

$$\nu\left(\bigcup_{k\in\mathbb{N}} E_k\right) = \sum_{k\in\mathbb{N}} \nu(E_k) = \sum_{k\in\mathbb{N}}\sum_{j=1}^{m} \alpha_j^{E_k} A_j = \sum_{j=1}^{m}\left(\sum_{k\in\mathbb{N}} \alpha_j^{E_k}\right) A_j\,,$$

which proves—again by the linear independence of $A_1, \ldots, A_m$—that

$$\upsilon_j\left(\bigcup_{k\in\mathbb{N}} E_k\right) = \sum_{k\in\mathbb{N}} \upsilon_j(E_k)\,,$$

for every $j = 1, \ldots, m$. Lastly, if $\nu(E) = 0$, then the linear independence of A_1, …, A_m yields $\upsilon_j(E) = 0$ for every j, whence each $\upsilon_j \ll_{\text{ac}} \nu$. ■

In proposition 3.1.8, it is not generally true that the signed measures υ_j are in fact positive measures.

Definition 3.1.9. *A measurement basis $\mathcal{B}_\nu = \{A_1, \ldots, A_m\}$ for a quantum probability measure ν is* perfect *if there exist (positive) probability measures μ_1, …, μ_m on $(X, \mathcal{O}(X))$ such that*

$$\nu(E) = \sum_{j=1}^{m} \mu_j(E) A_j\,, \text{ for all } E \in \mathcal{O}(X)\,.$$

It is not difficult to verify that every measurement basis $\mathcal{B}_\omega$ of a projective quantum probability measure ω is perfect. Indeed, by the spectral theorem we can write any Hermitian operator A as a linear combination of pairwise orthogonal projections: $A = \sum_{j=1}^{m} \lambda_j P_j$, where $\{\lambda_j\}_{j=1}^{m}$ is the spectrum of A. If A is positive, then $\lambda_j \geq 0$ for all j. So, if $\omega(E) = \sum_{j=1}^{m} \omega_j A_j$, then the spectrum of $\omega(E)$ is $\{\omega_j(E)\}_j = \{0, 1\}$ (since ω is a projection). Furthermore, if all elements of the measurement basis

$\{A_j\}_{j=1}^m$ are projections and $\sum_j A_j = \mathrm{I}$ (by proposition 3.1.6), then the A_j's are pairwise orthogonal, hence we can simultaneously diagonalize them. Thus we have

$$\omega(E) = \sum_j \omega_j(E) A_j = \mathrm{diag}(\omega_1(E), \ldots, \omega_2(E), \ldots, \omega_m(E)),$$

where the number of $\omega_j(E)$'s corresponds to the rank of A_j. Since ω is positive, it follows immediately that $\omega_j(E) \geq 0$ for all j.

In quantum information theory it is a standard practise to define a POVM associated with an n-outcome quantum measurement to be an n-tuple of positive operators $M_1, \ldots, M_n$ such that $\sum_{j=1}^n M_j = \mathrm{I}$. In this scenario, the underlying quantum probability measure $\nu : \mathcal{O}(X) \to \mathfrak{B}(\mathcal{H})$ is the unique POVM for which $\nu(\{x_j\}) = M_j$ for each j, where $X = \{x_1, \ldots, x_n\}$. Observe that $\mathcal{O}(X)$ is the power set of X and that

$$\mathcal{R}_\nu = \left\{ \sum_{j \in E} M_j \ : \ E \in \mathcal{O}(X) \right\} \subset \mathrm{Span}\{M_1, \ldots, M_n\} .$$

Thus, $\mathrm{Span}\{M_1, \ldots, M_n\} = \mathcal{T}_\nu$. However, any measurement basis $\{A_1, \ldots, A_m\}$ for $\mathcal{T}_\nu$ is drawn from the set $\mathcal{R}_\nu$ rather than from $\{M_1, \ldots, M_n\}$. This is an essential difference in our approach from other studies, where the POVM is analyzed by studying the spanning set $\{M_1, \ldots, M_n\}$. By way of the observation above it is clear that $\dim \mathcal{T}_\nu \leq |X|$, the cardinality of the sample space X.

Proposition 3.1.10. *[FFP13, Prop. IV.5] If $\{M_1, \ldots, M_n\}$ is a POVM associated with a quantum probability measure ν on a sample space X of cardinality n, then the following statements are equivalent:*

1. $\dim \mathcal{T}_\nu = n$;

2. $\{M_1, \ldots, M_n\}$ *is a perfect measurement basis for* ν*.*

Proof. If $\dim \mathcal{T}_\nu = n$, then $\mathrm{Span}\{M_1, \ldots, M_n\}$ is n-dimensional, which implies that $\{M_1, \ldots, M_n\}$ is a basis of $\mathcal{T}_\nu$. Moreover, $\{M_1, \ldots, M_n\}$ is clearly a measurement basis and

$$\nu = \sum_{j=1}^{n} \delta_{\{x_j\}} M_j \,,$$

where $\delta_{\{x_j\}}$ is the Dirac probability measure with mass concentrated at the point set $\{x_j\}$. The converse is obvious. ■

3.2 Ordering POVMs by Absolute Continuity

Recall the notion of absolute continuity introduced earlier:

Definition 3.2.1. *If* ν_1 *and* ν_2 *are quantum probability measures, then* ν_2 *is* absolutely continuous *with respect to* ν_1*, denoted by* $\nu_2 \ll_{\mathrm{ac}} \nu_1$*, if* $\nu_2(E) = 0$ *for every* $E \in \mathcal{O}(X)$ *for which* $\nu_1(E) = 0$*.*

In classical measure theory, absolute continuity is characterized by the Radon-Nikodým theorem. In this section a similar result will be established for POVMs. To do so, a notion of integration is is required.

3.2.1 Quantum random variables and integration

Definition 3.2.2. *A* quantum random variable *is a function* $\psi : X \to \mathfrak{B}(\mathcal{H})$ *that is* measurable *in the sense that the complex-valued functions*

$$x \mapsto \mathrm{Tr}\,(\rho\psi(x))$$

are measurable for every state $\rho \in \mathfrak{B}(\mathcal{H})_t$.

Equivalently, $\psi : X \to \mathfrak{B}(\mathcal{H})$ is measurable if, for every pair of vectors $|\xi\rangle, |\eta\rangle \in \mathcal{H}$, the complex-valued function $x \mapsto \langle \xi \,|\, \psi(x) \,|\, \eta \rangle$ is measurable. Our aim in this section is to define, using the procedure set out in [FZ07], a positive-preserving operator-valued integral $\int_X \psi \, d\nu$ for any measurable function $\psi : X \to \mathfrak{B}(\mathcal{H})$ and any quantum probability measure $\nu : \mathcal{O}(X) \to \mathfrak{B}(\mathcal{H})$.

Every positive operator $H \in \mathfrak{B}(\mathcal{H})$ has a unique positive square root $H^{1/2}$. Thus, if $\psi : X \to \mathfrak{B}(\mathcal{H})$ is a function for which $\psi(x)$ is a positive operator for every $x \in X$, then $\psi^{1/2} : X \to \mathfrak{B}(\mathcal{H})$ denotes the function $\psi^{1/2}(x) = (\psi(x))^{1/2}$.

The following observation will be useful.

Proposition 3.2.3. *[FPS11, Prop. II.1] If* $\psi : X \to \mathfrak{B}(\mathcal{H})$ *is a positive quantum random variable, then* $\psi^{1/2}$ *is a (positive) quantum random variable.*

Proof. Assume first that $\psi(x)$ is positive and invertible for every $x \in X$. Because sums and products of scalar-valued measurable functions are measurable, if one invokes an iterative procedure to compute $\psi(x)^{1/2}$—such as the one in [Hig97, Algorithm 2],

which is a Newton-type iteration combined with a Cholesky factorisation—then for each state ρ the function

$$x \longmapsto \mathrm{Tr}\left(\rho\psi(x)^{1/2}\right)$$

is a pointwise limit of a sequence of measurable functions. Thus, $\psi^{1/2}$ is a quantum random variable. In the case where $\psi(x)$ is not invertible for all $x \in X$, then $\psi^{1/2}$ is a pointwise limit of $x \mapsto (\psi(x) + \frac{1}{n}\,\mathrm{I})^{1/2}$ as n tends to infinity and, hence, is measurable. ■

3.2.2 The Principal Radon-Nikodým Derivative

Given ν, a probability measure μ is obtained from ν via

$$\mu(E) \;=\; \frac{\mathrm{Tr}\,(\nu(E))}{d},\ \text{for every } E \in \mathcal{O}(X)\,. \tag{3.1}$$

Because the trace functional maps nonzero positive operators to strictly positive real numbers, the measures μ and ν are mutually absolutely continuous: $\mu \ll_{\mathrm{ac}} \nu$ and $\nu \ll_{\mathrm{ac}} \mu$.

Assume that $\{|e_1\rangle, \ldots, |e_d\rangle\}$ is a fixed orthonormal basis of $\mathcal{H}$. Because $\nu \ll_{\mathrm{ac}} \mu$, each of the d^2 complex measures $\nu_{ij} : \mathcal{O}(X) \to \mathbb{C}$, defined by $\nu_{ij}(E) = \langle e_i \,|\, \nu(E) \,|\, e_j\rangle$, has the property that $\nu_{ij} \ll_{\mathrm{ac}} \mu$. Hence, by the (classical) Radon-Nikodým theorem, there is a unique $\dfrac{d\nu_{ij}}{d\mu} \in L^1(X,\mu)$ such that

$$\nu_{ij}(E) \;=\; \int_E \frac{d\nu_{ij}}{d\mu}\,d\mu,\ \text{for all } E \in \mathcal{O}(X)\,.$$

These scalar Radon-Nikoým derivatives give rise to a quantum random variable $\frac{d\nu}{d\mu}$: $X \to \mathfrak{B}(\mathcal{H})$ via

$$\frac{d\nu}{d\mu} = \sum_{i,j=1}^{d} \frac{d\nu_{ij}}{d\mu} \otimes |e_i\rangle\langle e_j| \,. \tag{3.2}$$

Let $\xi_k \in \mathbb{C}$ for $k = 1, \ldots, d$. Notice that for any $\xi = \sum_{k=1}^{d} \xi_k |e_k\rangle \in \mathcal{H}$, we have

$$\left\langle \xi \,\middle|\, \frac{d\nu}{d\mu} \,\middle|\, \xi \right\rangle = \sum_{i,j=1}^{d} \frac{d\nu_{ij}}{d\mu} \xi_i^* \xi_j \,.$$

Hence, for all $|\xi\rangle \in \mathcal{H}$ and $E \in \mathcal{O}(X)$,

$$\int_E \left\langle \xi \,\middle|\, \frac{d\nu}{d\mu}(x) \,\middle|\, \xi \right\rangle d\mu(x) = \sum_{i,j=1}^{d} \left(\int_E \frac{d\nu_{ij}}{d\mu} \, d\mu \right) \xi_i^* \xi_j = \langle \xi \,|\, \nu(E) \,|\, \xi \rangle \geq 0 \,.$$

This proves that $\frac{d\nu}{d\mu}(x)$ is a positive operator for μ-almost all $x \in X$; for such x let $\left(\frac{d\nu}{d\mu}(x)\right)^{1/2}$ denote the positive square root (in $\mathfrak{B}(\mathcal{H})$) of the positive operator $\frac{d\nu}{d\mu}(x)$. Now define $\left(\frac{d\nu}{d\mu}\right)^{1/2} : X \to \mathfrak{B}(\mathcal{H})$ to be $\left(\frac{d\nu}{d\mu}(x)\right)^{1/2}$ at those $x \in X$ for which $\frac{d\nu}{d\mu}(x)$ is a positive operator, and zero otherwise.

Definition 3.2.4. *If ν is a quantum random variable and if μ is the induced classical probability measure defined in equation (3.1), then the measurable function $\frac{d\nu}{d\mu}$ defined in equation (3.2) is called the* principal Radon-Nikodým derivative of ν.

Unlike the classical case, whenever $d > 1$ the principal Radon-Nikodým derivative of ν depends on the pre-selected choice of orthonormal basis $\{|e_1\rangle, \ldots, |e_d\rangle\}$ of $\mathcal{H}$. If one had chosen a different orthonormal basis, say $\{|e_1'\rangle, \ldots, |e_d'\rangle\}$, then the resulting principal Radon-Nikodým derivative computed in this new basis is simply that of $\alpha \circ \nu$

in the originally selected basis, where α is the automorphism induced by the unitary operator that transforms the basis $\{|e'_1\rangle, \ldots, |e'_d\rangle\}$ to the basis $\{|e_1\rangle, \ldots, |e_d\rangle\}$. The following proposition is even more general.

Proposition 3.2.5. *[FPS11, Prop. II.2] Assume that $\Phi : \mathfrak{B}(\mathcal{H})_t \to \mathfrak{B}(\mathcal{H})_t$ is a unital quantum channel. Let μ^ν and $\mu^{\Phi\circ\nu}$ be the probability measures induced by the quantum probability measures ν and $\Phi \circ \nu$ in accordance with equation (3.1). Then there is a classical probability measure μ such that*

1. $\mu = \mu^\nu = \mu^{\Phi\circ\nu}$ *and*
2. $\dfrac{d(\Phi \circ \nu)}{d\mu} = \Phi \circ \dfrac{d\nu}{d\mu}$.

Proof. The channel Φ is trace preserving, so for any $E \in \mathcal{O}(X)$

$$\mu^{\Phi\circ\nu}(E) = \frac{1}{d}\operatorname{Tr}(\Phi(\nu(E))) = \frac{1}{d}\operatorname{Tr}(\nu(E)) = \mu^\nu(E)\,.$$

The desired measure is $\mu = \mu^{\Phi\circ\nu} = \mu^\nu$.

Let $A = \sum_{i,j=1}^d \alpha_{ij}|i\rangle\langle j| \in \mathfrak{B}(\mathcal{H})$ and consider $A\nu A^\dagger$. If $\mu(E) = 0$, then $\nu(E) = 0$ and $A\nu(E)A^\dagger = 0$; therefore $A\nu A^\dagger \ll_{\text{ac}} \mu$. Fix i, j and consider the (i, j)-coordinate measure of $A\nu A^\dagger$:

$$\omega_{ij} = \sum_{l=1}^d \sum_{k=1}^d \alpha_{il}\alpha^*_{jk}\nu_{lk}.$$

Since $A\nu A^\dagger \ll_{\text{ac}} \mu$, we have $\omega_{ij} \ll_{\text{ac}} \mu$ and so we may consider the Radon-Nikodým derivative

$$\frac{d\omega_{ij}}{d\mu} = \sum_{l=1}^d \sum_{k=1}^d \alpha_{il}\alpha^*_{jk}\left(\frac{d\nu_{lk}}{d\mu}\right) = \left(A\frac{d\nu}{d\mu}A^\dagger\right)_{ij}.$$

Therefore, $\dfrac{d(A\nu A^\dagger)}{d\mu} = A\dfrac{d\nu}{d\mu}A^\dagger$.

Consider the Kraus decomposition of the channel Φ:

$$\Phi(\rho) = \sum_{j=1}^{q} A_j Y A_j^\dagger,\ \rho \in \mathfrak{B}(\mathcal{H})_t,\ \text{where}\ \sum_{j=1}^{q} A_j^\dagger A_j = \sum_{j=1}^{q} A_j A_j^\dagger = \mathrm{I}\ .$$

By linearity of the scalar Radon-Nikodým derivative, $\dfrac{d(\Phi\circ\nu)}{d\mu} = \Phi\circ\dfrac{d\nu}{d\mu}$. ■

3.2.3 Integrable Functions

If $f, \psi : X \to \mathfrak{B}(\mathcal{H})$ are quantum random variables such that $\psi(x) \in \mathfrak{B}(\mathcal{H})_+$ for all $x \in X$, then $\psi^{1/2}$ is measurable (proposition 3.2.3) and, thus, the function $\psi^{1/2} f \psi^{1/2}$ is Borel measurable.

Definition 3.2.6. *Assume that* $\dfrac{d\nu}{d\mu}$ *is the principal Radon-Nikodým derivative of* ν.

1. *If* $f : X \to \mathfrak{B}(\mathcal{H})$ *is a quantum random variable, then* f *is said to be* ν-integrable *if, for every state* ρ*, the complex-valued function*

$$f_\rho(x) \;=\; \mathrm{Tr}\left(\rho \left(\frac{d\nu}{d\mu}(x)\right)^{1/2} f(x) \left(\frac{d\nu}{d\mu}(x)\right)^{1/2}\right)\ ,\ x \in X,$$

is μ*-integrable.*

2. *The* integral *of a* ν*-integrable function* $f : X \to \mathfrak{B}(\mathcal{H})$ *is defined to be the unique operator acting on* $\mathcal{H}$ *having the property that*

$$\mathrm{Tr}\left(\rho \int_X f\, d\nu\right) \;=\; \int_X f_\rho\, d\mu$$

for every state ρ *of* $\mathcal{H}$.

Proposition 3.2.7. *[FPS11, Example II.3] If $0 \leq f(x) \leq \mathrm{I}$ for all $x \in X$, then $0 \leq \int_X f \, d\nu \leq \mathrm{I}$.*

Proof. Let μ be the principal Radon-Nikodým derivative of ν. Recall that if A and B are positive operators with $A \leq B$, then $S^\dagger A S \leq S^\dagger B S$ for every $S \in \mathfrak{B}(\mathcal{H})$. Thus, because $0 \leq f(x) \leq \mathrm{I}$, for every state ρ we have

$$0 \leq \rho^{1/2} \left(\frac{d\nu}{d\mu}(x)\right)^{1/2} f(x) \left(\frac{d\nu}{d\mu}(x)\right)^{1/2} \rho^{1/2} \leq \rho^{1/2} \left(\frac{d\nu}{d\mu}(x)\right) \rho^{1/2}$$

for μ-almost all $x \in X$. Thus, for every $\rho \in \mathrm{S}(\mathcal{H})$, we have

$$\int_X f_\rho \, d\mu \leq \int_X \mathrm{Tr}\left(\rho \frac{d\nu}{d\mu}\right) d\mu$$

and so $0 \leq \displaystyle\int_X f \, d\nu \leq \int_X \left(\frac{d\nu}{d\mu}\right) d\nu = \nu(X) = \mathrm{I} \in \mathfrak{B}(\mathcal{H})$. ∎

The example below illustrates the use of definition 3.2.6 as well as proposition 3.2.7.

