Entanglement in single-shot quantum channel discrimination

by

Daniel Puzzuoli

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Mathematics (Quantum Information)

Waterloo, Ontario, Canada, 2018

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Gilad Gour
Professor, Department of Mathematics and Statistics,
University of Calgary

Supervisor: John Watrous
Professor, Cheriton School of Computer Science,
University of Waterloo

Internal Member: Joseph Emerson
Associate Professor, Department of Applied Mathematics,
University of Waterloo

Internal-External Member: Ashwin Nayak
Professor, Department of Combinatorics and Optimization,
University of Waterloo

Other Member(s): Vern Paulsen
Professor, Department of Pure Mathematics,
University of Waterloo

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

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

Abstract

Single-shot quantum channel discrimination is the fundamental task of determining, given only a single use, which of two known quantum channels is acting on a system. In this thesis we investigate the well-known phenomenon that entanglement to an auxiliary system can provide an advantage in this task. In particular, we consider the questions: (1) How much entanglement is in general necessary to achieve an optimal discrimination strategy? (2) What is the maximal advantage provided by entanglement?

Given a linear map $\Psi : \mathrm{L}(\mathbb{C}^n) \to \mathrm{L}(\mathbb{C}^m)$, its multiplicity maps are defined as the family of linear maps $\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)} : \mathrm{L}(\mathbb{C}^n \otimes \mathbb{C}^k) \to \mathrm{L}(\mathbb{C}^m \otimes \mathbb{C}^k)$, where $\mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}$ is the identity on $\mathrm{L}(\mathbb{C}^k)$. Due to the Holevo-Helstrom theorem, the optimal performance using an auxiliary system of dimension k is quantified in terms of the norm of $\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}$, where Ψ is a linear map that depends on the parameters of the discrimination problem. Hence, the advantage provided by entanglement is represented in the growth of the norm of $\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}$ with k, a classic phenomenon in the theory of operator algebras.

We formalize question (1) by investigating, relative to the input and output dimensions of the channels to be discriminated, how large of an auxiliary system is necessary to achieve an optimal strategy. Mathematically, this is connected to when the norm of $\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}$ stops growing with k. It is well-known that an auxiliary system of dimension equal to the input is always sufficient to achieve an optimal strategy, and that this is sometimes necessary when the output dimension is at least as large as the input. We prove that, even when the output dimension is arbitrarily small compared to the input, it is still sometimes necessary to use an auxiliary system as large as the input to achieve an optimal strategy.

For question (2), we investigate, with respect to a fixed input dimension, how large the gap between the optimal performances with and without entanglement can be. Mathematically, this is quantified by the rate of growth of the norm of $\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}$ in k. It is known that matrix transposition has the fastest possible growth, and we prove that it is essentially the unique linear map with this property. We use this to prove that a discrimination problem defined in terms of the Werner-Holevo channels is essentially the unique game satisfying a norm relation that states that the game can be won with certainty using entanglement, but is hard to win without entanglement.

Along the way, we prove characterizations of the structure of maximal entanglement, as measured according to the entanglement negativity, as well as relative to a large class of entanglement measures. We also give various characterizations of complete trace-norm isometries, and reversible quantum channels.

Acknowledgements

I am extremely grateful to my supervisor John Watrous, for his exceptional guidance and insight into problems in quantum information. All of the work in this thesis is either a direct product of, or a continuation of ideas and questions raised by, collaboration with John.

I am also grateful to my "pseudo-supervisor" Vern Paulsen for always making time for me as if I were his own student, to discuss research and teach me about the world of operator algebras.

I would like to thank Chi-Kwong Li for his enthusiasm and insight into past work, and current/future collaborations.

I would also like to thank various collaborators and friends, who have been a key part of making graduate school stimulating and enjoyable; Holger Haas, Ian Hincks, Connor Paddock, Ben Lovitz, George Nichols, Christopher Wood, and others.

I would like to thank my parents Shirley and Gino Puzzuoli, for their never-ending support and encouragement.

Last but not least, I would like to thank my wife Hina Bandukwala, for making life more fun and interesting, and for helping me to be a better person.

None of the work in this thesis, or any of the other work I've done throughout my graduate studies – or, fundamentally, anything I've done or ever will do – would be possible without the input and help of others.