Example 3.2.8. *[FPS11, Example II.4] The principal Radon-Nikodým derivative of $\nu = \displaystyle\sum_{j=1}^{n} \delta_{x_j} H_j$ and the corresponding integral formula.*

Here, we assume that $H_1, \ldots, H_n \in \mathfrak{B}(\mathcal{H})_+$ are nonzero and satisfy $\sum_j H_j = \mathrm{I}$ and that $\{x_1, \ldots, x_n\}$ is a set of n distinct points of X. The quantum probability measure ν is defined by

$$\nu(E) = \sum_{j=1}^{n} \delta_{x_j}(E) H_j, \; E \in \mathcal{O}(X)\,.$$

If χ_E denotes the characteristic (or indicator) function of any measurement event $E \in \mathcal{O}(X)$, then

$$\frac{d\nu}{d\mu} = \sum_{j=1}^{n} \left(\frac{d}{\mathrm{Tr}(H_j)} \chi_{\{x_j\}} \right) H_j$$

and

$$\int_X f \, d\nu = \sum_{j=1}^{n} H_j^{1/2} f(x_j) H_j^{1/2},$$

for every measurable function $f : X \to \mathfrak{B}(\mathcal{H})$.

By the previous proposition, this implies that if $0 \leq f(x) \leq \mathrm{I}$ for all $x \in X$, then $0 \leq \sum_{j=1}^{n} H_j^{1/2} f(x_j) H_j^{1/2} \leq \mathrm{I}$. ♢

3.2.4 A Radon-Nikodým Theorem

If $H \in \mathfrak{B}(\mathcal{H})_+$, then H^{-1} shall denote the unique positive operator for which $\ker H^{-1} = \ker H$ and $H^{-1}H = HH^{-1} = Q$, the projection onto the range of H. Thus, if H is invertible, then H^{-1} is the inverse of H. Once we have this notion for positive operators, a similar notion of generalised inverse for positive operator valued functions can be made.

The following theorem characterizes the order induced by absolute continuity.

Theorem 3.2.9. *[FPS11, Theorem 2.7] The following statements are equivalent for quantum probability measures $\nu_1, \nu_2 : \mathcal{O}(X) \to \mathfrak{B}(\mathcal{H})$:*

1. $\nu_2 \ll_{\mathrm{ac}} \nu_1$;

2. *there exists a bounded Borel function* $g : X \to \mathfrak{B}(\mathcal{H})$*, unique up to sets of* ν_1*-measure zero, such that*

$$\nu_2(E) \;=\; \int_E g \, d\nu_1, \quad \textit{for every } E \in \mathcal{O}(X)\,. \tag{3.3}$$

If the equivalent conditions above hold and if μ_j *is the probability measure induced by* ν_j*, then* $\mu_2 \ll_{\mathrm{ac}} \mu_1$ *and*

$$g \;=\; \left(\frac{d\mu_2}{d\mu_1}\right)\left[\left(\frac{d\nu_1}{d\mu_1}\right)^{-1/2}\left(\frac{d\nu_2}{d\mu_2}\right)\left(\frac{d\nu_1}{d\mu_1}\right)^{-1/2}\right]. \tag{3.4}$$

Proof. Assume that $\nu_2 \ll_{\mathrm{ac}} \nu_1$. If $\mu_1(E) = \frac{1}{d}\operatorname{Tr}(\nu_1(E)) = 0$, then $\nu_1(E) = 0$. By assumption $\nu_2(E) = 0$ and therefore $\mu_2(E) = \frac{1}{d}\operatorname{Tr}(\nu_2(E)) = 0$, which proves that $\mu_2 \ll_{\mathrm{ac}} \mu_1$. Therefore, for any $E \in \mathcal{O}(X)$, we have

$$\nu_2^{(i,j)}(E) = \langle e_i \,|\, \nu_2(E) \,|\, e_j \rangle \ll_{\mathrm{ac}} \langle e_i \,|\, \nu_1(E) \,|\, e_j \rangle = \nu_1^{(i,j)}(E).$$

Coordinate-wise we obtain $\nu_2^{(i,j)} \ll_{\mathrm{ac}} \mu_1$. By applying the chain rule for the classical Radon-Nikodým derivative, we obtain

$$\frac{d\nu_2^{(i,j)}}{d\mu_1} = \frac{d\nu_2^{(i,j)}}{d\mu_2}\frac{d\mu_2}{d\mu_1}$$

Hence,

$$\frac{d\nu_2}{d\mu_1} = \frac{d\nu_2}{d\mu_2}\frac{d\mu_2}{d\mu_1},$$

where $\dfrac{d\nu_2}{d\mu_2} : X \to \mathfrak{B}(\mathcal{H})_+$ and $\dfrac{d\mu_2}{d\mu_1} : X \to \mathbb{R}_+$. With g as above,

$$\left(\frac{d\nu_1}{d\mu_1}\right)^{1/2} g \left(\frac{d\nu_1}{d\mu_1}\right)^{1/2} \;=\; \left(\frac{d\mu_2}{d\mu_1}\right)\frac{d\nu_2}{d\mu_2}\,.$$

Thus, for any state ρ and $E \in \mathcal{O}(X)$,

$$\begin{aligned}
\mathrm{Tr}\left(\rho \int_E g d\nu_1\right) &= \int_E \mathrm{Tr}\left(\rho \left(\frac{d\nu_1}{d\mu_1}\right)^{1/2} g \left(\frac{d\nu_1}{d\mu_1}\right)^{1/2}\right) d\mu_1 \\
&= \int_E \left[\mathrm{Tr}\left(\rho \frac{d\nu_2}{d\mu_2}\right)\right] \frac{d\mu_2}{d\mu_1}\, d\mu_1 \\
&= \int_E \mathrm{Tr}\left(\rho \frac{d\nu_2}{d\mu_2}\right) d\mu_2 \\
&= \mathrm{Tr}(\rho\, \nu_2(E)).
\end{aligned}$$

Therefore, by definition of the integral, $\nu_2(E) = \int_E g\, d\nu_1$ for every $E \in \mathcal{O}(X)$. Passing to the d^2 coordinate measures and using the uniqueness of the classical Radon-Nikodým derivative, one deduces that g is unique up to sets of ν_1-measure zero.

Conversely, assume such a function $g : X \to \mathfrak{B}(\mathcal{H})$ exists such that

$$\nu_2(E) = \int_E g\, d\nu_1, \quad \text{for every } E \in \mathcal{O}(X)\,.$$

If $\nu_1(E) = 0$, then $\nu_2(E) = \int_E g\, d\nu_1 = 0$ and thus $\nu_2 \ll_{\mathrm{ac}} \nu_1$. ■

The function g in (3.4) is called a *non-principal Radon-Nikodým derivative* of ν_2 with respect to ν_1.

3.3 Ordering POVMs by Cleanness

Definition 3.3.1. ([Pel11]) *If ν_1 and ν_2 are quantum probability measures on $(X, \mathcal{O}(X))$ with values in $\mathfrak{B}(\mathcal{H}_1)$ and $\mathfrak{B}(\mathcal{H}_2)$ respectively, then ν_1 is* cleaner than ν_2, *denoted by $\nu_2 \ll_{\mathrm{cl}} \nu_1$, if $\nu_2 = \Phi^\dagger \circ \nu_1$ for some quantum channel $\Phi : \mathfrak{B}(\mathcal{H}_2)_t \to \mathfrak{B}(\mathcal{H}_1)_t$.*

Note that if $\nu_2 \ll_{\text{cl}} \nu_1$ via a quantum channel Φ, then the measurement statistics satisfy

$$\text{Tr}_{\mathcal{H}_2}\left(\rho\,\nu_2(E)\right) \;=\; \text{Tr}_{\mathcal{H}_1}\left(\Phi(\rho)\nu_1(E)\right), \quad \forall\,\rho \in \mathfrak{B}(\mathcal{H}_2)_t,\; E \in \mathcal{O}(X)\,.$$

Our first objective is to characterise the order relation $\nu_2 \ll_{\text{cl}} \nu_1$. The theorem below can be seen as the finite-dimensional simplification of [FFP13, Theorem III.1].

Theorem 3.3.2. *If $\mathcal{B}_\nu = \{\nu(E_1), \ldots, \nu(E_m)\}$ is a measurement basis for a quantum probability measure $\nu : \mathcal{O}(X) \to \mathfrak{B}(\mathcal{H})$, then the following statements are equivalent for a quantum probability measure $\nu' : \mathcal{O}(X) \to \mathfrak{B}(\mathcal{H}')$:*

1. *$\nu' \ll_{\text{cl}} \nu$;*

2. *for all $L_0, \ldots, L_m \in \mathbb{M}_p(\mathbb{C})$ and every $p \in \mathbb{N}$,*

$$\left\| \sum_{j=0}^{m} \nu'(E_j) \otimes L_j \right\| \leq \left\| \sum_{j=0}^{m} \nu(E_j) \otimes L_j \right\| . \tag{3.5}$$

Proof. (1) $\Rightarrow$ (2). Because $\nu' = \Phi^\dagger \circ \nu$ for some quantum channel Φ, by letting $\psi = \Phi^\dagger$ and by using the fact that because ucp maps are completely contractive [Pau02, Proposition 3.6], $\left\| \sum_{j=0}^{m} \psi(\nu(E_j)) \otimes L_j \right\| \leq \left\| \sum_{j=0}^{m} \nu(E_j) \otimes L_j \right\|$ for all $L_0, \ldots, L_m \in \mathbb{M}_p(\mathbb{C})$ and every $p \in \mathbb{N}$. That is, inequality (3.5) holds.

(2) $\Rightarrow$ (1). Assume that for all $L_0, \ldots, L_m \in \mathbb{M}_p(\mathbb{C})$ and every $p \in \mathbb{N}$, inequality (3.5) holds.

Consider now the operator systems

$$\mathcal{X}_\nu = \text{Span}\,\{A_1, \ldots, A_m\} \quad \text{and} \quad \mathcal{X}_{\nu'} = \text{Span}\,\{B_1, \ldots, B_m\},$$

where $A_j = \nu(E_j)$ and $B_j = \nu'(E_j)$ for $j = 1, \ldots, m$, and let $\psi_0 : \mathcal{X}_\nu \to \mathcal{X}_{\nu'}$ denote the linear function defined by

$$\psi_0\left(\sum_{j=1}^{m} \alpha_j A_j\right) = \sum_{j=1}^{m} \alpha_j B_j \,,$$

for $\alpha_1, \ldots, \alpha_m \in \mathbb{C}$. Inequality (3.5) shows that the unital linear map ψ_0 is completely contractive. Hence, by Arveson's extension theorem [Pau02, Theorem 7.5], ψ_0 has a completely contractive extension to a ucp map $\psi : \mathfrak{B}(\mathcal{H}) \to \mathfrak{B}(\mathcal{H}')$, which maps the operator system $\mathcal{X}_\nu$ into $\mathcal{X}_{\nu'}$. Now let $\Phi = \psi^\dagger$ to obtain $\nu' = \Phi^\dagger \circ \nu$. ■

We turn to a useful preliminary result (lemma 3.3.4) related to the inequalities in statement (2) of theorem 3.3.2 above. To do so, we recall the concept of joint spectrum for commuting Hermitian operators. Recall that if $A_1, \ldots, A_m$ are pairwise commuting Hermitian operators acting on $\mathcal{H}$, then there are orthonormal basis vectors $|\phi_1\rangle$, ..., $|\phi_d\rangle$ such that each $|\phi_\ell\rangle$ is an eigenvector of each operator A_j.

Definition 3.3.3. *Suppose that $|\phi_1\rangle$, ..., $|\phi_d\rangle$ form an orthonormal basis of $\mathcal{H}$ consisting of (joint) eigenvectors of m commuting Hermitian operators $A_1, \ldots, A_m \in \mathfrak{B}(\mathcal{H})$. The* joint spectrum *of the m-tuple $(A_1, \ldots, A_m)$ is the set*

$$\mathrm{Sp}(A_1, \ldots, A_m) = \left\{ \begin{pmatrix} \lambda_{1\ell} \\ \lambda_{2\ell} \\ \vdots \\ \lambda_{m\ell} \end{pmatrix} : 1 \leq \ell \leq d \right\},$$

where $A_j|\phi_\ell\rangle = \lambda_{j\ell}|\phi_\ell\rangle$ for all $j = 1, \ldots, m$ and $\ell = 1, \ldots, d$.

Let $\mathrm{C}^*(A_1, \ldots, A_m)$ denote the unital C*-algebra generated by the pairwise commuting Hermitian operators $A_1, \ldots, A_m$. Thus, $\mathrm{C}^*(A_1, \ldots, A_m)$ is abelian and for

each ℓ the function $\varrho_\ell : \mathrm{C}^*(A_1, \ldots, A_m) \to \mathbb{C}$ for which $\varrho_\ell(A_j) = \lambda_{j\ell}$, for all j, is a homomorphism. In fact, the Gelfand theory of abelian C*-algebras implies that if $\varrho : \mathrm{C}^*(A_1, \ldots, A_m) \to \mathbb{C}$ is a homomorphism, then $\varrho = \varrho_\ell$ for some ℓ. Thus, let $\mathfrak{M} = \{\varrho_1, \ldots, \varrho_d\}$, which is called the *maximal ideal space* of $\mathrm{C}^*(A_1, \ldots, A_m)$, and consider each A_j as a (continuous) function $\mathfrak{M} \to \mathbb{C}$ whereby $A_j(\varrho) = \varrho(A_j)$.

Lemma 3.3.4. *[FFP13, Lemma III.3] If $A_1, \ldots, A_m \in \mathfrak{B}(\mathcal{H})$ are positive operators such that $\sum_{j=1}^m A_j = \mathrm{I}$, then for every Hilbert space $\mathcal{K}$ and operators $L_1, \ldots, L_m \in \mathfrak{B}(\mathcal{K})$ we have*

$$\left\| \sum_{j=1}^m A_j \otimes L_j \right\| \leq \max_{1 \leq j \leq m} \|L_j\| .$$

If, moreover, $A_1, \ldots, A_m$ are pairwise commuting with joint spectrum $\Lambda \equiv \mathrm{Sp}(A_1, \ldots, A_m) \subset \mathbb{R}^m$, then

$$\left\| \sum_{j=1}^m A_j \otimes L_j \right\| = \max_{\lambda \in \Lambda} \left\| \sum_{j=1}^m \lambda_j L_j \right\| .$$

In particular, if $A_1, \ldots, A_m$ are nonzero projections, then

$$\left\| \sum_{j=1}^m A_j \otimes L_j \right\| = \max_{1 \leq j \leq m} \|L_j\| .$$

Proof. Let $G_j = (A_j^{1/2} \otimes \mathrm{I}_\mathcal{K}) \in \mathfrak{B}(\mathcal{H} \otimes \mathcal{K})$; thus, $G_1, \ldots, G_m \in \mathfrak{B}(\mathcal{H} \otimes \mathcal{K})$ are positive operators such that $\sum_j G_j^2 = \mathrm{I} \in B(\mathcal{H} \otimes \mathcal{K})$. Let us pass to matrices of operators: let $L = (\mathrm{I}_\mathcal{H} \otimes L_1) \oplus \cdots \oplus (\mathrm{I}_\mathcal{H} \otimes L_m)$ and

$$G = \begin{pmatrix} G_1 & 0 & \cdots & 0 \\ G_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ G_m & 0 & \cdots & 0 \end{pmatrix} .$$

Thus, $\displaystyle\sum_{j=1}^{m} A_j \otimes L_j = \sum_{j=1}^{m} G_j(\mathrm{I}_{\mathcal{H}} \otimes L_j)G_j$ and

$$\left\| \sum_{j=1}^{m} G_j(\mathrm{I}_{\mathcal{H}} \otimes L_j)G_j \right\| \;=\; \|G^{\dagger}LG\| \;\leq\; \|G\|^2\|L\| \;=\; \|G^{\dagger}G\| \max_{1\leq j\leq m} \|\,\mathrm{I}_{\mathcal{H}} \otimes L_j\|\,.$$

Because $\|G^{\dagger}G\| = \left\|\sum_{j=1}^{m} G_j^2\right\| = 1$ and $\|\,\mathrm{I}_{\mathcal{H}} \otimes L_j\| = \|L_j\|$, the desired general inequality follows.

Suppose now that $A_1, \ldots, A_m$ are pairwise commuting. Hence, every operator A_j is a complex-valued continuous function on the maximal ideal space $\mathfrak{M}$ of $\mathrm{C}^*(A_1, \ldots, A_m)$ via $\varrho \mapsto \varrho(A_j)$. Because the C*-algebra $C(\mathfrak{M}) \otimes \mathfrak{B}(\mathcal{K})$ is isometrically isomorphic to the C*-algebra of continuous functions on $\mathfrak{M}$ with values in $\mathfrak{B}(\mathcal{K})$, we deduce that

$$\left\| \sum_{j=1} A_j \otimes L_j \right\| \;=\; \max_{\varrho\in\mathfrak{M}} \left\| \sum_{j=1} \varrho(A_j) \otimes L_j \right\| \;=\; \max_{\lambda\in\Lambda} \left\| \sum_{j=1}^{m} \lambda_j L_j \right\| .$$

Lastly, if $A_1, \ldots, A_m$ are nonzero projections, then the equation $\sum_{j=1}^{m} A_j = \mathrm{I}$ implies that $A_1, \ldots, A_m$ are pairwise orthogonal and, hence, pairwise commuting. Fix i and choose any ϱ in the maximal ideal space $\mathfrak{M}$ of $\mathrm{C}^*(A_1, \ldots, A_m)$ for which $\varrho(A_i) = 1$. (Such a ϱ exists because the spectrum of A_i contains 1.) Then $\mathrm{I} = A_i + \sum_{j\neq i} A_j$ implies that $\varrho(\mathrm{I}) = \varrho(A_i) + \sum_{j\neq i} \varrho(A_j)$, and so $\sum_{j\neq i} \rho(A_j) = 0$. As each $\varrho(A_j) \geq 0$, we obtain $\varrho(A_j) = 0$ for every $j \neq i$. Hence, the joint spectrum Λ of $(A_1, \ldots, A_m)$ contains only the canonical coordinate vectors $|1\rangle, \ldots, |m\rangle$ of $\mathbb{R}^m$ and,

therefore,

$$\max_{\lambda\in\Lambda} \|\sum_{j=1}^{m} \lambda_j L_j\| \;=\; \max_{1\leq j\leq m} \|L_j\|\,,$$

which completes the proof. ■

3.4 Clean POVMs

Definition 3.4.1. *A quantum probability measure ν is* clean *if $\nu' \ll_{\mathrm{cl}} \nu$ for every quantum probability measure ν' satisfying $\nu \ll_{\mathrm{cl}} \nu'$.*

The main result of this section, theorem 3.4.3 below, is a characterisation of clean POVMs and can be seen to be the finite-dimensional version of [FFP13, Theorem III.4]. Recall from the introduction to this chapter that we are using a definition of clean ordering that differs from the original definition given in [BKD+05].

Definition 3.4.2. *If $Q \in \mathfrak{B}(\mathcal{H})_+$, then $\lambda_{\min}(Q)$ and $\lambda_{\max}(Q)$ are the nonnegative real numbers*

$$\lambda_{\min}(Q) \;=\; \min\{\lambda \,:\, \lambda \in \mathrm{Sp}(Q)\} \quad \textit{and} \quad \lambda_{\max}(Q) \;=\; \max\{\lambda \,:\, \lambda \in \mathrm{Sp}(Q)\}\,,$$

where $\mathrm{Sp}(Z)$ denotes the spectrum of an operator $Z \in \mathfrak{B}(\mathcal{H})$.

Theorem 3.4.3. *The following statements are equivalent for a quantum probability measure ν:*

1. *ν is clean;*

2. *there are a Hilbert space* $\mathcal{K}$ *and a projective quantum probability measure* $\omega : \mathcal{O}(X) \to \mathfrak{B}(\mathcal{K})$ *such that the operator systems* $\mathcal{T}_\nu$ *and* $\mathcal{T}_\omega$ *are unitally completely order isomorphic;*

3. *for all Hilbert spaces* $\mathcal{K}$*, all operators* $L_1, \ldots, L_m \in \mathfrak{B}(\mathcal{K})$*, and every measurement basis* $\mathcal{B}_\nu = \{A_1, \ldots, A_m\}$ *for* ν*, the following two properties hold:*

   (a) $\mathcal{B}_\nu$ *has trivial residual, and*

   (b) $\left\| \sum_{j=1}^{m} A_j \otimes L_j \right\| = \max_{1 \leq j \leq m} \|L_j\|$*;*

4. *for every measurement basis* $\mathcal{B}_\nu = \{A_1, \ldots, A_m\}$ *for* ν*, the following two properties hold:*

   (a) $\mathcal{B}_\nu$ *has trivial residual, and*

   (b) $\lambda_{\max}(A_j) = 1$ *and* $\lambda_{\min}(A_j) = 0$ *for every* $j = 1, \ldots, m$*;*

5. *for each measurement basis* $\mathcal{B}_\nu = \{A_1, \ldots, A_m\}$ *for* ν*, the following two properties hold:*

   (a) $\mathcal{B}_\nu$ *has trivial residual, and*

   (b) *there exist a subspace* $\mathcal{H}_0 \subset \mathcal{H}$*, positive operators* $Y_1, \ldots, Y_m \in \mathfrak{B}(\mathcal{H}_0)$*, and pairwise-orthogonal projections* $Q_1, \ldots, Q_m \in \mathfrak{B}(\mathcal{H}_0^\perp)$ *such that, for every* $j = 1, \ldots, m$*,* $\|Y_j\| \leq 1$ *and* $A_j = Q_j \oplus Y_j \in B(\mathcal{H}_0^\perp \oplus \mathcal{H}_0)$*;*

Proof. (1) $\Rightarrow$ (2). Suppose that $\{B_1, \ldots, B_m\}$ is a linear basis for $\mathcal{T}_\nu$, where $B_j = \nu(F_j)$ for some $F_j \in \mathcal{O}(X)$. By Naimark's Dilation theorem there are a Hilbert space

$\mathcal{K}$, an isometry $V : \mathcal{H} \to \mathcal{K}$, and a projective measurement $\omega : \mathcal{O}(X) \to \mathfrak{B}(\mathcal{K})$ such that $\nu(E) = V^*\omega(E)V$ for all $E \in \mathcal{O}(X)$. If $\Phi : \mathfrak{B}(\mathcal{H})_t \to \mathfrak{B}(\mathcal{K})_t$ denotes the quantum channel $\Phi(R) = VRV^*$ and if $\phi = \Phi^*$, then Naimark's Dilation theorem takes the form $\nu = \phi \circ \omega$. Because ν is clean, $\omega = \psi \circ \nu$ for a unital completely positive linear map $\psi : \mathfrak{B}(\mathcal{H}) \to \mathfrak{B}(\mathcal{K})$. Thus $\mathcal{T}_\omega$ and $\mathcal{T}_\nu$ necessarily have the same dimension, namely m.

Let $Q_j = \omega(F_j) \in \mathcal{T}_\omega$ for each j. By the proof of proposition 3.1.6, there are pairwise disjoint sets $E_1, \ldots, E_m \in \mathcal{O}(X)$ such that, if $P_j = \omega(E_j)$ for each j, then $\{P_1, \ldots, P_m\}$ is a measurement basis for $\mathcal{T}_\omega$ with trivial residual. For each j let $A_j = \nu(E_j)$ and note that $P_j = \psi(A_j)$. Thus, $\{A_1, \ldots, A_m\}$ is a measurement basis for $\mathcal{T}_\nu$ with trivial residual.

Moreover, by lemma 3.3.4, for every $p \in \mathbb{N}$ and all $L_1, \ldots, L_m \in \mathbb{M}_p(\mathbb{C})$, we have

$$\begin{aligned} \max_{1\le j\le m} \|L_j\| &= \left\| \sum_{j=1}^m P_j \otimes L_j \right\| = \left\| \sum_{j=1}^m \psi(A_j) \otimes L_j \right\| \\ &\le \left\| \sum_{j=1}^m A_j \otimes L_j \right\| \\ &\le \max_{1\le j\le m} \|L_j\| . \end{aligned}$$

Thus, $\psi_{|\mathcal{T}_\nu}$ is a unital completely contractive map of $\mathcal{T}_\nu$ onto $\mathcal{T}_\omega$, and so $\psi_{|\mathcal{T}_\nu}$ is a unital complete order isomorphism [Pau02, Proposition 2.11].

(2) $\Rightarrow$ (3). Fix a measurement basis $\mathcal{B}_\nu = \{A_1, \ldots, A_m\}$ for ν. By assumption, there are a Hilbert space $\mathcal{K}$ and a projective quantum probability measure $\omega : \mathcal{O}(X) \to \mathfrak{B}(\mathcal{K})$ such that $\mathcal{T}_\nu$ and $\mathcal{T}_\omega$ are unitally completely order isomorphic. If

$\psi : \mathcal{T}_\nu \to \mathcal{T}_\omega$ denotes this complete order isomorphism, then for any finite-dimensional Hilbert space $\mathcal{K}$ and $L_1, \ldots, L_m \in \mathfrak{B}(\mathcal{K})$, we have, by lemma 3.3.4, that

$$\left\| \sum_{j=1}^{m} A_j \otimes L_j \right\| = \left\| \sum_{j=1}^{m} \psi(A_j) \otimes L_j \right\| = \max_{1 \leq j \leq m} \|L_j\| .$$

(3) $\Rightarrow$ (4). Let $\mathcal{B}_\nu = \{A_1, \ldots, A_m\}$ be a measurement basis for ν. Fix i and let $L_i = \mathrm{I}$ and $L_j = 0$ for $j \neq i$. Then, by assumption, $\|A_i \otimes L_i\| = \|L_i\|$ and so $\|A_i\| = 1$. Since $\lambda_{\max}(A_i) = \|A_i\|$, we deduce that $\lambda_{\max}(A_i) = 1$. Let $|\phi\rangle$ be a unit eigenvector of A_i corresponding to the eigenvalue 1. Then, because $\mathrm{I} - A_i = \sum_{j \neq i} A_j$, we deduce that

$$0 \leq \sum_{j \neq i} \langle \phi \,|\, A_j \,|\, \phi \rangle = \langle \phi \,|\, \mathrm{I} \,|\, \phi \rangle - \langle \phi \,|\, A_i \,|\phi\rangle = 1 - 1 = 0.$$

Thus, $\|A_j^{1/2} |\phi\rangle\| = 0$ for each $j \neq i$, which implies that $A_j |\phi\rangle = \vec{0}$ for $j \neq i$. Thus, 0 in the spectrum of A_j for all $j \neq i$. That is, $\lambda_{\min}(A_j) = 0$ for all $j \neq i$. Our choice of i was arbitrary; thus, for every $j = 1, \ldots, m$ we necessarily have $\lambda_{\max}(A_j) = 1$ and $\lambda_{\min}(A_j) = 0$.

(4) $\Rightarrow$ (5). Fix a measurement basis $\mathcal{B}_\nu = \{A_1, \ldots, A_m\}$ for ν. Decompose $\mathcal{H}$ as $\mathcal{H} = \ker(A_1 - \mathrm{I}) \oplus \mathcal{H}_1$. Because both A_j and $\mathrm{I} - A_j$ are positive, and because $\sum_{j=1}^{m} A_j = \mathrm{I}$, the decomposition $\mathcal{H} = \ker(A_1 - \mathrm{I}) \oplus \mathcal{H}_1$ implies that

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & G_1 \end{pmatrix} \quad \text{and} \quad A_j = \begin{pmatrix} 0 & 0 \\ 0 & G_j \end{pmatrix} \text{ for all } j \geq 2 ,$$

for some positive contractions $G_1, \ldots, G_m \in \mathfrak{B}(\mathcal{H}_1)$ satisfying $\sum_{j=1}^{m} G_j = \mathrm{I}$. Because 1 is an eigenvalue of A_2, the matrix representation of A_2 above shows that $\ker(A_2 - \mathrm{I})$ must lie within the subspace $\mathcal{H}_1$. Thus, we decompose the Hilbert space $\mathcal{H}$ further

as $\mathcal{H} = \ker(A_1 - \mathrm{I}) \oplus \ker(A_2 - \mathrm{I}) \oplus \mathcal{H}_2$ so that

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & K_1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & K_2 \end{pmatrix}, \text{ and } A_j = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & K_j \end{pmatrix}$$

for all $j \geq 3$ and for some positive contractions $K_1, \ldots, K_m \in \mathfrak{B}(\mathcal{H}_2)$ satisfying $\sum_{j=1}^{m} K_j = \mathrm{I}$. As before, A_3 has an eigenvalue 1; the corresponding eigenspace $\ker(A_3 - \mathrm{I})$ must lie in $\mathcal{H}_3$, by the matrix representation of A_3 above. Therefore, an iteration of the argument used to this point produces, after m steps in total, a decomposition of $\mathcal{H}$ as

$$\mathcal{H} = \left(\bigoplus_{j-1}^{m} \ker(A_j - \mathrm{I}) \right) \bigoplus \mathcal{H}_m \,,$$

for a subspace $\mathcal{H}_m \subset \mathcal{H}$, and where, for every j, $A_j = Q_j \oplus Y_j$ for some positive contractions $Y_j \in \mathfrak{B}(\mathcal{H}_m)$ such that $\sum_{j=1}^{m} Y_j = \mathrm{I} \in \mathfrak{B}(\mathcal{H}_m)$ and where each Q_j is the projection with range $\ker(A_j - \mathrm{I})$.

(5) $\Rightarrow$ (1). Suppose that ν' is a quantum probability measure and $\nu \ll_{\mathrm{cl}} \nu'$. The proof of proposition 3.1.6 shows that there exists pairwise-disjoint measurable sets $E_1, \ldots, E_m \in \mathcal{O}(X)$ such that, if $A_j = \nu(E_j)$ for each j, then $\{A_1, \ldots, A_m\}$ is a measurement basis for ν. By assumption (5), there exist a subspace $\mathcal{H}_0 \subset \mathcal{H}$, positive operators $Y_1, \ldots, Y_m \in \mathfrak{B}(\mathcal{H}_0)$, and pairwise-orthogonal projections $Q_1, \ldots, Q_m \in \mathfrak{B}(\mathcal{H}_0^{\perp})$ such that, for every $j = 1, \ldots, m$, $\|Y_j\| \leq 1$ and $A_j = Q_j \oplus Y_j \in \mathfrak{B}(\mathcal{H}_0^{\perp} \oplus \mathcal{H}_0)$.

Therefore, for any finite-dimensional Hilbert space $\mathcal{K}$ and operators $L_1, \ldots, L_m \in \mathfrak{B}(\mathcal{K})$,

$$\begin{aligned}\left\|\sum_{j=1}^{m} A_j \otimes L_j\right\| &= \left\|\sum_{j=1}^{m} A_j \otimes L_j\right\| = \left\|\sum_{j=1}^{m} (Q_j \oplus Y_j) \otimes L_j\right\| \\ &= \max\left\{\left\|\sum_{j=1}^{m} Q_j \otimes L_j\right\|, \left\|\sum_{j=1}^{m} Y_j \otimes L_j\right\|\right\} \\ &= \max_{1\leq j\leq m} \|L_j\|,\end{aligned}$$

since, by lemma 3.3.4,

$$\left\|\sum_{j=1}^{m} Q_j \otimes L_j\right\| = \max_{1\leq j\leq m} \|L_j\| \quad \text{and} \quad \left\|\sum_{j=1}^{m} Y_j \otimes L_j\right\| \leq \max_{1\leq j\leq m} \|L_j\|.$$

Therefore, by lemma 3.3.4 again,

$$\left\|\sum_{j=1}^{m} \nu'(E_j) \otimes L_j\right\| \leq \max_{1\leq j\leq m} \|L_j\| = \left\|\sum_{j=1}^{m} \nu(E_j) \otimes L_j\right\|.$$

Hence, theorem 3.3.2 implies that $\nu' \ll_{\text{cl}} \nu$. ■

3.5 Observations and Applications

3.5.1 Clean 1-0 measurements

A *1-0 measurement* is one in which the sample space of outcomes is given by $X = \{0, 1\}$ and the space of events by $\mathcal{O}(X) = \{\emptyset, \{0\}, \{1\}, X\}$. A direct application of theorem 3.4.3 leads to the following equivalent statements for a quantum probability measure ν on $(X, \mathcal{O}(X))$:

1. ν is clean;

2. $\lambda_{\min}(\nu(\{0\})) = 0$ and $\lambda_{\max}(\nu(\{0\})) = 1$.

This spectral condition above also appears in another notion of optimality for quantum measurements. By the notation $\nu_2 \ll_{\mathrm{f}} \nu_1$ one means that ν_2 is a *fuzzy version* of ν_1, which is to say that there is a confidence mapping Λ on $\mathcal{O}(X)$ such that $\mathrm{Tr}\,(\rho\nu_2(E)) = \mathrm{Tr}\,(\rho\nu_1(\Lambda(E)))$ for all measurable sets E and density operators ρ [Hei05]. With respect to this partial order, ν is optimal if $\nu \ll_{\mathrm{f}} \nu'$ implies $\nu' \ll_{\mathrm{f}} \nu$. In the case where $X = \{0, 1\}$ and $\mathcal{O}(X)$ is the power set of X, a 1-0 quantum measurement ν is optimal with respect to $\ll_{\mathrm{f}}$ if and only if $\nu(\{0\})$ satisfies the spectral property (2) above [Hei05, Proposition 3]. Hence, we deduce that a 1-0 measurement is clean if and only if it is optimal with respect to the fuzzy ordering $\ll_{\mathrm{f}}$.

3.5.2 Clean qubit measurements are projective

Furthermore, if $\dim \mathcal{H} = 2$, then theorem 3.4.3 implies that ν is clean if and only if ν is a projection-valued measure. In contrast, a qubit measurement ν with sample space $X = \{x_1, \ldots, x_n\}$ is clean in the original sense of [BKD+05, Kah07] if and only if $\nu(\{x_j\})$ has rank-1 for every $j = 1, \ldots, n$ [BKD+05, Theorem 11.2], [Kah07]. Thus, the stricter criteria for cleanness used herein leads to a smaller class of clean quantum measurements than was found in [BKD+05, Kah07].

3.5.3 Quantity of information versus quality of information

It was noted by Buscemi *et al* in [BKD+05] that in passing to cleaner quantum measurements there is a trade off between the quantity of information and quality of information afforded by such measurements. As a specific example, we indicate below that quantum measurements that yield the greatest amount of information are never clean.

A quantum probability measure ν on $(X, \mathcal{O}(X))$ with values in $\mathfrak{B}(\mathcal{H})$ is *informationally complete* if, for any two states $\rho_1, \rho_2 \in \mathfrak{B}(\mathcal{H})_t$, the equation

$$\mathrm{Tr}\,(\rho_1 \nu(E)) \;=\; \mathrm{Tr}\,(\rho_2 \nu(E))$$

holds for every event $E \in \mathcal{O}(X)$ if and only if $\rho_1 = \rho_2$.

The next result is known, but we are unaware of a reference to its proof; therefore, we include a proof of the proposition here.

Proposition 3.5.1. *The following statements are equivalent for a quantum probability measure $\nu : \mathcal{O}(X) \to \mathfrak{B}(\mathcal{H})$:*

1. *ν is informationally complete;*

2. *$\mathcal{T}_\nu = \mathfrak{B}(\mathcal{H})$.*

Proof. Let $\mathcal{T}_\nu = \mathrm{span}_{\mathbb{C}}\{\nu(\mathrm{E}) \,|\, \mathrm{E} \in \mathcal{O}(\mathrm{X})\}$ and consider the subspace $\mathcal{T}_\nu{}^{\perp} \subset \mathfrak{B}(\mathcal{H})$ given by

$$\{z \in \mathfrak{B}(\mathcal{H}) \,|\, \mathrm{Tr}(z^\dagger y) = 0 \text{ for all } y \in \mathcal{T}_\nu\}.$$

Observe that $\mathcal{T}_\nu{}^\perp = \{z \in \mathfrak{B}(\mathcal{H}) \mid \mathrm{Tr}(z^\dagger \nu(E)) = 0 \text{ for all } E \in \mathcal{O}(X)\}$. Thus, $\mathfrak{B}(\mathcal{H}) = \mathcal{T}_\nu \oplus \mathcal{T}_\nu{}^\perp$. We claim that $\mathcal{T}_\nu{}^\perp = \{0\}$.