Table of Contents

List of Tables

List of Figures

Inspirational quote

The physicists say it's not physics,
And the mathematicians say it's not mathematics;
And they are right.

— Karol Życzkowski, on his book *Geometry of Quantum States* during the CMS winter meeting 2016

Chapter 1

Introduction

Quantum channel discrimination is the general task of determining which quantum channel is acting on a quantum system. Many versions of this task exist, and vary depending on the number of uses, types of channels, and resources available. For example, one may consider when perfect discrimination is possible given a finite number of channel uses [1, 17], the influence of memory effects [10], the benefits of adaptive strategies [21], the effects of locality in multiparty settings [16, 41], and also asymptotic versions [23, 3]. Parameter estimation in experiments is another version of this problem [19].

In this thesis we study a particular version of this task, called *single-shot quantum channel discrimination*, in which the goal is to determine, given only a single use, which of two known channels is acting on a system. The individual performing the task must choose a state to feed into the channel, then perform a measurement on the output to guess which channel acted on the state. In general, it can be useful to probe the channels using a state that is entangled to some auxiliary system, then perform a joint measurement on the output and auxiliary systems together. This fact was suggested (somewhat implicitly) in [34] and (more explicitly) in [35], and also proved not to hold for the restricted case of unitary channels in [2] and [9]. See, for example, [9, 57, 56, 47, 28] for investigations on the advantages of using entanglement in this setting, and [55, 53, 18, 68] for other work in the single-shot channel discrimination setting.

Mathematically, optimal performance in this task is quantified using various norms, depending on which resources are available. As such, operational questions in single-shot quantum channel discrimination correspond mathematically to questions about properties of these norms. Let $\mathrm{L}(\mathbb{C}^n)$ denote the linear maps on $\mathbb{C}^n$ (i.e. $n \times n$ matrices with complex entries), and denote the trace-norm of a matrix $A \in \mathrm{L}(\mathbb{C}^n)$ as $\|A\|_1 = \mathrm{Tr}(\sqrt{A^*A})$, where A^* is the usual adjoint of A. For a linear map $\Psi : \mathrm{L}(\mathbb{C}^n) \to \mathrm{L}(\mathbb{C}^m)$, there are three norms of primary relevance in this thesis:

- The induced trace-norm:

$$\|\Psi\|_1 = \max\{\|\Psi(X)\|_1 : X \in \mathrm{L}(\mathbb{C}^n), \|X\|_1 = 1\}, \tag{1.1}$$

- The induced Hermitian trace-norm:

$$\|\Psi\|_{1,H} = \max\{\|\Psi(X)\|_1 : X \in \mathrm{L}(\mathbb{C}^n), \|X\|_1 = 1, X = X^*\}, \tag{1.2}$$

- The completely bounded trace-norm:

$$|||\Psi|||_1 = \sup_{k \geq 1} \|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_1 = \|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)}\|_1, \tag{1.3}$$

  where $\mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}$ is the identity map on $\mathrm{L}(\mathbb{C}^k)$.

As we will outline in Chapter 3, optimal performance in a channel discrimination game using entanglement to an auxiliary system of dimension k is quantified by the norm

$$\left\|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_{1,H}, \tag{1.4}$$

with $\Psi = \lambda\Phi_0 - (1-\lambda)\Phi_1$, where Φ_0 and Φ_1 are the channels to be discriminated, and $\lambda \in [0,1]$ is a probability parameter in the game. Hence, the potential advantage provided by entanglement in this setting is mathematically captured by the potential growth of the above expression with k.

The work in this thesis is motivated by questions regarding the advantage provided by entanglement in this context. Generally speaking, we investigate the following questions:

1. In general, how much entanglement is necessary to achieve the optimal performance in this task?

2. What is the largest possible advantage that entanglement can provide?

Mathematically, the first question concerns when $\left\|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_{1,H}$ stops growing with k, and the second question concerns how much this quantity can grow with k. As these quantities are necessarily unbounded if the input and output dimensions of the channels are allowed to be arbitrary, we investigate these questions relative to fixed (finite) input and output dimensions.