Choose $z \in \mathcal{T}_\nu{}^\perp$. Observe that $y \in \mathcal{T}_\nu$ implies that $y^\dagger \in \mathcal{T}_\nu$, and so $z \in \mathcal{T}_\nu{}^\perp$ implies $z^\dagger \in \mathcal{T}_\nu{}^\perp$. Thus, $\Re(z) = \frac{1}{2}(z + z^\dagger)$ and $\Im(z) = \frac{1}{2i}(z - z^\dagger)$ are elements of $\mathcal{T}_\nu{}^\perp$. Therefore, without loss of generality, we assume that $z^\dagger = z$. In $\mathfrak{B}(\mathcal{H})$ every Hermitian operator is a difference of positive operators $z = a - b$ for some $a, b \in \mathfrak{B}(\mathcal{H})_+$. Because $1 = \nu(X) \in \mathcal{T}_\nu$, $\mathrm{Tr}(z) = \mathrm{Tr}(z1) = 0$, whence $\mathrm{Tr}(a) = \mathrm{Tr}(b)$. Let $\lambda = \mathrm{Tr}(a) = \mathrm{Tr}(b)$. If $\lambda = 0$, then $a = b = 0$ because the trace is a faithful positive linear functional, which implies $z = 0$. If $\lambda \neq 0$, then let $\rho_1 = \frac{1}{\lambda}a$ and $\rho_2 = \frac{1}{\lambda}b$ to obtain $\frac{1}{\lambda}z = \rho_1 - \rho_2 \in \mathcal{T}_\nu{}^\perp$, where $\rho_1, \rho_2 \in \mathrm{S}(\mathcal{H})$. For every $E \in \mathcal{O}(X)$, $0 = \mathrm{Tr}\left(\frac{1}{\lambda} z\nu(E)\right) = \mathrm{Tr}((\rho_1 - \rho_2)\nu(E))$. Because ν is informationally complete, $\rho_1 = \rho_2$, and so $a = b$ and, thus, $z = 0$.

Conversely, assume $\mathcal{T}_\nu = \mathfrak{B}(\mathcal{H})$ and suppose $\mathrm{Tr}(\rho_1 \nu(E)) = \mathrm{Tr}(\rho_2 \nu(E))$ for every $E \in \mathcal{O}(X)$. Since $\mathcal{T}_\nu = \mathfrak{B}(\mathcal{H})$, this is equivalent to $\mathrm{Tr}(\rho_1 x) = \mathrm{Tr}(\rho_2 x)$ for every $x \in \mathfrak{B}(\mathcal{H})$. It follows immediately that $\rho_1 = \rho_2$, hence ν is informationally complete. ■

Corollary 3.5.2. *If a quantum probability measure ν is informationally complete, then $|\mathcal{O}(X)| \geq d^2$.*

Proof. Note that $d^2 = \dim(\mathfrak{B}(\mathcal{H})) = \dim\left(\mathrm{span}_{\mathbb{C}}\{\nu(\mathrm{E}) \mid \mathrm{E} \in \mathcal{O}(\mathrm{X})\}\right)$ and so there are at least d^2 subsets $E_j \in \mathcal{O}(X)$ for which $\{\nu(E_j)\}_j$ is a spanning set for $\mathfrak{B}(\mathcal{H})$. From such a set a basis must exist. Thus, there are at least d^2 distinct Borel sets E_j. ■

Corollary 3.5.3. *If a quantum probability measure ν is informationally complete,*

then

$$\{y \in \mathfrak{B}(\mathcal{H}) \,|\, \nu(E)y = y\nu(E) \text{ for all } E \in \mathcal{O}(X)\} = \{\lambda \cdot \mathrm{I} \,|\, \lambda \in \mathbb{C}\}.$$

Proof. Let $\mathcal{N} = \{y \in \mathfrak{B}(\mathcal{H}) \,|\, \nu(E)y = y\nu(E) \text{ for all } E \in \mathcal{O}(X)\}$, which is a von Neumann subalgebra of $\mathfrak{B}(\mathcal{H})$. Thus, $\mathcal{N}$ is spanned by its projections. Therefore, it is enough to prove that if $p \in \mathcal{N}$ is a projection, then $p = 0$ or $p = 1$.

Select an arbitrary projection $p \in \mathcal{N}$ and without loss of generality assume $p \neq 1$. Thus, there exists a nonzero $\xi \in \mathrm{ran}(1 - \mathrm{p})$. Choose any $\eta \in \mathrm{ran(p)}$. Because $\mathfrak{B}(\mathcal{H})$ is transitive, there is a nonzero operator $z \in \mathfrak{B}(\mathcal{H})$ such that $z\xi = \eta$. Because ν is informationally complete, proposition 3.5.1 implies that $z = \sum_{j=1}^{k} \alpha_j \nu(E_j)$ for some $\alpha_1, ..., \alpha_k \in \mathbb{C}$ and $E_1, ..., E_k \in \mathcal{O}(X)$. Thus,

$$\begin{aligned} \eta &= p(\eta) = pz(1-p)\xi \\ &= \left(\sum_{j=1}^{k} \alpha_j \nu(E_j) p\right)(1-p)\xi \\ &= zp(1-p)\xi = 0. \end{aligned}$$

Thus every vector η in the range of p is zero, whence $p = 0$. ∎

Proposition 3.5.4. *No informationally complete quantum measurement with values in $\mathfrak{B}(\mathcal{H})$, where $\mathcal{H}$ is a Hilbert space of finite dimension $d \geq 2$, is clean.*

Proof. By proposition 3.5.1, if ν is informationally complete, then the only operators that commute with every effect $\nu(E)$, $E \in \mathcal{O}(X)$, are the scalar operators. On the other hand, statement (5b) of theorem 3.4.3 asserts that if ν is clean and $\dim \mathcal{H} \geq 2$, then there is a non-trivial projection in the commutant of $\mathcal{T}_\nu$. ∎

Chapter 4

Majorization and Trumping

The idea of majorization was first introduced by R. Muirhead (a mathematician/physicist/engineer) in 1903 and developed by Lorenz starting in 1905. Lorenz was an economist interested in inequalities of wealth and inequalities of income. A good reference for majorization, including its history and references to earlier works, is [MOA].

Recent work by Lo and Propescu [LP01] on local operations and classical communication has been used by Nielsen [Nie99] to link quantum entanglement with majorization, thus enabling the use of deep and plentiful results of majorization to gain further insight into comparisons and measures of entanglement. Jonathan and Plenio [JP99] introduced the idea of *trumping*, a generalization of majorization which allows us to make use of majorization theory for a larger number of vectors.

We introduce these topics as well as another generalization of majorization, called

power majorization, and discuss the link between trumping and power majorization. We prove some results on the geometric structure of the set of vectors that are power majorized by a fixed vector y. These results can be found in [KPP12]; we give further details herein.

Finally, we approach the notion of trumping through the use of general Dirichlet polynomials, Mellin transforms, and completely monotone sequences. This is accomplished via convex sequence inequality results by [Mer05, Gav05, Nie05], which allow us to link an inequality for hyper-convex sequences with zeros of a general Dirichlet polynomial and the complete monotonicity of a quotient involving said polynomial. Our result [PP13] can be seen as a succinct generalization of a major result of Turgut [Tur07], which gives alternate conditions for trumping. Through some analysis and reinterpretation, we obtain the result of [Tur07] as a corollary.

4.1 Majorization

Let $x = (x_1, x_2, ..., x_d) \in \mathbb{R}^d$ and let $x^{\downarrow} = (x_1^{\downarrow}, x_2^{\downarrow}, ..., x_d^{\downarrow}) \in \mathbb{R}^d$ be the vector consisting of the elements of x reordered so that

$$x_1^{\downarrow} \geq x_2^{\downarrow} \geq x_3^{\downarrow} \geq ... \geq x_d^{\downarrow}.$$

Similarly $x^{\uparrow} = (x_1^{\uparrow}, x_2^{\uparrow}, ..., x_d^{\uparrow}) \in \mathbb{R}^d$ is the vector consisting of the elements of x reordered so that $x_1^{\uparrow} \leq x_2^{\uparrow} \leq x_3^{\uparrow} \leq ... \leq x_d^{\uparrow}$.

Definition 4.1.1. *Let $x = (x_1, x_2, ..., x_d), y = (y_1, y_2, ..., y_d) \in \mathbb{R}^d$. We say that x is majorized by y, written $x \prec y$, if*

$$\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow} \quad 1 \leq k \leq d,$$

with equality when $k = d$.

If equality does not necessarily hold when $k = d$, we say that x is *sub-majorized* by y and we write $x \prec_w y$, where the w stands for "weak".

A similar definition holds if we order the components of the vectors in *non-decreasing* order: x is majorized by y if

$$\sum_{j=1}^{k} x_j^{\uparrow} \geq \sum_{j=1}^{k} y_j^{\uparrow} \quad 1 \leq k \leq d,$$

with equality when $k = d$. If equality does not necessarily hold when $k = d$, we say that x is *super-majorized* by y and we write $x \prec^w y$. Note that sub- and super-majorization are not equivalent in general.

Consider a linear transformation $T : \mathbb{R}^d \to \mathbb{R}^d$ having the following matrix representation:

$$T = t\,\mathrm{id} + (1-t)P, \tag{4.1}$$

where $0 \leq t \leq 1$ and P is any permutation matrix that simply interchanges two coordinates (and leaves all other coordinates unchanged). Thus for all $y \in \mathbb{R}^d$, we have

$$T(y) = (y_1, \ldots, y_{j-1}, ty_j + (1-t)y_k, y_{j+1}, \ldots, (1-t)y_j + ty_k, y_{k+1}, \ldots y_d),$$

for some $j, k \in \{1, \ldots, d\}$.

This linear transformation T is known as a *T-transform*. In the field of economics, B. C. Arnold is credited for coining the term "Robin Hood transfer" [Arn87], as such a transfer of t percent of the wealth or income of the rich to the poor effectively diminishes inequality in a population.

The following alternative characterizations of majorization are well-known (see [Mui03], [HLP52]).

Theorem 4.1.2. *Let $x, y \in \mathbb{R}^d$. The following are equivalent:*

(i) $x \prec y$;

(ii) $x = T_1 \cdots T_r y$, *where T_i is a T-transform for all $i \in \{1, \ldots r\}$ and $r < d$;*

(iii) $\sum_{i=1}^d \phi(x_i) \leq \sum_{i=1}^d \phi(y_i)$ *for all convex functions $\phi : \mathbb{R} \to \mathbb{R}$.*

For sub- and super-majorization, we have that $x \prec_w y \Leftrightarrow \sum_{i=1}^d \phi(x_i) \leq \sum_{i=1}^d \phi(y_i)$ for all increasing convex functions $\phi : \mathbb{R} \to \mathbb{R}$ and $x \prec^w y \Leftrightarrow \sum_{i=1}^d \phi(x_i) \leq \sum_{i=1}^d \phi(y_i)$ for all decreasing convex functions $\phi : \mathbb{R} \to \mathbb{R}$.

We now present the result of Lo and Propescu, that simplifies LOCC in that it is sufficient to consider only one-way (and not two-way) communication between Alice and Bob. This theorem was used to prove the main result in [Nie99].

Theorem 4.1.3. *[LP01] Let $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ be a state shared by Alice and Bob and suppose Bob has a POVM $\{P_\ell^B\}$. Then there exists a POVM $\{P_\ell^A\}$ for Alice and, for each outcome ℓ, a tensor product of local unitary transformations $U_\ell^A \otimes U_\ell^B$ such that, for each ℓ,*

$$(\mathrm{id} \otimes P_\ell^B)|\psi\rangle = (U_\ell^A \otimes U_\ell^B)(P_\ell^A \otimes \mathrm{id})|\psi\rangle.$$

Physically, it follows from this theorem that an arbitrary protocol transforming the state $|\psi^A\rangle$ to the state $|\psi^B\rangle$ using local operations and two-way classical communication can be simulated by a one-way communication protocol from Alice to Bob.

Theorem 4.1.4. *(Uhlmann's theorem) Let ρ, ρ' be density matrices such that $\rho' = \sum_i p_i U_i \rho U_i^\dagger$, where $\{p_i\}$ form a probability distribution and the U_i are unitary. Then the vector of eigenvalues of ρ' is majorized by the vector of eigenvalues of ρ: $\lambda_{\rho'} \prec \lambda_\rho$.*

To state the result of [Nie99] linking entanglement with majorization we will use the notation found therein. Suppose $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. We denote by $\rho_\psi \equiv \mathrm{tr}_B(|\psi\rangle\langle\psi|)$ the state of Alice's system, and λ_ψ the vector of eigenvalues of ρ_ψ. We say that $|\psi\rangle \to |\phi\rangle$, read "$|\psi\rangle$ transforms to $|\phi\rangle$" if $|\psi\rangle$ can be transformed into $|\phi\rangle$ by local operations and potentially unlimited two-way classical communication. Then we have:

Theorem 4.1.5. *[Nie99] We can transform $|\psi\rangle$ to $|\phi\rangle$ using local operations and classical communication if and only if λ_ψ is majorized by λ_ϕ. That is,*

$$|\psi\rangle \to |\phi\rangle \text{ iff } \lambda_\psi \prec \lambda_\phi. \tag{4.2}$$

This theorem is significant because any entangled state can be transformed via LOCC into a state that is less or equally entangled. Thus this theorem states that $|\psi\rangle$ is at least as entangled as $|\phi\rangle$ if and only if the eigenvalues of $|\psi\rangle$ are majorized by the eigenvalues of $|\phi\rangle$.

Nielsen's result makes repeated use of the Schmidt decomposition 1.2.2 and basic eigenvalue calculations: in particular, if $|x\rangle = \sum_i \sqrt{\lambda_i}|i_A\rangle \otimes |i_B\rangle$, then ρ_x has eigenvalues λ_i. We write $|x\rangle \sim |y\rangle$ if $|x\rangle$ and $|y\rangle$ are the same up to local unitary operations by Alice and Bob. The Schmidt decomposition implies that $|x\rangle \sim |y\rangle$ if and only if ρ_x and ρ_y have the same spectrum of eigenvalues.

Nielsen's theorem highlights the importance of majorization in quantum information theory. Unfortunately, there are many vectors that are *incomparable* in the sense that one vector does not majorize the other.

Suppose $x_i, y_i \in \mathbb{N}$ for all $i : 1 \leq i \leq d$. Let

$$a_n = \#\{i : y_i = n\} - \#\{i : x_i = n\} \quad \forall n \in \mathbb{N}. \tag{4.3}$$

Then $\sum_{n=1}^{\infty} a_n \Psi(n) = \sum_{i=1}^{d} \Psi(y_i) - \sum_{i=1}^{d} \Psi(x_i)$ and hence $x \prec y$ if and only if $\sum_{n=1}^{\infty} a_n \Psi(n) \geq 0$ for all convex functions $\Psi : \mathbb{R} \to \mathbb{R}$ or equivalently $\sum_{n=1}^{\infty} a_n \phi(n) \geq 0$ for all convex sequences $\{\phi(n)\}_{n=1}^{\infty}$. This is a classical description of majorization. We now consider general d-tuples of real numbers $\{a_n\}$ for which $\sum_{n=1}^{\infty} a_n \Psi(n) \geq 0$ and do not restrict ourselves only to the integer values of $\{a_n\}$ which arise from equation (4.3).

The following characterization of such sequences $\{a_n\}$ was given in [Gav05], which built on the work of [Mer05].

Theorem 4.1.6. *Let $a_0, \ldots, a_d$ be $d+1$ fixed real numbers such that $\sum_{i=1}^{d} a_i^2 > 0$. The inequality*

$$\sum_{i=0}^{d} a_i \phi(i) \geq 0$$

holds for every convex sequence $\{\phi(i)\}$ if and only if the polynomial $\sum_{i=1}^{d} a_i t^i$ has a double root at $t = 1$ and all the coefficients of the polynomial

$$\frac{\sum_{i=0}^{d} a_i t^i}{(t-1)^2} = \sum_{i=0}^{d-2} c_i t^i \tag{4.4}$$

are nonnegative.

As a simple example, consider the vectors $x = (5, 5, 5, 5), y = (2, 2, 6, 10)$. It can easily be verified that $x \prec y$. Consider $n \in \mathbb{N}$ for which $x_i = n$ or $y_i = n$ and let a_n be defined as in equation (4.3). We have

$$\sum_{\{n \,|\, x_i = n \text{ or } y_i = n\}} a_n t^n = \sum_i t^{y_i} - \sum_i t^{x_i} = t^{10} + t^6 - 4t^5 + 2t^2.$$

We can check that this polynomial has a double root at $t = 1$ by dividing by $t^2 - 2t + 1$ to obtain the polynomial $t^8 + 2t^7 + 3t^6 + 4t^5 + 6t^4 + 4t^3 + 2t^2$, which has no further roots at $t = 1$. Furthermore, we can clearly see all the coefficients of this new polynomial are non-negative. By theorem 4.1.6, we have $\sum_{0 \leq i \leq m} a_i \phi(i) \geq 0$ for every convex sequence $\{\phi(i)\}$ and hence $x \prec y$.

We note that theorem 4.1.6 can be seen to be a reformulation of the following earlier result.

Lemma 4.1.7. *[KN63] The inequality* $\sum_{0\le i\le m} a_i\phi(i) \ge 0$ *for all convex functions* ϕ *is equivalent to the following three conditions.*

(i) $\sum_{0\le i\le m} a_i = 0$*;*

(ii) $\sum_{0\le i\le m} ia_i = 0$*;*

(iii) $\sum_{j=0}^{k}\sum_{i=0}^{j} a_i \ge 0 \quad 0 \le k \le m.$

To see that theorem 4.1.6 is equivalent to lemma 4.1.7, we first note that condition (i) of the lemma is equivalent to the polynomial $\sum_i a_i t^i$ having a root at $t = 1$. Condition (ii) of the lemma is precisely the derivative of the polynomial $\sum_i a_i t^i$ evaluated at $t = 1$, so conditions (i) and (ii) are equivalent to the polynomial $\sum_i a_i t^i$ having a double root at $t = 1$. Finally, re-writing equation (4.4) in terms of the a_i, we find $a_i = c_i - 2c_{i-1} + c_{i-2}$, where we let $c_{-1} = c_{-2} = 0$. We then find $\sum_{j=0}^{k}\sum_{i=0}^{j} a_i = c_0 + (c_1 - c_0) + (c_2 - c_1) + (c_3 - c_2) + \cdots + (c_k - c_{k-1}) = c_k$. Thus condition (iii) of the lemma is precisely the statement that all c_i are non-negative.

Lemma 4.1.7 is itself a special case of a more general result from the same paper.

Lemma 4.1.8. *[KN63] Let* μ *be a signed measure on* $[a, b]$*. Let* $\mu_1(x) = \int_a^x d\mu$ *and* $\mu_2(x) = \int_a^x \mu_1(t)\,dt$*. Then the inequality* $\int_a^b \phi\,d\mu \ge 0$ *is satisfied for all convex functions* ϕ *defined on* $[a, b]$ *if and only if the following three conditions hold:*

(i) $\int_a^b d\mu = 0$*;*

(ii) $\int_a^b x\,d\mu = 0$*;*

(iii) $\mu_2(x) \geq 0 \quad \forall x \in [a, b]$.

By considering any measure μ with support strictly on some set of integers $\{i\}$ having associated point masses a_i, we then obtain the discrete analogue—lemma 4.1.7 considered above. Lemma 4.1.8 can also be used to characterize finite sequences $\{a_i\}_{i \in I}$ where I is a finite set of real numbers (and not necessarily integers) such that $\sum_{i \in I} a_i \phi(i) \geq 0$ for all convex functions ϕ by considering the signed measure μ supported on I with $\mu(i) = a_i$.

4.2 Trumping

Trumping is a generalization of majorization in that, given two incomparable vectors x and y, it is sometimes possible [JP99] to find a third vector z, having all positive components, such that $x \otimes z$ is majorized by $y \otimes z$, thus allowing one to compare a larger set of vectors than was possible using majorization alone. In [JP99], the authors explain the physical relevance of trumping: the third vector z can be seen as a resource that allows one to transform one state into another via local operations and classical information; this third vector z remains unaltered after being used, yet the transformation could not occur without its presence.

Since 2005, finding useful characterizations of trumping has been identified as an open problem in quantum information ([Wer05, Problem 4]). Nielsen's theorem 4.1.5 gives a useful, concise characterization of when one pure bipartite state can

be converted into another pure bipartite state using only LOCC. The goal is to characterize trumping in a similar manner. Specifically: "to give a similarly efficient criterion to decide which pure bipartite states can be converted into each other using a catalyst" [Wer05, Problem 4]. The works of Aubrun-Nechita [AN08], Turgut [Tur07], Daftuar-Klimesh [DK01], and Klimesh [Kli04, Kli07], have all been cited in [Wer05, Problem 4] as partial results toward this open problem.

The *catalyst* z is often taken to have all positive components, which avoids the potential issue of its components summing to zero. Formally, we have:

Definition 4.2.1. *Let* $(x_1, x_2, ..., x_d)$ *and* $(y_1, y_2, ..., y_d)$ *be two* d*-tuples of real numbers. We say that* x *is* trumped *by* y *and write* $x \prec_T y$ *if there exists a unit vector* $z \in \mathbb{R}^n$ *with positive components such that* $x \otimes z \prec y \otimes z$.

Sub-trumping (written $x \prec_{wT} y \Leftrightarrow x \otimes z \prec_w y \otimes z$) and super-trumping ($x \prec_T^w y$) can be defined in an analogous manner to sub- and super-majorization.

Recently, much work has been done on the subject of trumping. In [DK01], it was shown that the dimension of the catalyst is in general unbounded. In [AN08], the authors link trumping with ℓ^p-norm inequalities.

Remark 4.2.2. *If we are attempting to compare two vectors* x *and* y *to see if* x *is majorized or trumped by* y*, we can effectively "delete" corresponding zeros from both vectors without affecting the relation, as adding zero has no affect on a sum. In this way, we can assume without loss of generality that the vector* x *has no zero components. Similarly, if we wish to compare two vectors* $x \in \mathbb{R}^{d_1}$ *and* $y \in \mathbb{R}^{d_2}$*, where*

$d_1 \neq d_2$, we can append zeros to the vector that lives in the smaller dimensional space until the vectors have equal dimension. In this way, we can effectively compare two vectors of different dimensions.

Recall the formula for the von Neumann entropy of a density matrix x with eigenvalues x_i: $\sigma(x) = -\sum_{i=1}^{d} x_i \log x_i$. We follow the convention that $0 \log 0 \equiv 0$, which is precisely what one obtains when considering the limit of $x_i \log x_i$ as x_i approaches 0. Let us also define $A_\nu(x) = \left(\frac{1}{d}\sum_{i=1}^{d} x_i^\nu\right)^{\frac{1}{\nu}}$ for real numbers $\nu \neq 0$. When $\nu = 0$, we obtain the geometric mean $A_0(x) = \left(\prod_{i=1}^{d} x_i\right)^{\frac{1}{d}}$.

We now present the main result from [Tur07].

Theorem 4.2.3. *[Tur07] For two real d-dimensional vectors x and y of non-negative components such that x has non-zero elements and the vectors are distinct (i.e. $x^\uparrow \neq y^\uparrow$), the relation $x \prec_T y$ is equivalent to the following three strict inequalities:*

(T1) $A_\nu(x) > A_\nu(y), \quad \forall \nu \in (-\infty, 1),$

(T2) $A_\nu(x) < A_\nu(y), \quad \forall \nu \in (1, \infty),$

(T3) $\sigma(x) > \sigma(y).$

Note that for dimensions two and three, trumping is equivalent to majorization: Clearly majorization implies trumping, and for the reverse direction, first note that for any dimension d, $x \prec_T y$ implies that we have $x_1^\uparrow \geq y_1^\uparrow$ and $x_d^\uparrow \leq y_d^\uparrow$. Additionally, for an n-dimensional catalyst z, trumping implies

$$\begin{aligned} \sum_{i,j=1}^{d,n} x_i z_j &= \sum_{i,j=1}^{d,n} y_i z_j \\ \Rightarrow \sum_{i=1}^{d} x_i \left(\sum_{j=1}^{n} z_j \right) &= \sum_{i=1}^{d} y_i \left(\sum_{j=1}^{n} z_j \right) \\ \Rightarrow \sum_{i=1}^{d} x_i &= \sum_{i=1}^{d} y_i. \end{aligned}$$

The two-dimensional case becomes obvious. For $d = 3$, by combining $x_3^{\uparrow} \leq y_3^{\uparrow}$ with $x_1^{\uparrow} + x_2^{\uparrow} + x_3^{\uparrow} = y_1^{\uparrow} + y_2^{\uparrow} + y_3^{\uparrow}$, we obtain $x_1^{\uparrow} + x_2^{\uparrow} \geq y_1^{\uparrow} + y_2^{\uparrow}$, and the result follows.

The above holds similarly for sub- and super-trumping implying sub- and super-majorization, respectively, for $d = 2$.

For dimensions greater than 3, trumping does not necessarily imply majorization; that is, there are instances of "non-trivial" trumping. For example, for $d = 4$, we have $x = (0.4, 0.4, 0.1, 0.1) \prec_T (0.5, 0.25, 0.25, 0) = y$, with catalyst $z = (0.6, 0.4)$, yet $x \not\prec y$ [JP99].

The literature on trumping is not entirely consistent with regard to terminology; synonyms for trumping include: entanglement-assisted local transformation (ELQCC) [JP99], entanglement-assisted local operations and classical communication (ELOCC) [AN08], and catalytic majorization [AN08].

In [Kli04, Kli07], Klimesh establishes a theorem showing that trumping is equivalent to a series of inequalities for a family of additive Schur-convex functions. A function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be *Schur-convex* when $x \prec y$ implies $f(x) \leq f(y)$. Similarly, f is Schur concave when $x \prec y$ implies $f(x) \geq f(y)$. A function $f : \mathbb{R}^d \to \mathbb{R}$

is additive if $f(a \otimes b) = f(a) + f(b)$ for all $a, b \in \mathbb{R}^d$.

For a d-dimensional probability vector x, let

$$f_r(x) = \begin{cases} \ln \sum_{i=1}^d x_i^r & (r > 1); \\ \sum_{i=1}^d x_i \ln x_i & (r = 1); \\ -\ln \sum_{i=1}^d x_i^r & (0 < r < 1); \\ -\sum_{i=1}^d \ln x_i & (r = 0); \\ \ln \sum_{i=1}^d x_i^r & (r < 0). \end{cases}$$

If any of the components of x are 0, we take $f_r(x) = \infty$ for $r \leq 0$.

Note that these functions are indeed Schur-convex and additive for all r. Both properties are straightforward to check by calculation, although one must recall that $\sum_{i=1}^d x_i = \sum_{j=1}^n z_j = 1$ in order to show that f_1 is additive.

Theorem 4.2.4. *([Kli04, Kli07]) Let $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ be d-dimensional probability vectors. Suppose that x and y do not both contain components equal to 0 and that $x^\uparrow \neq y^\uparrow$. Then $x \prec_T y$ if and only if $f_r(x) < f_r(y)$ for all real numbers r.*

The conditions in the theorems of Klimesh and Turgut are easily seen to be equivalent. To see this, first note that the function $a \mapsto a^{1/\nu}$ (where $\nu > 1$ and $a \geq 0$), as well as the natural logarithm, are both injective and increasing. Then condition (T2) of theorem 4.2.3 can be seen to be equivalent to $\ln \sum_{i=1}^d x_i^r < \ln \sum_{i=1}^d y_i^r$, which is precisely $f_r(x) < f_r(y)$ for $r > 1$. Condition (T3) is precisely $\sum_{i=1}^d x_i \ln x_i < \sum_{i=1}^d y_i \ln y_i$

(that is, $f_1(x) < f_1(y)$). Furthermore, condition (T1) can be broken down into three cases:

1. $0 < \nu < 1$, in which case we can manipulate the condition to become $\ln \sum_{i=1}^{d} x_i^r > \ln \sum_{i=1}^{d} y_i^r$, giving $f_r(x) < f_r(y)$ for $0 < r < 1$;

2. $\nu = 0$; thus (T1) becomes a geometric mean inequality, which we can simplify as

$$\begin{aligned}
\left(\Pi_{i=1}^{d} x_i\right)^{1/d} > \left(\Pi_{i=1}^{d} y_i\right)^{1/d} &\Leftrightarrow \Pi_{i=1}^{d} x_i > \Pi_{i=1}^{d} y_i \\
&\Leftrightarrow \ln \Pi_{i=1}^{d} x_i > \ln \Pi_{i=1}^{d} y_i \\
&\Leftrightarrow \sum_{i=1}^{d} \ln x_i > \sum_{i=1}^{d} \ln y_i \\
&\Leftrightarrow -\sum_{i=1}^{d} \ln x_i < -\sum_{i=1}^{d} \ln y_i \\
&\Leftrightarrow f_0(x) < f_0(y);
\end{aligned}$$

3. $\nu < 0$, in which case condition (T1) becomes

$$\begin{aligned}
\left(\frac{1}{d}\sum_{i=1}^{d} x_i^{\nu}\right)^{\frac{1}{\nu}} > \left(\frac{1}{d}\sum_{i=1}^{d} y_i^{\nu}\right)^{\frac{1}{\nu}} &\Leftrightarrow \frac{1}{\left(\frac{1}{d}\sum_{i=1}^{d} x_i^{\nu}\right)^{\frac{1}{|\nu|}}} > \frac{1}{\left(\frac{1}{d}\sum_{i=1}^{d} y_i^{\nu}\right)^{\frac{1}{|\nu|}}} \\
&\Leftrightarrow \left(\frac{1}{d}\sum_{i=1}^{d} y_i^{\nu}\right)^{\frac{1}{|\nu|}} > \left(\frac{1}{d}\sum_{i=1}^{d} x_i^{\nu}\right)^{\frac{1}{|\nu|}} \\
&\Leftrightarrow \sum_{i=1}^{d} y_i > \sum_{i=1}^{d} x_i \\
&\Leftrightarrow \ln \sum_{i=1}^{d} y_i > \ln \sum_{i=1}^{d} x_i \\
&\Leftrightarrow f_r(x) < f_r(y), r < 0.
\end{aligned}$$

We now introduce the concept of power majorization, a generalized notion of majorization studied in [Cla84, Ben86, All88] (to name a few), and we connect power majorization with trumping.

Definition 4.2.5. *Let x and y be vectors of non-negative components. We say that x is* power majorized *by y, denoted $x \preceq_p y$, if $x_1^p + \cdots + x_d^p \leq y_1^p + \cdots + y_d^p$ for all $p \geq 1, p \leq 0$ and the inequality switches direction when $0 \leq p \leq 1$. In particular, we note that equality holds when $p = 0, 1$. We define* strict power majorization*, denoted $x \prec_p y$, to be power majorization with strict inequality, and equality if and only if $p = 0, 1$.*

Power majorization is unfortunately not as well-behaved of a partial order on vectors as majorization is, in the following sense: If x is majorized by y and $\sum_{i=1}^d \phi(x_i) = \sum_{i=1}^d \phi(y_i)$ for some strictly convex function ϕ, then $x^\uparrow = y^\uparrow$ (see [FH77], which generalizes a result from [Sch23]). This is not the case with power majorization. Indeed, let $x = (1+\delta, 1+\delta, 7+\delta, 8+\delta), y = (\delta, 3+\delta, 5+\delta, 9+\delta)$ [Ben11, Theorem 2]. With $\delta \approx 0.2952$, this yields $x \preceq_p y$, and $\sum_{i=1}^d \phi(x_i) = \sum_{i=1}^d \phi(y_i)$ for $\phi(c) = c^2$, but $x^\uparrow \neq y^\uparrow$.

Proposition 4.2.6. *Power majorization can be expressed in terms of Klimesh's functionals: Let x and y be vectors in $\mathbb{R}^d$ with positive components. Then $x \preceq_p y$ if and only if $f_r(x) \leq f_r(y)$ for all $r \in \mathbb{R}$.*

Proof. For any positive scalars a, b, $a \leq b \Rightarrow \ln(a) \leq \ln(b)$; it follows immediately

from this basic property and the definition of power majorization that $x \preceq_p y$ implies that $f_r(x) \leq f_r(y)$ whenever $r \neq 0, 1$. Let $g(r) = \sum_{i=1}^d x_i^r - \sum_{i=1}^d y_i^r$. For $r = 0$, consider $\lim_{r\to 0} \frac{g(r)}{r}$, which is a limit of the form $\frac{0}{0}$. Using l'Hôpital's rule, we find

$$\begin{aligned}\lim_{r\to 0} \frac{g(r)}{r} &= \lim_{r\to 0} \frac{g'(r)}{1} \\ &= \lim_{r\to 0} \sum_{i=1}^d x_i^r \ln(x_i) - \sum_{i=1}^d y_i^r \ln(y_i) \\ &= \sum_{i=1}^d \ln(x_i) - \sum_{i=1}^d \ln(y_i) = -f_0(x) + f_0(y) \geq 0.\end{aligned}$$

The fact that this equation is ≥ 0 follows by considering the left- and right-hand side of the limit separately: $g(r) = \sum_{i=1}^d x_i^r - \sum_{i=1}^d y_i^r$ is non-negative on $(0, 1)$ and non-positive on $(-\infty, 0) \cup (1, \infty)$, thus yielding $\frac{g(r)}{r} \geq 0$ for $r \to 0^+$ and $r \to 0^-$. Thus we have $-f_0(x) + f_0(y) \geq 0$, or, $f_0(x) \leq f_0(y)$.

Similarly, by evaluating $\lim_{r\to 1} \frac{g(r)}{r-1}$, we obtain $g'(1) = f_1(x) - f_1(y) \leq 0$. In other words, $f_1(x) \leq f_1(y)$. The converse, namely $f_r(x) \leq f_r(y)$ for all $r \in \mathbb{R}$ implies $x \preceq_p y$, is immediate. ■

We note that if x is strictly power-majorized by y, we will have $f_0(x) \leq f_0(y)$ and $f_1(x) \leq f_1(y)$ but these inequalities may not be strict. An example considered in [Tur07] shows that the third inequality in theorem 4.2.3 does not follow from the other inequalities. That is, strict inequality in $f_1(x) < f_1(y)$ is necessary for trumping.

Proposition 4.2.7. *Let x and y be vectors in $\mathbb{R}^d$ with positive components with $x \prec_p y$, then $x \prec_T y$ provided that $\prod_{i=1}^d x_i \neq \prod_{i=1}^d y_i$ and $\prod_{i=1}^d x_i^{x_i} \neq \prod_{i=1}^d y_i^{y_i}$.*

Indeed,

$$f_0(x) \neq f_0(y) \Leftrightarrow \sum_{i=1}^{d} \ln x_i \neq \sum_{i=1}^{d} \ln y_i \Leftrightarrow \ln \Pi_{i=1}^{d} x_i \neq \ln \Pi_{i=1}^{d} y_i \Leftrightarrow \Pi_{i=1}^{d} x_i \neq \Pi_{i=1}^{d} y_i.$$

The inequality $f_1(x) \neq f_1(y)$ gives us $\prod_{i=1}^{d} x_i^{x_i} \neq \prod_{i=1}^{d} y_i^{y_i}$. Proposition 4.2.7 is especially useful for proving trumping relations between d-tuples of integers.

4.3 Geometry of Trumping and Power Majorization

Consider the sets $S(y) = \{x \in (0,\infty)^d : x \prec y\}$ and $T(y) = \{x \in (0,\infty)^d : x \prec_T y\}$. This notation was used in [DK01]. For our purposes, we similarly define $P(y) = \{x \in (0,\infty)^d : x \preceq_p y\}$. The geometric properties of $S(y)$ have been extensively studied (see [GG77, Dah10]). After trumping was introduced in [JP99], several interesting properties of $T(y)$ were found in [DK01]; however, much is still unknown. The geometric properties of $P(y)$ appear not to have been studied in the literature. In this section we will study the geometric relationship between $T(y)$ and $P(y)$. It is clear that $S(y) \subseteq T(y) \subseteq P(y)$. We begin with the following closure relation [KPP12].

Theorem 4.3.1. *Let y be a d-vector all of whose components are positive, then the set $P(y)$ is the closure in $\mathbb{R}^d$ of the set $T(y)$.*

Proof. Since all of the functions $f_r(x)$ are continuous on $(0,\infty)^d$, $T(y) \subseteq P(y)$ implies that $\overline{T(y)} \subseteq P(y)$. Indeed, let $x \in \overline{T(y)}$. Then there exists a sequence (x_n) in $T(y)$

such that $x_n \to x$ as $n \to \infty$. Since the functions f_r are continuous for all $r \in \mathbb{R}$, it follows that $f_r(x_n) \to f_r(x)$. For some fixed r, choose n such that $|f_r(x_n) - f(x)| < \epsilon$. We then have:

$$\begin{aligned} f_r(y) - f_r(x_n) &> 0 \text{ since } (x_n) \subset T(y) \\ \Rightarrow f_r(y) - f_r(x) + f_r(x) - f_r(x_n) &> 0 \\ \Rightarrow f_r(y) - f_r(x) + \epsilon &> 0. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we obtain $f_r(y) - f_r(x) \geq 0$, which, by proposition 4.2.6, implies $x \in P(y)$.