Such mathematical questions find their home within the theory of operator algebras, in which the growth of the norms of the family of maps

$$\{\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)} : k \geq 1\}, \tag{1.5}$$

sometimes called the *multiplicity maps* of Ψ, have been extensively studied [60, 64, 46, 38]. Within this context, the operator norm and completely bounded norm are used, rather than the trace-norm and completely bounded trace-norm. However, due to the duality of these norms, in finite dimensions, many questions can be equivalently phrased in either norm.

Regarding the above two questions, we build on previously known results and examples. For the first question, it is well-known that an auxiliary system as large as the input dimension of the channels is always sufficient and is sometimes necessary, and we show in Chapter 7 that this holds even when the output dimension is arbitrarily small compared to the input. For the second, in Chapter 8 we prove that transposition is the unique linear map saturating the bound [46, Exercise 3.10]

$$\left\|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_1 \leq k\|\Psi\|_1, \tag{1.6}$$

and leverage this to prove that the Werner-Holevo channel discrimination game is the unique game satisfying a norm relation that implies the game can be won with certainty using arbitrary entanglement, but is hard to win without entanglement.

The proofs of these results depend on an understanding of the structure of maximal entanglement, as well as the notions of complete trace-norm isometries and reversible quantum channels. In Chapter 5, we structurally characterize maximally entangled matrices and quantum states as measured by the negativity, as well as a class of entanglement measures that we call weak entanglement measures. We extend this characterization to multi-party settings, which enables the result in Chapter 7 via a monogamy of entanglement type argument. In Chapter 6 we give various characterizations of complete trace-norm isometries and reversible quantum channels. The main contribution of this chapter is a characterization of complete trace-norm isometries in terms of algebraic relations that are specially suited to proving the results in Chapter 8. We also give norm characterizations of the existence of error correcting codes.

Before getting into the results, the preliminary chapters are organized as follows

- In Chapter 2, we give a general background of the mathematics, quantum theory, and notation used in this thesis.
- In Chapter 3, single-shot quantum channel discrimination is formally introduced, and the relevant background theorems are given.
- In Chapter 4, we more formally introduce the problems we will study, and give a more detailed description of the remaining chapters.

Lastly, in Chapter 9, we present some natural lines of continuation of this work, and prove some partial results.

Most of the results in this thesis are drawn from [50] (co-authored with John Watrous) and [49].

Chapter 2

General background and notation

In this chapter we give a general background on the fundamental objects we will use in this thesis: complex Euclidean spaces, a particular version of finite dimensional quantum theory, and the Choi matrix. This background is not meant to be a comprehensive introduction, but is meant to review the concepts we will use while establishing notation.

2.1 Complex Euclidean spaces

In this thesis we will work primarily in $\mathbb{C}^n$, with the usual inner-product which we will denote with $\langle \cdot, \cdot \rangle$, using the convention that it is conjugate linear in the first argument. Occasionally, the symbols $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$, and $\mathcal{W}$ will be used to denote complex Euclidean spaces, either when it is unnecessary to refer to the dimension, or when it is convenient to have a label.

For $x \in \mathbb{C}^n$, the *Euclidean norm* is denoted $\|x\| = \sqrt{\langle x, x \rangle}$. The *elementary vectors* are denoted $e_i \in \mathbb{C}^n$ for $1 \leq i \leq n$, where the vector e_i has a 1 in the i^{th} entry and 0 in all other entries. The *standard basis* for $\mathbb{C}^n$ is the set of elementary vectors $\{e_i\}_{i=1}^n$.

The set of linear maps taking $\mathbb{C}^n \to \mathbb{C}^m$ is denoted $\mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$, and we write $\mathrm{L}(\mathbb{C}^n) = \mathrm{L}(\mathbb{C}^n, \mathbb{C}^n)$. We will not differentiate between $\mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$ and the set of $m \times n$ matrices. For a matrix $A \in \mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$, we will write $T(A)$ or A^{T} to denote its transpose, which is an element of $\mathrm{L}(\mathbb{C}^m, \mathbb{C}^n)$. Transposition is a linear map on $\mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$, and for clarity we will often denote the transpose on $\mathrm{L}(\mathbb{C}^n)$ as T_n, or as $T_{\mathcal{X}}$ for $\mathrm{L}(\mathcal{X})$.