Now suppose $x \in P(y)$. If all the entries of x are the same, then tensoring x with a catalyst will not change the majorization inequality. In other words, this is a case of trivial trumping: x is in fact an element of $S(y) \subseteq T(y)$. If there is at least one entry x_j of x that is different from all other entries of x, then there exists a vector $x' \neq x$ where x' is some permutation of x. Now let $z(t) = tx + (1-t)x'$. The function f_r is either strictly convex or is the logarithm of a strictly convex function; hence, if $t \in (0,1)$, we have $f_r(z(t)) < f_r(x)$ by theorem 4.1.2. But $f_r(x) \leq f_r(y)$ by assumption (again using proposition 4.2.6), so we have $f_r(z(t)) < f_r(y)$. That is, $z(t) \in T(y)$. As $x = \lim_{t \to 0^+} z(t)$, it follows that $x \in \overline{T(y)}$. ■

It is well-known that $T(y)$ is convex (see, e.g. [DK01]). Since $P(y) = \overline{T(y)}$, it follows that $P(y)$ is a convex set. Thus the set $P(y)$ is a closed convex set, and so Kreĭn-Milman Theorem implies that $P(y)$ is the convex hull of its extreme points.

We would therefore like to characterize the extreme points of $P(y)$.

Rado's theorem for majorization [MOA, p. 34] states that x is majorized by y if and only if x lies in the convex hull of vectors Py, where P is any permutation matrix. Put another way, $x \prec y$ if and only if x is contained in the convex hull of $(y_{\sigma(1)}, \ldots, y_{\sigma(d)})$, where σ is any permutation on d elements. We will derive an analogue of Rado's theorem for power majorization.

We require the following lemma.

Lemma 4.3.2. *[DK01] Let x and y be d-vectors all of whose components are positive. Let $x_{\max}$, $x_{\min}$, $y_{\max}$, $y_{\min}$ be the values of the maximum and minimum entries of x and y respectively. Suppose x is a boundary point of $T(y)$, then either $x_{\max} = y_{\max}$ or $x_{\min} = y_{\min}$.*

We now have the necessary tools to prove the following theorem [KPP12]; we add details that were omitted in the paper.

Theorem 4.3.3. *Let y be a d-vector all of whose components are positive and let $x \in P(y)$. Then the following are equivalent:*

1. *x is an extreme point of $P(y)$.*
2. *$f_r(x) = f_r(y)$ for some $r \in \mathbb{R}$.*
3. *Either x is not trumped by y or there exists some d-by-d permutation matrix P such that $x = Py$.*

Proof. The equivalence of (2) and (3) are the results of Turgut and Klimesh, so we prove the equivalence of (1) and (2). To show (2) implies (1) let $f_r(x) = f_r(y)$ for some $r \in \mathbb{R}$. Suppose by contradiction that we can write x as a non-trivial convex combination of vectors in $P(y)$; namely $x = \sum_i \lambda_i x_i$, where λ_i form a probability distribution with $\lambda_i \neq 0, 1$ and $x_i \in P(y)$ for all i. Since $f_r(x)$ is either strictly convex or is the logarithm of a strictly convex function for any r, we have $f_r(x) = f_r(\sum_i \lambda_i x_i) < \sum_i \lambda_i f_r(x_i)$. By proposition 4.2.6 we have $f_r(x_i) \leq f_r(y)$ for all i, from which it follows that $f_r(x) < \sum_i \lambda_i f_r(y) = f_r(y)$, a contradiction. Thus x is indeed an extreme point.

We now show that (1) implies (3) by proving the contrapositive. Our proof is by induction on d. We note that since $T(y) \subseteq P(y)$, if x is an interior point of $T(y)$, it cannot be an extreme point of $P(y)$. For the base case of $d = 2$, if $x \prec_T y$ and x is in the interior of $T(y)$, then we are done. So, assume x is on the boundary of $T(y)$. Lemma 4.3.2 states that either $x_{max} = y_{max}$ or $x_{min} = y_{min}$. Since x and y each have only two components, if one of these statements is true, it fixes the other statement to be true (since $x_{max} + x_{min} = y_{max} + y_{min}$). Thus $x^{\uparrow} = y^{\uparrow}$. But this implies there exists some d-by-d permutation matrix P such that $x = Py$, contrary to our assumption. Thus x cannot be a boundary point of $T(y)$. We conclude that x is not an extreme point of $P(y)$.

Now suppose (1) implies (3) for $d = n$ and let x, y be $(n+1)$-tuples of positive numbers. Suppose $x \prec_T y$ and x is not a rearrangement of y. Again, if x is in the

interior of $T(y)$, it cannot be an extreme point of $P(y)$. So suppose x is a boundary point of $T(y)$, then by lemma 4.3.2, we must have $x_i = y_j$ for some $1 \leq i, j \leq n+1$. Let $\tilde{x}$ and $\tilde{y}$ be the n-tuples formed by removing x_i and y_j from x and y respectively. Then $\tilde{x} \prec_T \tilde{y}$ and by our induction hypothesis, $\tilde{x}$ is not an extreme point of $P(\tilde{y})$. Hence there exists $w, z \in P(\tilde{y})$, $w^{\uparrow} \neq z^{\uparrow}$ such that $\tilde{x} = \lambda w + (1-\lambda)z$ for some $\lambda \in (0,1)$. Then we can write the full vector x as $x = \lambda(w_1, w_2, ..., w_{i-1}, x_i, w_i, ..., w_n) + (1-\lambda)(z_1, z_2, ..., z_{i-1}, x_i, z_i, ..., z_n)$. Since the latter two vectors are in $P(y)$, our result follows. ■

4.4 Examples of Trumping

There are many examples of non-trivial power majorization in the literature; that is, examples of vectors x, y such that $x \preceq_p y$ but x is not majorized by y. We will look at an infinite family of such pairs x and y considered in [Ben10], where the author made use of [BJ00, Lemma 1 and Theorem 4]. We show that these examples are in fact examples of trumping by modifying the very results of [BJ00] used in [Ben10] in order to suit our purposes. In particular, if we include strictness in one of the hypotheses of [BJ00, Lemma 1], we obtain strictness in the conclusion of the lemma (see lemma 4.4.1, below); similarly, if we include strictness in the hypothesis of [BJ00, Theorem 4], we obtain strictness in the conclusion (see theorem 4.4.2, below). The proofs of the modified versions of these two results follow the original proofs save for these slight modifications, and so will be omitted.

Lemma 4.4.1. *Suppose that $a < b < c < d$ and that p, q, r are non-negative numbers. Suppose further that*

$$b - a \leq d - c \text{ and } p \lneq r$$

$$q(c - b) = p(b - a) + r(d - c).$$

Let g be a convex function on $[a, d]$ with $g(c) \gneq g(b)$. Then

$$q \int_b^c g \lneq p \int_a^b g + r \int_c^d g.$$

Let $M_n(f)$ denote the midpoint Riemann sum with n subintervals for $\int_a^b f(t)dt$. In particular, when $[a, b] = [0, 1]$, we have $M_n(f) = \frac{1}{n} \sum_{r=1}^{n} f(\frac{2r-1}{2n})$. We can use lemma 4.4.1 to prove a strengthened version of [BJ00, Theorem 4] following the original proof using this new lemma.

Theorem 4.4.2. *Suppose that f' is either convex or concave on $[a, b]$. If f is strictly convex, then $M_n(f)$ strictly increases with n; if f is strictly concave, then $M_n(f)$ strictly decreases with n.*

Example 4.4.3. *In [Ben10], Bennett studies the following system of inequalities:*

$$\frac{1^p}{1^{p+1}} \leq \frac{1^p + 3^p}{2^{p+1}} \leq \frac{1^p + 3^p + 5^p}{3^{p+1}} \leq \cdots \frac{1^p + 3^p + \cdots + (2n-1)^p}{n^{p+1}} \leq \cdots \tag{4.5}$$

for $p > 1, p < 0$ and reversed for $0 < p < 1$.

As in [Ben10], we can show that this system leads to power majorization by considering a particular inequality (say, the second one) and cross-multiplying, yielding

$3^p + 3^p + 3^p + 9^p + 9^p + 9^p \leq 2^p + 2^p + 6^p + 6^p + 10^p + 10^p$ *for* $p > 1, p < 0$ *and reversed for* $0 < p < 1$. *Letting* $x = (3, 3, 3, 9, 9, 9)$ *and* $y = (2, 2, 6, 6, 10, 10)$, *we obtain* $x \prec_p y$. *It was noted in [Ben10] that the power majorizations found in this way are* not majorizations, *save for the first inequality. To see why this is the case, we consider the* n*-th inequality of the system:*

$$\frac{1^p + 3^p + \cdots + (2n-1)^p}{n^{p+1}} \leq \frac{1^p + 3^p + \cdots + (2n+1)^p}{(n+1)^{p+1}},$$

which yields, upon cross-multiplying, $n+1$ *copies of* $(n+1)^p + (3(n+1))^p + \cdots + \big((2n-1)(n+1)\big)^p$ *being less than (for appropriate* p*)* n *copies of* $n^p + (3n)^p + \cdots + (n(2n+1))^p$. *These sums are already arranged in increasing order. We note that the first* n *inequalities of majorization are met, but the* $(n+1)$*-th inequality is flipped:*

$$\begin{aligned} n+1 &> n \\ 2(n+1) &> 2n \\ &\vdots \\ n(n+1) &> n^2 \\ (n+1)(n+1) &< n^2 + 3n, \end{aligned}$$

for all $n > 1$, *thus we see explicitly where majorization fails.*

The system considered in [Ben10] is written with non-strict inequalities; however, we claim that the inequalities are strict for $p \neq 0, 1$, *hence equation (4.5) gives us an infinite number of trumping relations.*

We can rewrite the second inequality of equation (4.5) as

$$\frac{\left(\frac{1}{2}\right)^p + \left(\frac{3}{2}\right)^p}{2} \leq \frac{\left(\frac{1}{3}\right)^p + \left(\frac{3}{3}\right)^p + \left(\frac{5}{3}\right)^p}{3}. \tag{4.6}$$

Now, returning to our Riemann sums results, consider approximating the integral

$$\frac{1}{2}\int_0^2 x^p \, dx$$

using midpoint Riemann sums. If we divide the interval $[0,2]$ *into two equal intervals* $[0,1] \cup [1,2]$*, the midpoints are* $1/2$ *and* $3/2$*; if we divide the interval* $[0,2]$ *into three equal intervals* $[0,2/3] \cup [2/3,4/3] \cup [4/3,2]$*, the midpoints are* $1/3, 3/3$*, and* $5/3$*. Thus, asking whether inequality (4.5) is strict amounts to asking whether estimates of an integral of the function* $f(x) = x^p$ *using midpoint Riemann sums strictly improves as* n *increases. For instance,* $M_2(x^p) \leq M_3(x^p)$ *(where* $[a,b] = [0,2]$*) gives us* $\frac{\left(\frac{1}{2}\right)^p + \left(\frac{3}{2}\right)^p}{2} < \frac{\left(\frac{1}{3}\right)^p + \left(\frac{3}{3}\right)^p + \left(\frac{5}{3}\right)^p}{3}$ *which is the second inequality in (4.5). The full system of inequalities in (4.5) is* $M_1(x^p) \leq M_2(x^p) \leq M_3(x^p) \leq \ldots$ *for* $p \in (-\infty, 0) \cup (1, \infty)$ *and with the inequalities reversed for* $p \in (0,1)$*.*

As stated earlier, we can show that these inequalities are strict by using theorem 4.4.2. We observe that $f(x) = x^p$ *is strictly convex for* $p > 1, p < 0$ *and strictly concave for* $0 < p < 1$*, yielding* $M_n(f) > M_m(f)$ *for all* $m < n$ *when* $p > 1, p < 0$ *and the inequality reverses for* $0 < p < 1$*—precisely what is needed for strict power majorization. Thus the system of inequalities (4.5) is strict, and in particular,* $x = (3,3,3,9,9,9)$ *is strictly power majorized by* $y = (2,2,6,6,10,10)$*. To prove trumping, we use proposition 4.2.7 and note that* x *is composed of odd numbers and* y *is composed*

of even numbers so it follows that $\prod_{i=1}^{d} x_i \neq \prod_{i=1}^{d} y_i$ *and* $\prod_{i=1}^{d} x_i^{x_i} \neq \prod_{i=1}^{d} y_i^{y_i}$. *The same idea works for any pair of n-tuples generated by the sequence of inequalities in (4.5) since every term in the numerator is odd and the denominators alternate between odd and even, meaning one of the tuples will consist entirely of odd numbers and the other entirely of even numbers. This gives us an infinite sequence of pairs of vectors* (x, y) *where* x *is trumped by, but not majorized by,* y.

Example 4.4.4. *Using the machinery of the previous example, is it possible that we can always create vectors* x, y *such that* x *is not majorized by* y *but* x *is trumped by* y*? (We call such a relation* non-trivial trumping*). The following example shows that this is not always the case.*

Indeed, if $p > 1$ *or* $p < 0$, *we have the system of strict inequalities [Ben05]*

$$\frac{1^p}{3^p} < \frac{1^p + 3^p}{5^p + 7^p} < \frac{1^p + 3^p + 5^p}{7^p + 9^p + 11^p} < \cdots < \frac{1^p + \cdots + (2n-1)^p}{(2n+1)^p + \cdots + (4n-1)^p} < \cdots$$

The inequality reverses for $0 < p < 1$.

For sake of example, consider the second inequality:

$$\frac{1^p + 3^p}{5^p + 7^p} < \frac{1^p + 3^p + 5^p}{7^p + 9^p + 11^p}.$$

Cross-multiplying, we obtain

$$7^p + 9^p + 11^p + 21^p + 27^p + 33^p < 5^p + 7^p + 15^p + 21^p + 25^p + 35^p$$

for $p > 1$ *or* $p < 0$, *with reverse inequality for* $0 < p < 1$. *In [Ben05], the author shows that this is an example of strict power majorization; however, one can easily*

check that $(7, 9, 11, 21, 27, 33)$ *is majorized by* $(5, 7, 15, 21, 25, 35)$*, and in that sense this is a trivial example.*

4.5 Dirichlet Polynomials, Completely Monotonone Functions, Mellin transforms, and Trumping

4.5.1 Dirichlet Polynomials

Definition 4.5.1. *A* general Dirichlet polynomial *is a polynomial of the form*

$$\sum_{n=1}^{k} a_n e^{-\lambda_n s},$$

where, herein, we take $a_n, s \in \mathbb{R}$ *(in general, they can be complex) and* $\{\lambda_n\}$ *is a strictly increasing sequence of positive numbers that tends to infinity.*

We obtain a Dirichlet polynomial

$$\sum_{n=1}^{k} \frac{a_n}{n^s},$$

via $\lambda_n = \log n$*.*

Our results make use of general Dirichlet polynomials; we do not address the case when the number of summands is infinite, which would likely introduce convergence issues in our results.

4.5.2 Completely Monotone Functions

Definition 4.5.2. *Let I be a real interval. A C^∞-function f is said to be* completely monotone *on I if $(-1)^n f^{(n)}(x) \geq 0$ for all $x \in I$ and all $n = 0, 1, 2, \ldots$.*

The following elementary observation will be useful to us later on:

Lemma 4.5.3. *Any non-zero entire function that is completely monotone on $(0, \infty)$ must be strictly positive on $\mathbb{R}$.*

Proof. Let f be a non-zero entire function that is completely monotone on $(0, \infty)$. Then f is non-negative and non-increasing on $[0, \infty)$. It follows that if there exists $c \in [0, \infty)$ such that $f(c) = 0$, then f must be zero on $[c, \infty)$. However, the zero set of a non-zero entire function cannot have a limit point. Indeed, let z_0 be a limit point of the zero set of f. Then by continuity z_0 is a zero of f. Let m be the multiplicity of that zero. Hence $g(z) \equiv \frac{f(z)}{(z-z_0)^m}$ is entire and non-zero at z_0. By continuity of g, g is non-zero in a neighbourhood of z_0, hence $f(z) \equiv g(z)(z - z_0)^m$ is also non-zero in a neighbourhood of z_0, and so, again by continuity, cannot be zero at z_0, a contradiction.

Since the derivatives of f are continuous of all orders, $(-1)^n f^{(n)}(0) \geq 0$ and, using the MacLaurin series $f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = \sum_{n=0}^{\infty} \frac{(-1)^n f^{(n)}(0)}{n!} (|x|)^n \geq 0$ for $x < 0$, we can argue as above that $f(x)$ is strictly greater than zero on $(-\infty, 0)$ as well.

■

The example $f(x) = \frac{1}{1+x}$, which is negative on $(-\infty, -1)$, shows us that the

requirement that f be entire in the above lemma cannot be removed. Completely monotone functions are necessarily positive, decreasing, and convex. Bernstein's theorem [Ber29] on monotone functions characterizes completely monotone functions f on $(0,\infty)$ via

$$f(s) = \int_0^\infty e^{-st}\, d\mu(t),$$

the Laplace transform of a positive measure. In this way the completely monotone functions on $(0,\infty)$ are seen to be the cone generated by e^{-st}.

We recall that the Mellin transform of a function f on $(0,\infty)$ is the function $\phi(s) = \int_0^\infty f(t)t^{s-1}dt$. The Mellin and Laplace transforms are closely related. One such connection is the following observation:

Proposition 4.5.4. *Let $f \in L^1((0,\infty))$ be zero outside of $[0,1]$. Then the Mellin transform of $f(x)$ is the Laplace transform of $f(e^{-x})$.*

Indeed, the Mellin transform of $f(u)$ is

$$\int_0^\infty f(u)u^{s-1}\, dx.$$

Using the substitution $u = e^{-x} \Rightarrow du = -e^{-x}\, dx$. The upper and lower limits in terms of x change to 0 and ∞ respectively, and we obtain

$$\begin{aligned}\int_0^\infty f(u)u^{s-1}\, dx &= \int_\infty^0 f(e^{-x})(e^{-x})^{s-1}(-e^{-x})\, dx \\ &= \int_0^\infty f(e^{-x})e^{-xs}\, dx,\end{aligned}$$

which is exactly the Laplace transform of the function $f(e^{-x})$.

This observation, together with Bernstein's characterization of completely monotone functions on $(0,\infty)$, leads to the following useful characterization of complete positivity of Mellin transforms of functions supported on $[0,1]$.

Corollary 4.5.5. *Let $f \in L^1((0,\infty))$ be zero outside of $[0,1]$. Then the Mellin transform of f is completely monotone on $(0,\infty)$ iff f is non-negative almost everywhere.*

Finally, we will require the fact that the product of two completely monotone functions on I is a completely monotone function on I. For the convenience of the reader, we prove this well-known fact below.