The adjoint of a matrix $A \in \mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$ is denoted $A^* \in \mathrm{L}(\mathbb{C}^m, \mathbb{C}^n)$, and may be defined equivalently as the conjugate transpose of A, or as the unique matrix satisfying

$$\langle A^* y, x \rangle = \langle y, Ax \rangle \tag{2.1}$$

for all $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^m$.

The *elementary matrices* are denoted $E_{ij} \in \mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$ for $1 \leq i \leq m$ and $1 \leq j \leq n$, where the ij^{th} entry of E_{ij} is 1 and the rest are 0.

For complex Euclidean spaces $\mathcal{X}$, we will work with various special subsets of $\mathrm{L}(\mathcal{X})$:

- $\mathrm{Herm}(\mathcal{X}) = \{A \in \mathrm{L}(\mathcal{X}) : A = A^*\}$, the set of *Hermitian*, or *self-adjoint*, matrices. For a matrix $A \in \mathrm{L}(\mathcal{X})$, it holds that A is Hermitian if and only if $\langle x, Ax \rangle \in \mathbb{R}$ for all $x \in \mathcal{X}$.

- $\mathrm{Pos}(\mathcal{X}) = \{P \in \mathrm{L}(\mathcal{X}) : P \geq 0\}$, where we write $P \geq 0$ if $\langle x, Px \rangle \geq 0$ for all $x \in \mathcal{X}$, the set of *positive semidefinite*, or simply *positive*, matrices. For $P \geq 0$, as $\langle x, Px \rangle \in \mathbb{R}$ for all $x \in \mathcal{X}$, it holds that P is necessarily self-adjoint, and hence $\mathrm{Pos}(\mathcal{X}) \subset \mathrm{Herm}(\mathcal{X})$.

- For $m \geq n$, $\mathrm{U}(\mathbb{C}^n, \mathbb{C}^m) = \{A \in \mathrm{L}(\mathbb{C}^n, \mathbb{C}^m) : A^* A = \mathbb{1}_n\}$, the set of *isometries* mapping $\mathbb{C}^n$ into $\mathbb{C}^m$. That is, $\mathrm{U}(\mathbb{C}^n, \mathbb{C}^m)$ is exactly the set of matrices that satisfy $\|Ax\| = \|x\|$ for all $x \in \mathbb{C}^n$. We also write $\mathrm{U}(\mathbb{C}^n) = \mathrm{U}(\mathbb{C}^n, \mathbb{C}^n)$, in which case, an element $U \in \mathrm{U}(\mathbb{C}^n)$ is called a *unitary* and satisfies $U^* U = U U^* = \mathbb{1}_n$.

For a matrix $A \in \mathrm{L}(\mathbb{C}^n)$, the *trace of* A is given by

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} \langle e_i, A e_i \rangle. \tag{2.2}$$

A standard exercise is to show that the orthonormal basis $\{e_i\}_{i=1}^n$ can be replaced by any orthonormal basis for $\mathbb{C}^n$ in the above definition, and the value will not change. The Hilbert-Schmidt inner product is defined on all $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ by

$$\langle A, B \rangle = \mathrm{Tr}(A^* B). \tag{2.3}$$

The *tensor product* of $\mathbb{C}^n$ and $\mathbb{C}^m$ is denoted $\mathbb{C}^n \otimes \mathbb{C}^m$, and may be defined concretely as the linear span of the symbols $\{e_i \otimes e_j : 1 \leq i \leq n, 1 \leq j \leq m\}$, where the e_i and e_j are, respectively, the elementary basis vectors of $\mathbb{C}^n$ and $\mathbb{C}^m$. The set of vectors $\{e_i \otimes e_j\}$ is the elementary basis of $\mathbb{C}^n \otimes \mathbb{C}^m$. Given any two vectors $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^m$, we define

$$x \otimes y = \sum_{i=1}^{n} \sum_{j=1}^{m} \langle e_i, x \rangle \langle e_j, y \rangle e_i \otimes e_j, \tag{2.4}$$

and it is straightforward to verify that $(x, y) \mapsto x \otimes y$ is a bilinear operation. The inner product on $\mathbb{C}^n \otimes \mathbb{C}^m$ is defined on the elementary vectors as

$$\langle e_a \otimes e_i, e_b \otimes e_j \rangle = \langle e_a, e_b \rangle \langle e_i, e_j \rangle, \tag{2.5}$$