Lemma 4.5.6. *Let I be a real interval. Then the product of completely monotone functions on I is itself completely monotone on I.*

Proof. We only need to show that if f and g be completely monotone functions on I then the product fg is completely monotone on I. Indeed, using the product rule for higher order derivatives, we find

$$\begin{aligned}(-1)^n(fg)^{(n)}(x) &= (-1)^n \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)} g^{(k)}(x) \\ &= \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} f^{(n-k)}(x) (-1)^k g^{(k)}(x).\end{aligned}$$

Both $(-1)^{n-k} f^{(n-k)}(x)$ and $(-1)^k g^{(k)}(x)$ are non-negative for all $x \in I$ and all $n-k$ (k, respectively) $= 0, 1, 2, \ldots$. It follows that fg is completely monotone on I by definition.

■

4.5.3 Connection to Majorization & Trumping

Consider the following Dirichlet polynomial: $\zeta(s) = \sum_i \frac{1}{y_i^s} - \sum_i \frac{1}{x_i^s}$, which we can write as $\sum_n \frac{a_n}{n^s}$ with a_n as in equation (4.3). We will relax our assumption that the x_i, y_i need to be integers. For non-integer values we can simply use a general Dirichlet polynomial $\sum_{n=1}^{k} a_n e^{-\lambda_n s}$. For ease and consistency of notation we will still use $\sum_n \frac{a_n}{n^s}$ where n is summed over all real numbers in $(1, \infty)$ rather than the natural numbers. Note that is still a finite sum since $a_n = \#\{i : y_i = n\} - \#\{i : x_i = n\}$.

By considering this generalized Dirichlet polynomial, lemma 4.1.8 can be re-written as

Theorem 4.5.7. *Let x and y be vectors of the same length all of whose entries lie in $(1, \infty)$ and let $\zeta(s) = \sum_i \frac{1}{y_i^s} - \sum_i \frac{1}{x_i^s}$. Then $x \prec y$ if and only if*

(i) $\zeta(0) = 0$;

(ii) $\zeta(-1) = 0$;

(iii) $\frac{\zeta(s)}{s(s+1)}$ *is completely monotone on* $(0, \infty)$.

Proof. Items (i) and (ii) can easily be seen to correspond to items (i) and (ii) of lemma 4.1.8. We show the correspondence of the respective items (iii).

Using a well-known result by Abel in Analytic number theory, we have

$$\zeta(s) = \sum_{n=0}^{x} \frac{a_n}{n^s} = \frac{\mu_1(t)}{x^s} + s \int_1^x \mu_1(t) t^{-s-1}\, dt,$$

with x large enough to capture all the terms and $\mu_1(t) = \sum_{0 \le n \le t} a_n$.

Since $\mu_1(t) = 0$ for t sufficiently large, the term $\frac{\mu_1}{x^s} \to 0$ as $x \to \infty$. We then have

$$\begin{aligned}\zeta(s) &= s\int_1^x \mu_1(t)t^{-s-1}\,dt \\ &= t^{-(s+1)}\mu_2(t)\,|_1^\infty + s(s+1)\int_1^\infty \mu_2(t)t^{-(s+2)}\,dt,\end{aligned}$$

where we have used integration by parts. Since $\mu_2(t) = 0$ for t sufficiently large, the first term on the RHS tends to 0 as $t \to \infty$ (it is also equal to zero at $t = 1$). We therefore obtain

$$\zeta(s) = s(s+1)\int_1^\infty \mu_2(t)t^{-(s+2)}\,dt \quad \text{for } s > 0.$$

Dividing by $s(s+1)$, we obtain

$$\frac{\zeta(s)}{s(s+1)} = \int_1^\infty \mu_2(t)t^{-(s+2)}\,dt.$$

By making the change of variables $t = e^x$, we obtain

$$\frac{\zeta(s)}{s(s+1)} = \int_0^\infty \mu_2(e^x)e^{-x}e^{-sx}\,dx. \tag{4.7}$$

This is precisely the Laplace transform of the function $\mu_2(e^x)e^{-x}$.

Suppose $\mu_2(\cdot) \geq 0$, so that $\mu_2(e^x)e^{-x} \geq 0$. By Bernstein's theorem, the Laplace transform of a positive measure is completely monotone. A Dirichlet polynomial is always the Laplace transform of a function (not simply a general measure); moreover, Laplace transforms are unique up to differences on sets of measure 0. Therefore, provided $\mu_2(\cdot) \geq 0$, we have that $\frac{\zeta(s)}{s(s+1)}$ is completely monotone on $(0, \infty)$. Furthermore, if $\frac{\zeta(s)}{s(s+1)}$ is completely monotone on $(0, \infty)$ and if equation (4.7) holds, again using the

uniqueness of the Laplace transform as well as Bernstein's theorem, it follows that $\mu_2(\cdot) \geq 0$.

Note that $\mu_2(\cdot) \geq 0$ is precisely condition (iii) of lemma 4.1.7. ■

If there exists a catalyst c such that $x \otimes c \prec y \otimes c$, then condition (iii) of theorem 4.5.7 says that if $\tilde{\zeta}$ is the corresponding Dirichlet polynomial then $\frac{\tilde{\zeta}(s)}{s(s+1)}$ is completely monotone on $(0, \infty)$. It is easy to see that $\tilde{\zeta}(s)$ is given by

$$\begin{aligned} \tilde{\zeta}(s) &= \sum_{i,j} \frac{1}{(y_i c_j)^s} - \frac{1}{(x_i c_j)^s} \\ &= \left(\sum_i \frac{1}{y_i^s} - \frac{1}{x_i^s} \right) \left(\sum_j \frac{1}{c_j^s} \right). \end{aligned}$$

Thus we can write $\tilde{\zeta}(s) = \zeta(s)\zeta_2(s)$ where $\zeta_2(s) = \sum_j \frac{1}{c_j^s}$ is the Dirichlet polynomial corresponding to the non-zero catalyst vector c.

Note that, by virtue of its coefficients being non-negative, ζ_2 is completely monotone on $(0, \infty)$. This leads us to the following observation, which follows from theorem 4.5.7 and the definition of trumping.

Proposition 4.5.8. *Let x and y be vectors of the same length all of whose entries lie in $(1, \infty)$. Let $\zeta(s) = \sum_i \frac{1}{y_i^s} - \sum_i \frac{1}{x_i^s}$. The following statements are equivalent:*

1. *$x \prec_T y$;*

2. *There exists a Dirichlet polynomial $\zeta_2 \not\equiv 0$ with non-negative coefficients such that*

$$\frac{\zeta(s)\zeta_2(s)}{s(s+1)}$$

is completely monotone on $(0,\infty)$.

We can reformulate the conditions (T1)-(T3) of Turgut's theorem (theorem 4.2.3) by using the general Dirichlet polynomial $\zeta(s) = \sum_n a_n e^{-\lambda_n s}$. We note that, in terms of the functionals f_r from theorem 4.2.4, we have $\zeta(s) = \exp\{f_{-s}(y) - f_{-s}(x)\}$ when $s \notin (-1,0)$; $\zeta(s) = \exp\{f_{-s}(x) - f_{-s}(y)\}$ when $s \in (-1,0)$; $\zeta'(0) = f_0(y) - f_0(x)$ and $\zeta'(-1) = f_1(x) - f_1(y)$. We can use these observations to rewrite Turgut's theorem in terms of Dirichlet polynomials.

The following proposition can be taken as a succinct restatement of Turgut's theorem (theorem 4.2.3).

Proposition 4.5.9. *Let* x *and* y *be vectors of the same length all of whose entries lie in* $(1,\infty)$ *and let* $\zeta(s) = \sum_i \frac{1}{y_i^s} - \sum_i \frac{1}{x_i^s}$. *Then the following statements are equivalent:*

1. $x \prec_T y$;

2. ζ *is a general Dirichlet polynomial with simple zeros at* -1 *and* 0 *such that*

$$\frac{\zeta(s)}{s(s+1)}$$

is positive for all $s \in \mathbb{R}$;

The proof follows from theorem 4.5.19.

If we do not assume $\frac{\zeta(s)}{s(s+1)}$ is completely monotone on $(0,\infty)$, then the integral $\int_0^\infty \mu_2(e^x)e^{-x}e^{-sx}\,dx$ could be positive *without* μ_2 being positive. For this case, it is possible that $\frac{\zeta(s)}{s(s+1)} > 0$ without $\frac{\zeta(s)}{s(s+1)}$ being completely monotone on $(0,\infty)$. This would be precisely the case of non-trivial trumping (that is, trumping that is not majorization).

4.5.4 Higher Order Convexity

Consider a sequence $\chi = (\chi_n)_n$. For any natural number r, we define the *r-th difference operator* [PS58] by $\Delta^r\chi_n = \sum_{j=0}^{r}(-1)^j\binom{r}{j}\chi_{n-j}$, with $\chi_k \equiv 0$ for negative k. Thus, for example, $\Delta\chi_n = \chi_n - \chi_{n-1}$ and $\Delta^2\chi_n = \chi_n - 2\chi_{n-1} + \chi_{n-2}$. We say that a sequence $\chi = (\chi_n)_n$ is *r-convex* if $\Delta^r\chi_n \geq 0$ for all $n = 1, 2, \ldots$.

We can also define higher order continuous functions using divided differences.

Definition 4.5.10. *Let I be an interval and $f : I \to \mathbb{R}$ and let $\{x_j\}$ be distinct elements of I. Then we define the first order divided difference as $f[x_0, x_1] = \frac{f(x_1)-f(x_0)}{x_1-x_0}$. The second order divided difference is $f[x_0, x_1, x_2] = \frac{f[x_1,x_2]-f[x_1,x_0]}{x_2-x_0}$. Higher order divided differences are defined in a similar inductive manner:*

$$f[x_0, ..., x_n] = \frac{f[x_1, ..., x_n] - f[x_{n-1}, ..., x_0]}{x_n - x_0}.$$

An equivalent characterization is that $f[x_0, ..., x_n]$ is the coefficient of x^n in the Lagrange interpolating polynomial to f at the nodes $x_0, x_1, ..., x_n$.

Definition 4.5.11. *[Bul71] Let I be an interval and $f : I \to \mathbb{R}$, f is said to be a*

convex function *of order r (or r-convex function) on I if $f[x_0, ..., x_r] \geq 0$ whenever $\{x_j\}_{j=0}^r$ is a $(r+1)$-tuple of distinct numbers in I.*

We note that the term "1-convex" is equivalent to "increasing", and "2-convex" is equivalent to "convex", for both higher order convex sequences and higher order convex functions.

Observation 4.5.12. *There is a close relationship between r-convex functions and r-convex sequences. If f is an r-convex function on $(0, \infty)$, then $\{f(n)\}_{n\in\mathbb{N}}$ is an r-convex sequence.*

We define $\mu_k(x) = \int_a^x \mu_{k-1}(t)\, dt$ for $k = 1, 2, 3, \ldots$, where $\mu_0 := \mu$. The r-convex analogues to lemmas 4.1.8 and 4.1.7 are as follows.

Lemma 4.5.13. *The inequality*

$$\int_a^b \phi(t)\, d\mu \geq 0 \quad \text{for all } r\text{-convex functions } \phi \tag{4.8}$$

is equivalent to the following two conditions.

(i) $\int_a^b x^k\, d\mu = 0$ *for* $k = 0, 1, ..., r-1$;

(ii) $(-1)^r \mu_r(x) \geq 0$ *for all* $x \in [a, b]$.

Proof. The proof follows the original proof from [KN63]. If inequality (4.8) holds, then since $\psi_\pm := \pm x^k$ are r-convex for $k = 0, 1, ..., r-1$, item (i) follows immediately. The $k = 0$ case of item (i) implies $\mu_1(b) = 0$, and the $k = 0$ and $k = 1$ cases of item

(i) together imply $\mu_2(b) = 0$, these cases together with the $k = 2$ case in turn implies $\mu_3(b) = 0$. We can repeat this to obtain $\mu_k(b) = 0$, for all $1 \leq k \leq r$. Applying integration by parts r times to $\int_a^b \phi(t)\, d\mu$ yields

$$\int_a^b \phi(t)\, d\mu = (-1)^r \int_a^b \phi^{(r)} \mu_r(x)\, dx. \tag{4.9}$$

The assumption that equation (4.9) is non-negative, together with the assumption that ϕ is r-convex, gives item (ii).

These two items are also sufficient: For an r-th differentiable function ϕ satisfying items (i) and (ii), equation (4.9) holds, which, assuming ϕ is r-convex, implies $\int_a^b \phi(t)\, d\mu \geq 0$. ■

The discrete version of the above lemma is stated below for completeness.

Lemma 4.5.14. *The inequality*

$$\sum_{0 \leq i \leq m} a_i \chi_n \geq 0 \quad \text{for all } r\text{-convex sequences } (\chi_n)_n \tag{4.10}$$

is equivalent to the following two conditions.

(i) $\sum_{0 \leq n \leq m} n^k a_n = 0$ *for* $k = 0, 1, ..., r-1$*;*

(ii) If r *is even,* $\sum_{i_r=0}^{i_r+1} \sum_{i_{r-1}=0}^{i_r} \cdots \sum_{i_2=0}^{i_3} \sum_{i_1=0}^{i_2} a_{i_1} \geq 0$ *where* $0 \leq i_j \leq m$ *for all* $j \leq r+1$*, with the reverse inequality for odd* r*.*

In the discrete setting, it is notationally convenient to denote $\displaystyle\sum_{i_r=0}^{i_r+1} \sum_{i_{r-1}=0}^{i_r} \cdots \sum_{i_2=0}^{i_3} \sum_{i_1=0}^{i_2} a_{i_1}$ by μ_r, as in the continuous case.

Using Descartes' rule of signs, we can restate corollary 4.1 of [Nie05] ever so slightly, as follows.

Theorem 4.5.15. *Let $a = (a_0, a_1, \ldots, a_\ell)$ be a real vector and let $0 \leq r \leq \ell$. The inequality $\sum_{n=0}^{\ell} a_n \chi_n \geq 0$ holds for all r-convex sequences $\chi = (\chi_n)_n$ if an only if the polynomial $p(z) = \sum_{n=0}^{\ell} a_n z^n$ has a root of multiplicity r at $z = 1$ and all the coefficients of the polynomial $\frac{p(z)}{(z-1)^r}$ are non-negative.*

For our setting, we wish to prove a generalized version of this theorem for Dirichlet polynomials.

Theorem 4.5.16. *Let $a = (a_0, a_1, \ldots, a_\ell)$ be a real vector and let $0 \leq r \leq \ell$. The inequality $\sum_{n=0}^{\ell} a_n \chi_n \geq 0$ holds for all r-convex sequences $\chi = (\chi_n)$ if an only if the Dirichlet polynomials $\zeta(s) = \sum_n \frac{a_n}{n^s}$ has zeros at $s = 0, -1, -2, \ldots, -r+1$ and $\frac{(-1)^r \zeta(s)}{\Pi_{k=0}^{r-1}(s+k)}$ is completely monotone on $(0, \infty)$.*

Proof. The bulk of the proof follows the calculations preceeding theorem 4.5.7, which use a well-known result due to Abel, which we restate here for convenience.

$$\zeta(s) = \sum_{n=0}^{x} \frac{a_n}{n^s} = \frac{\mu_1(t)}{x^s} + s \int_1^x \mu_1(t) t^{-s-1}\, dt,$$

with x large enough to capture all the terms.

If $s > 0$, we note that the first term on the RHS vanishes. We can integrate the second term on the RHS by parts, and the first term of the result vanishes as well.

Continuing in this manner, integrating by parts $r-1$ times (where $r \geq 2 \in \mathbb{Z}$) yields

$$\zeta(s) = \Pi_{k=0}^{r-1}(s+k) \int_1^\infty \mu_r(t) t^{-(s+r)}\, dt.$$

Dividing by $\Pi_{k=0}^{r-1}(s+k)$, multiplying both sides of the equation by $(-1)^r$, and making the change of variables $t = e^x$, we obtain

$$\frac{(-1)^r \zeta(s)}{\Pi_{k=0}^{r-1}(s+k)} = \int_0^\infty (-1)^r \mu_r(e^x) e^{-rx} e^{-sx}\, dx. \tag{4.11}$$

It follows from lemma 4.5.14 that $(-1)^r \mu_r(e^x) e^{-rx} \geq 0$ for any r. Similar to before, we note that equation (4.11) is precisely the Laplace transform of the function $(-1)^r \mu_r(e^x) e^{-rx}$, and we conclude $\frac{(-1)^r \zeta(s)}{\Pi_{k=0}^{r-1}(s+k)}$ is a completely monotone function on $(0, \infty)$.

For the reverse direction, looking at equation (4.11) and noting Bernstein's theorem, the result follows from lemma 4.5.14. ■

Note that theorem 4.5.16 is a direct generalization of theorem 4.5.7 with the same setup as before, namely $a_n = \#\{i : y_i = n\} - \#\{i : x_i = n\}$ for all $n \in \mathbb{N}$ (and using a general Dirichlet polynomial in the case of non-integer values). The inequality $\sum_n a_n \chi_n \geq 0$ holding for all r-convex sequences (rather than for all convex sequences, as is the case for $x \prec y$) is a more generalized partial order on real vectors.

We note that we can write $\frac{\zeta(s)}{\Pi_{k=0}^{r-1}(s+k)}$ as a Mellin transform. We remind the reader of the truncated power notation used in the theory of splines $(1-nx)_+^{r-1} = (\max(1-nx, 0))^{r-1}$. We are now ready to state our result:

Proposition 4.5.17. *Let* $\zeta(s) := \sum_{n=1}^{\infty} \frac{a_n}{n^s}$. *Then* $\frac{\zeta(s)}{\prod_{k=0}^{r-1}(s+k)}$ *is the Mellin transform of* $\frac{1}{(r-1)!}\sum_{n=1}^{\infty} a_n(1-nx)_+^{r-1}$.

Proof. The proof is a straightforward calculation using integration by parts.