and is extended by linearity to all pairs of vectors in $\mathbb{C}^n \otimes \mathbb{C}^m$ (conjugate linear in the first argument). Given matrices $A \in \mathrm{L}(\mathbb{C}^n)$, and $B \in \mathrm{L}(\mathbb{C}^m)$, we define the linear map $A \otimes B \in \mathrm{L}(\mathbb{C}^n \otimes \mathbb{C}^m)$ to act as

$$(A \otimes B)(e_i \otimes e_j) = (Ae_i) \otimes (Be_j). \tag{2.6}$$

It is not difficult to verify that $\mathrm{L}(\mathbb{C}^n \otimes \mathbb{C}^m) = \mathrm{L}(\mathbb{C}^n) \otimes \mathrm{L}(\mathbb{C}^m)$, i.e. all linear maps in $\mathrm{L}(\mathbb{C}^n \otimes \mathbb{C}^m)$ lie in the span of maps of the form $A \otimes B$ for $A \in \mathrm{L}(\mathbb{C}^n)$ and $B \in \mathrm{L}(\mathbb{C}^m)$.

The *swap operator* in $\mathrm{L}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Y} \otimes \mathcal{X})$ is denoted $W_{\mathcal{X},\mathcal{Y}}$, and acts on $x \otimes y \in \mathcal{X} \otimes \mathcal{Y}$ as $W_{\mathcal{X},\mathcal{Y}}(x \otimes y) = y \otimes x$. It may also be explicitly written out as

$$W_{\mathcal{X},\mathcal{Y}} = \sum_{i=1}^{\dim(\mathcal{Y})} \sum_{j=1}^{\dim(\mathcal{X})} E_{i,j} \otimes E_{j,i}, \tag{2.7}$$

where both sets of elementary matrices $E_{i,j}$ and $E_{j,i}$ are of the appropriate dimension for ensuring $W_{\mathcal{X},\mathcal{Y}}$ lies in $\mathrm{L}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Y} \otimes \mathcal{X})$.

The set of linear maps taking $\mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y})$ is denoted $\mathrm{T}(\mathcal{X}, \mathcal{Y})$, and we write $\mathrm{T}(\mathcal{X}) = \mathrm{T}(\mathcal{X}, \mathcal{X})$. The identity in $\mathrm{T}(\mathcal{X})$ is denoted $\mathbb{1}_{\mathrm{L}(\mathcal{X})}$. There are many special subsets of $\mathrm{T}(\mathcal{X}, \mathcal{Y})$ to consider. For a linear map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, we say:

- Φ is *trace-preserving* if $\mathrm{Tr}(\Phi(A)) = \mathrm{Tr}(A)$ for all $A \in \mathrm{L}(\mathcal{X})$.
- Φ is *Hermiticity-preserving* if, for all $A \in \mathrm{Herm}(\mathcal{X})$, $\Phi(A) \in \mathrm{Herm}(\mathcal{Y})$. This condition is equivalent to the condition that $\Phi(X)^* = \Phi(X^*)$ for all $X \in \mathrm{L}(\mathcal{X})$.

- Φ is *positive*, or *positivity preserving*, if, for all $P \in \mathrm{Pos}(\mathcal{X})$, it holds that $\Phi(P) \in \mathrm{Pos}(\mathcal{Y})$. If Φ is positive, then it is necessarily also Hermiticity preserving: any Hermitian matrix $H \in \mathrm{Herm}(\mathcal{X})$ may be decomposed as $H = P - Q$ for $P, Q \geq 0$, and hence $\Phi(H) = \Phi(P) - \Phi(Q)$ is necessarily Hermitian as $\Phi(P), \Phi(Q) \geq 0$.

Given linear maps $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ and $\Psi \in \mathrm{T}(\mathcal{Z}, \mathcal{W})$, the map

$$\Phi \otimes \Psi \in \mathrm{T}(\mathcal{X} \otimes \mathcal{Z}, \mathcal{Y} \otimes \mathcal{W}) \tag{2.8}$$

is defined to act as

$$(\Phi \otimes \Psi)(A \otimes B) = \Phi(A) \otimes \Psi(B) \tag{2.9}$$

for $A \in \mathrm{L}(\mathcal{X})$ and $B \in \mathrm{L}(\mathcal{Z})$, and extending by linearity to all of $\mathrm{L}(\mathcal{X} \otimes \mathcal{Z})$. A special example of a map in $\mathrm{T}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{X})$ that we use special notation for is the *partial trace*, $\mathrm{Tr}_{\mathcal{Y}}$, which acts as

$$\mathrm{Tr}_{\mathcal{Y}}(A \otimes B) = \mathrm{Tr}(B)A. \tag{2.10}$$

Note that the symbol $\mathrm{Tr}_{\mathcal{Y}}$ makes no specific reference to the space $\mathcal{X}$.

Further classes of elements of $\mathrm{T}(\mathcal{X}, \mathcal{Y})$ may be defined in relation to their properties under tensor product:

- Φ is *k-positive* if $(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)})(P) \geq 0$ for all $P \in \mathrm{Pos}(\mathcal{X} \otimes \mathbb{C}^k)$.
- Φ is *completely-positive* if Φ is k-positive for all $k \geq 1$. We denote the set of completely positive maps in $\mathrm{T}(\mathcal{X}, \mathcal{Y})$ as $\mathrm{CP}(\mathcal{X}, \mathcal{Y})$.

Lastly, we will occasionally use a *vectorization* notation. We define a linear map $\mathrm{vec} : \mathrm{L}(\mathcal{X}, \mathcal{Y}) \to \mathcal{Y} \otimes \mathcal{X}$ to act on the basis of elementary matrices as

$$\mathrm{vec}(e_i e_j^*) = e_i \otimes e_j. \tag{2.11}$$

Note that the symbol vec does not make explicit reference to the input and output spaces; it will be clear from the context it is used what those spaces are. A simple identity we make use of is that

$$\mathrm{vec}(ABC) = (A \otimes C^{\mathsf{T}})\mathrm{vec}(B), \tag{2.12}$$

where A, B, and C are any matrices for which the product ABC is well-defined.

2.2 Finite dimensional quantum theory

In this thesis we will work in the following version of finite dimensional quantum theory, which may be viewed simply as an abstract probability theory with certain rules.

A quantum system is associated to a finite dimensional complex Euclidean space $\mathcal{X}$:

- The *states* of the system are called *density matrices*, and are the elements of the set
$$\mathrm{D}(\mathcal{X}) = \{\rho \in \mathrm{Pos}(\mathcal{X}) : \mathrm{Tr}(\rho) = 1\}. \tag{2.13}$$
It is immediate that this is a convex, and the spectral theorem implies that the rank-1 density matrices, which are of the form uu^* for a unit vector $u \in \mathcal{X}$, are the extreme points of the set. Rank-1 density matrices are called *pure states*.

- A measurement of the system with outcomes $\{1, \ldots, k\}$ is specified by a set of positive semidefinite matrices $\{\mu(1), \ldots, \mu(k)\} \subset \mathrm{Pos}(\mathcal{X})$, satisfying $\sum_{i=1}^{k} \mu(i) = \mathbb{1}_n$. The probability of observing outcome i is given by the expression
$$p(i) = \langle \mu(i), \rho \rangle. \tag{2.14}$$
The requirement that $\mu(i) \geq 0$ ensures $p(i) \geq 0$, and the requirement that $\sum_{i=1}^{k} \mu(i) = \mathbb{1}_n$ ensures that $\sum_{i=1}^{k} p(i) = 1$, i.e. the $p(i)$ are actually probabilities. The form of measurement presented here may be deduced from requiring that measurement is a linear map taking density matrices to probability distributions over the set of outcomes.

- In this thesis, we assume that the system is destroyed after measurement, and as such a single copy of a system may only ever be measured once.[1]

Another important building block of quantum theory is a rule for system composition. Given two quantum systems, individually associated respectively to the complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, the complex Euclidean space associated to the two systems together is $\mathcal{X} \otimes \mathcal{Y}$. Hence, the states of the composite system are elements of $\mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$.

[1] A key aspect of any formulation of quantum theory is the form of the state after measurement. In other formulations, the system is not destroyed, but the state will be somehow altered by measurement. We choose to formulate measurement to destroy the system, as in the context we will consider in this thesis, nothing can be gained from repeated measurement of the system, and so we may as well assume the system can only be measured once.