To simplify the manipulations of expressions, let $g_{q,r}$ be the Mellin transform of $x^q(1-nx)_+^{r-1}$. Note that the expression $\left((1-nx)_+\right)^0$ is 1 for $x \in [0, 1/n]$ and 0 for $x > 1/n$. This yields

$$\begin{aligned} g_{q,0} &= \int_0^{\infty} x^q \chi_{[0,1/n]}(x) x^{s-1}\, dx \\ &= \int_0^{1/n} x^{q+s-1}\, dx \\ &= \frac{n^{-(q+s)}}{q+s}. \end{aligned}$$

For $r-1 > 0$, we use integration by parts with $u = x^q(1-nx)^{r-1}$ and $dv = x^{s-1}dx$ so that

$$\begin{aligned} g_{q,r} &= \int_0^{\infty} x^q (1-nx)_+^{r-1} x^{s-1}\, dx \\ &= (1-nx)^{r-1}\frac{x^{q+s}}{q+s}\Big|_0^{1/n} + \frac{n(r-1)}{q+s}\int_0^{1/n}(1-nx)^{r-2}x^{q+s}\, dx \\ &= \frac{n(r-1)}{q+s}\int_0^{1/n}(1-nx)^{r-2}x^{q+s}\, dx \\ &= \frac{n(r-1)}{q+s} g_{q+1,r-1}. \end{aligned}$$

In general, we obtain

$$g_{q,r} = \frac{n^{-(q+s)}(r-1)!}{(q+s)(q+s+1)\cdots(q+s+r-1)}.$$

Letting $q = 0$, we have

$$g_{0,r} = \frac{n^{-s}}{\prod_{k=0}^{r-1}(s+k)}(r-1)!$$

Since the Mellin transform of $(1-nx)_+^{r-1}$ is $g_{0,r}$, we see that the Mellin transform of $\frac{1}{(r-1)!}\sum_{n=1}^{\infty} a_n(1-nx)_+^{r-1}$ is

$$\begin{aligned}\frac{1}{(r-1)!}\sum_{n=1}^{\infty} a_n g_{0,r} &= \frac{\sum_{n=1}^{\infty} a_n n^{-s}}{\prod_{k=0}^{r-1}(s+k)} \\ &= \frac{\zeta(s)}{\prod_{k=0}^{r-1}(s+k)},\end{aligned}$$

as desired. ■

We can use this result to generalize Turgut's theorem. We first state a lemma from [Tur07]. This lemma is a special case of a result of Pólya [Pól28]. This result was rediscovered in [Tur07].

Lemma 4.5.18. *Let $f(x)$ be a real polynomial which is positive on $(0,\infty)$. Then there exists two real polynomials $g(x)$ and $h(x)$ with all coefficient nonnegative such that $f(x)g(x) = h(x)$.*

Theorem 4.5.19. *Let $\zeta(s)$ be a generalized Dirichlet polynomial with simple zeros at $s = 0, -1, ..., -(r-2), -(r-1)$. Then $\frac{\zeta(s)}{\prod_{k=0}^{r-1}(s+k)}$ is positive on the real line if and only if there exists a generalized Dirichlet polynomial $\zeta_2(s) \not\equiv 0$ with non-negative coefficients such that $\frac{\zeta(s)\zeta_2(s)}{\prod_{k=0}^{r-1}(s+k)}$ is completely monotone on $(0,\infty)$.*

Proof. Suppose $g(s) = \frac{\zeta(s)\zeta_2(s)}{\prod_{k=0}^{r-1}(s+k)}$ is completely monotone on $(0,\infty)$. By lemma 4.5.3, $g(s)$ must be positive on the real line. Since $\zeta_2(s)$ is strictly positive on the real line, $\frac{g(s)}{\zeta_2(s)} = \frac{\zeta(s)\zeta_2(s)}{\prod_{k=0}^{r-1}(s+k)}$ must be strictly positive on the real line.

The proof of other direction is modelled after the proof of the main result in [Tur07]. As in [Tur07], we begin with the special case that $\zeta(s) = \sum_{n=0}^{k} a_n e^{-\lambda_n s}$

where $\lambda_n = n\alpha$ where α is some positive number. Let $p(x) = \sum_{n=0}^{k} a_n x^n$, then $\zeta(s) = p(e^{-\alpha s})$. If the only real zeros of $\zeta(s)$ are simple zeros at $-(r-1), -(r-2), ..., -1, 0$ then the only positive real zeros of $p(x)$ are simple zeros at $1, e^{\alpha}, e^{2\alpha}, ..., e^{(r-1)\alpha}$. Hence $p(x) = f(x)\prod_{k=0}^{r-1}(e^{k\alpha} - x)$ where $f(x)$ is a real polynomial which is positive on $(0, \infty)$. Hence, by lemma 4.5.18, there exists two real polynomials $g(x)$ and $h(x)$ with all coefficients nonnegative such that $f(x)g(x) = h(x)$. Let $\zeta_2(s) = g(e^{-\alpha s})$, then $\zeta_2(s)$ is a generalized Dirichlet polynomial with nonnegative coefficients. Note that

$$\frac{\zeta(s)\zeta_2(s)}{\prod_{k=0}^{r-1}(s+k)} = f(e^{-\alpha s})g(e^{-\alpha s})\frac{\prod_{k=0}^{r-1}(e^{\alpha k} - e^{-\alpha s})}{\prod_{k=0}^{r-1}(s+k)} = h(e^{-\alpha s})\prod_{k=0}^{r-1}\frac{e^{\alpha k} - e^{-\alpha s}}{s+k}.$$

Since $h(e^{-\alpha s})$ is a generalized Dirichlet polynomial with nonnegative coefficients it is completely monotone on $(0, \infty)$.

The functions $e^{\alpha k} - e^{-\alpha s}$ and $\frac{1}{s+k}$ are completely monotone on $(0, \infty)$ for all $k \in \mathbb{N} \cup \{0\}$. Therefore the product $\frac{\zeta(s)\zeta_2(s)}{\prod_{k=0}^{r-1}(s+k)}$ is completely monotone on $(0, \infty)$ by lemma 4.5.6. The general case of this result now follows from an approximation argument similar to that of [Tur07]. ■

Turgut's theorem (theorem 4.2.3) can easily be seen as a corollary: when $r = 2$, the statement of our theorem is exactly proposition 4.5.8:

Corollary 4.5.20. *Let $\zeta(s)$ be a generalized Dirichlet polynomial with simple zeros at $s = 0, -1$. Then $\frac{\zeta(s)}{s(s+1)}$ is positive on the real line if and only if there exists a generalized Dirichlet polynomial $\zeta_2(s) \not\equiv 0$ with non-negative coefficients such that $\frac{\zeta(s)\zeta_2(s)}{s(s+1)}$ is completely monotone on $(0, \infty)$.*

Bibliography

[All88] G. D. Allen. Power majorization and majorization of sequences. *Result. Math.*, 14:211–222, 1988. 119

[AMTdW00] A. Ambainis, M. Mosca, A. Tapp, and R. de Wolf. Private quantum channels. *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 547–553, 2000. 25, 27, 51

[AN08] G. Aubrun and I. Nechita. Catalytic majorization and ℓ_p norms. *Comm. Math. Phys.*, 278, 2008. 114, 116

[Arn87] B. C. Arnold. *Majorization and the Lorenz Order: A Brief Introduction. Lecture Notes in Statistics*, volume 43. Springer-Verlag, Berlin, 1987. 108

[BBPS96] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev. A*, 53(4):2046–2052, 1996. 22

[Ben86] G. Bennett. Power majorization versus majorization. *Anal. Math.*, 12:283–286, 1986. 119

[Ben05] G. Bennett. An odd inequality. *Amer. Math. Monthly*, 112:Problem 11139, 2005. 129

[Ben10] G. Bennett. p-free ℓ^p inequalities. *Amer. Math. Monthly*, 117:334–351, 2010. 125, 126, 127

[Ben11] G. Bennett. A problem of Ghorbani. *Anal. Math.*, 37:239–244, 2011. 119

[Ber29] S. N. Bernstein. Sur les fonctions absolument monotones. *Acta Math.*, 52(1):1–66, 1929. 132

[Bha07] R. Bhatia. *Positive Definite Matrices*. Princeton University Press, Princeton, New Jersey, 2007. 13

[BHS05] S. Bartlett, P. Hayden, and R. Spekkens. Random subspaces for encryption based on private shared cartesian frame. *Phys. Rev. A*, 72:052329, 2005. 40, 41, 51

[BJ00] G. Bennett and G. Jameson. Monotonic averages of convex functions. *J. Math. Anal. Appl.*, 252:410–430, 2000. 125, 126

[BKD+05] F. Buscemi, M. Keyl, G. M. D'Ariano, P. Perinotti, and R. F. Werner. Clean positive operator valued measures. *J. Math. Phys.*, 46(8):082109, 17, 2005. 4, 6, 72, 73, 95, 101, 102

[BR03] P. O. Boykin and V. Roychowdhury. Optimal encryption of quantum bits. *Phys. Rev. A*, 67:042317, 2003. 25, 51

[BRS04] S. Bartlett, T. Rudolph, and R. Spekkens. Optimal measurements for relative quantum information. *Phys. Rev. A*, 70:032307, 2004. 51

[Bul71] P. S. Bullen. A criterion for n-convexity. *Pacific J. Math.*, 36(1):81–98, 1971. 138

[BŻ06] I. Bengtsson and K. Życzkowski. *Geometry of quantum states*. Cambridge University Press, Cambridge, 2006. 78

[BZ07] J. Bouda and M. Ziman. Optimality of private quantum channels. *J. Phys. A: Math. Theor.*, 40:5415–5426, 2007. 32

[CE77] M.-D. Choi and E. G. Effros. Injectivity and operator spaces. *J. Funct. Anal.*, 24:156–209, 1977. 23

[Cho75] M.-D. Choi. Completely positive linear maps on complex matrices. *Linear Algebra Appl.*, 10:285–290, 1975. 12

[CKPP11] A. Church, D. W. Kribs, R. Pereira, and S. Plosker. Private quantum channels, conditional expectations, and trace vectors. *Quantum Information & Computation (QIC)*, 11(9 & 10):774–483, 2011. 6

[Cla84] A. Clausing. A problem concerning majorization. *General Inequalities*, 4 (W. Walter, Ed.), Birkhäuser, Basel:405, 1984. 119

[Dah10] G. Dahl. Majorization permutahedra and $(0, 1)$-matrices. *Linear Algebra and its Applications*, 432:3265–3271, 2010. 121

[Dav96] K. Davidson. *C*-algebras by example*, volume 6. Fields Institute Monographs, American Math. Society, Providence, RI, 1996. 28

[DK01] S. K. Daftuar and M. Klimesh. The trumping relation and the structure of the bipartite entangled states. E-print: arXiv:quant-ph/0104058, 2001. 114, 121, 122, 123

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, 1935. 10

[FFP13] D. Farenick, R. Floricel, and S. Plosker. Approximately clean quantum probability measures. *J. Math. Phys.*, 54:052201, 2013. 7, 73, 75, 76, 78, 80, 91, 93, 95

[FH77] P. Fischer and J. A. R. Holbrook. Matrices doubly stochastic by blocks. *Can. J. Math.*, 29(3):559–577, 1977. 119

[FPS11] D. Farenick, S. Plosker, and J. Smith. Classical and nonclassical randomness in quantum measurements. *J. Math. Phys.*, 52:122204, 2011. 6, 82, 85, 87, 88

[FZ07] D. R. Farenick and F. Zhou. Jensen's inequality relative to matrix-valued measures. *J. Math. Anal. Appl.*, 327(2):919–929, 2007. 82

[Gav05] I. Gavrea. Some remarks on a paper by A. McD. Mercer. *J. Inequal. Pure Appl. Math*, 6(1, Art. 26), 2005. 106, 111

[Ge] L. Ge. On "Problems on von Neumann algebras by R. Kadison, 1967". *Acta Math. Sinica, English Series*, 19(3). 30

[GG77] P. Gaiha and S. K. Gupta. Adjacent vertices on a permutohedron. *SIAM Journal on Applied Mathematics*, 32(2):323–327, 1977. 121

[Hei05] T. Heinonen. Optimal measurements in quantum mechanics. *Phys. Letters A*, 346:77–86, 2005. 4, 72, 101

[Hig97] N. J. Higham. Stable iterations for the matrix square root. *Numer. Algorithms*, 15(2):227–242, 1997. 82

[HLP52] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge University Press, New York, NY, 2nd edition, 1952. 108

[Hol07] A. S. Holevo. Complementary channels and the additivity problem. *Theory Probab. Appl.*, 51:92–100, 2007. 19, 20

[JOKLP13] T. Jochym-O'Connor, D. W. Kribs, R. Laflamme, and S. Plosker. Private quantum subsystems. *Phys. Rev. Lett.*, 2013. 6

[JP99] D. Jonathan and M. B. Plenio. Minimal conditions for local pure-state entanglement manipulation. *Phys. Rev. Lett.*, 83:3566, 1999. 105, 113, 116, 121

[Kah07] J. Kahn. Clean positive operator-valued measures for qubits and similar cases. *J. Phys. A*, 40(18):4817–4832, 2007. 72, 73, 101

[KKS08] D. Kretschmann, D. W. Kribs, and R. Spekkens. Complementarity of private and correctable subsystems in quantum cryptography and error correction. *Phys. Rev. A.*, 78:032330, 2008. 2, 42, 50

[KL97] E. Knill and R. Laflamme. Theory of quantum error-correcting codes. *Phys. Rev. A*, 55:900–911, 1997. 41, 42

[Kli04] M. Klimesh. Entropy measures and catalysis of bipartite quantum state transformations. *Proceedings 2004 IEEE International Symposium on Information Theory*, page 357, 2004. 114, 116, 117

[Kli07] M. Klimesh. Inequalities that collectively completely characterize the catalytic majorization relation. E-print: arXiv:0709.3680v1, 2007. 114, 116, 117

[KLP05] D. W. Kribs, R. Laflamme, and D. Poulin. A unified and generalized approach to quantum error correction. *Phys. Rev. Lett.*, 94:180501, 2005. 42

[KLPL05] D. W. Kribs, R. Laflamme, D. Poulin, and M. Lesosky. Operator quantum error correction. E-print: http://arXiv:quant-ph/0504189v1, 2005. 69

[KLPL06] D. W. Kribs, R. Laflamme, D. Poulin, and M. Lesosky. Operator quantum error correction. *Quantum Inf. Comput.*, 6:383–399, 2006. 42

[KMNR07] C. King, K. Matsumoto, M. Nathanson, and M. B. Ruskai. Properties of conjugate channels with applications to additivity and multiplicativity. *Markov Process. Related Fields*, 13:391–423, 2007. 19, 20

[KN63] S. Karlin and A. Novikoff. Generalized convex inequalities. *Pacific J. Math.*, 13:1251–1279, 1963. 112, 139

[KP12] D. W. Kribs and S. Plosker. Private quantum codes: Introduction and connection with higher rank numerical ranges. *Linear and Multilinear Algebra; Proceedings of WONRA Special Edition*, 2012. 6, 44

[KPP12] D. W. Kribs, R. Pereira, and S. Plosker. Trumping and power majorization. *Linear and Multilinear Algebra*, 2012. 7, 106, 121, 123

[KR01] C. King and M. B. Ruskai. Minimal entropy of states emerging from noisy quantum channels. *IEEE Trans. Inform. Theory*, 47:192–209, 2001. 33

[Kra71] K. Kraus. General state changes in quantum theory. *Ann. Physics*, 64:311–335, 1971. 12

[LP01] H.-K. Lo and S. Propescu. Concentrating entanglement by local actions: Beyond mean values. *Phys. Rev. A*, 63:022301, 2001. 22, 105, 109

[LS93] L. J. Landau and R. F. Streater. On Birkhoff's theorem for doubly stochastic completely positive maps of matrix algebras. *Linear Algebra Appl.*, 193:107–127, 1993. 32

[Mer05] A. McD. Mercer. Polynomials and convex sequence inequalities. *J. Inneq. Pure Appl. Math.*, 6(1, Art. 8), 2005. 106, 111

[MOA] A. W. Marshall, I. Olkin, and B. C. Arnold. *Inequalities: Theory of Majorization and its Applications*. Springer: New York, 2nd edition. 105, 123

[Mui03] R. F. Muirhead. Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters. *Proc. Edinburgh Math. Soc.*, 21:144–157, 1903. 108

[MvN37] F. J. Murray and J. von Neumann. On rings of operators II. *Trans. Amer. Math. Soc.*, 41:208–248, 1937. 29

[NC00] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2000. 9, 11, 12, 13, 32, 34

[Nie99] M. Nielsen. Conditions for a class of entanglement transformations. *Phys. Rev. Lett.*, 83(2):436–439, 1999. 5, 105, 108, 109

[Nie05] M. Niezgoda. An extension of results of A. McD. Mercer and I. Gavrea. *J. Inequal. Pure Appl. Math.*, 6(4, Art. 107), 2005. 106, 141

[NP07] M. A. Nielsen and D. Poulin. Algebraic and information-theoretic conditions for operator quantum error-correction. *Phys. Rev. A*, 75, 2007. 42

[Pól28] G. Pólya. Über positive Darstellung von Polynomen. *Vierteljschr. Naturforsch. Ges. Zürich*, 73:141–145, 1928. 144

[Pau02] V. Paulsen. *Completely bounded maps and operator algebras*, volume 78 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. 23, 24, 91, 92, 97

[Pel11] J.-P. Pellonpää. Complete characterization of extreme quantum observables in infinite dimensions. *J. Phys. A*, 44(8):085304, 12, 2011. 72, 90

[Per03] R. Pereira. Trace vectors in matrix analysis. PhD dissertation, 2003. 30

[PP13] R. Pereira and S. Plosker. Dirichlet polynomials, majorization, and trumping. submitted to J. Phys. A., 2013. 7, 106

[PS58] T. Pati and S. R. Sinha. On the absolute summability factors of Fourier series. *Indian J. Math.*, 1(1):41–54, 1958. 138

[Sch23] I. Schur. Uber eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie. *Sitzungsber. d. Berl. Math. Gesellsch.*, 22:9–20, 1923. 119

[Sha48] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656, 1948. 1

[Sti55] W. F. Stinespring. Positive functions on C^*-algebras. *Proc. Amer. Math. Soc.*, 6:211–216, 1955. 17

[Stø73] E. Størmer. *Positive linear maps of C^*-algebras*. Foundations of quantum mechanics and ordered linear spaces. Advanced Study Inst., Marburg, 1973. Lecture Notes in Phys., Vol. 29, Springer, Berlin, 1974. 28

[Tom57] J. Tomiyama. On the projetion of norm one in W^*-algebras. *Proc. Japan Acad.*, 33:608–612, 1957. 28

[Tur07] S. Turgut. Catalytic transformations for bipartite pure states. *J. Phys. A: Math. Theor.*, 40:12185–12212, 2007. 7, 106, 114, 115, 120, 144, 145

[Ume54] H. Umegaki. Conditional expectation in an operator algebra. *Tohoku Math. J. II*, 6:177–181, 1954. 28, 29

[Wer89] R. F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277–4281, 1989. 10

[Wer05] R. F. Werner. Some open problems in quantum information theory: Catalytic majorization. http://qig.itp.uni-hannover.de/qiproblems/4, 2005. 5, 113, 114