The last element of quantum theory we need is that of the physical transformations between systems. Given two quantum systems associated to complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, the allowable physical transformations from one to the other are called *quantum channels* and are given by completely positive and trace-preserving linear maps in $\mathrm{T}(\mathcal{X},\mathcal{Y})$. We denote the set of quantum channels by $\mathrm{C}(\mathcal{X},\mathcal{Y})$. The reason for this choice of mathematical representation is that it is exactly the set of linear maps that take density matrices to density matrices, even when acting only on part of composite system. That is, given a quantum channel $\Phi \in \mathrm{C}(\mathcal{X},\mathcal{Y})$ and an arbitrary density matrix $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Z})$, it holds that:

- $(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\rho) \geq 0$, as Φ is completely positive.
- $\mathrm{Tr}((\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\rho)) = \mathrm{Tr}(\rho) = 1$, as Φ is trace-preserving.

Hence, $(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\rho) \in \mathrm{D}(\mathcal{Y} \otimes \mathcal{Z})$.

2.2.1 Entanglement and separability

An important concept in quantum information is that of *entanglement*. A natural form for a density matrix $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ is

$$\rho = \sum_{i=1}^{k} p(i)\rho_i \otimes \sigma_i, \tag{2.15}$$

where $p(i)$ is a probability distribution, and $\rho_i \in \mathrm{D}(\mathcal{X})$ and $\sigma_i \in \mathrm{D}(\mathcal{Y})$ for all i. Density matrices of this form are called *separable*, and operationally represent a state where, with probability $p(i)$, independently, the state of $\mathcal{X}$ is ρ_i and the state of $\mathcal{Y}$ is σ_i.

A state $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ is called *entangled* if it is not *separable*, i.e. if it does not have a decomposition of the form in Equation (2.15).

Definition 2.1. A pure state $uu^* \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^m)$ is called *maximally entangled* if, for $r = \min(n,m)$, there exists orthonormal sets

$$\{x_i\}_{i=1}^r \subset \mathbb{C}^n \text{ and } \{y_i\}_{i=1}^r \subset \mathbb{C}^m \tag{2.16}$$

for which

$$u = \sqrt{\frac{1}{r}} \sum_{i=1}^{r} x_i \otimes y_i. \tag{2.17}$$

Equivalently, in terms of vectorization notation, a unit vector $u \in \mathbb{C}^n \otimes \mathbb{C}^m$ is maximally entangled if, for $r = \min(n, m)$, $u = \sqrt{\frac{1}{r}}\mathrm{vec}(A)$ for $A \in \mathrm{L}(\mathbb{C}^m, \mathbb{C}^n)$, where A is an isometry if $m \leq n$, or A^* is an isometry if $m > n$.

We also say that a pure state $\rho = uu^* \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^m)$ is *maximally entangled* if u is maximally entangled.

For a complex Euclidean space $\mathcal{X}$, we let $\tau_{\mathcal{X}} \in \mathrm{D}(\mathcal{X} \otimes \mathcal{X})$ denote the *canonical maximally entangled state*, which is defined as

$$\tau_{\mathcal{X}} = \frac{1}{n} \sum_{i,j=1}^{\dim(\mathcal{X})} E_{i,j} \otimes E_{i,j} = \frac{1}{n}\mathrm{vec}(\mathbb{1}_{\mathcal{X}})\mathrm{vec}(\mathbb{1}_{\mathcal{X}})^*. \tag{2.18}$$

We will also write $\tau_n = \tau_{\mathbb{C}^n}$, when the dimension of the space is explicit. Observe as well that

$$(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes T_{\mathcal{X}})(\tau_{\mathcal{X}}) = \frac{1}{n} W_{\mathcal{X},\mathcal{X}}, \tag{2.19}$$

where $W_{\mathcal{X},\mathcal{X}} \in \mathrm{L}(\mathcal{X} \otimes \mathcal{X})$ is the swap operator.

In this thesis we will consider generalizations of maximal entanglement to general density matrices, not just pure states.

2.3 The Choi matrix

In [12], Choi introduced a now classic object for studying linear maps between matrices. For a linear map $\Phi \in \mathrm{T}(\mathbb{C}^n, \mathcal{Y})$, the *Choi matrix* of Φ, which we denote $J(\Phi)$, is defined as:

$$J(\Phi) = \sum_{a,b=1}^{n} \Phi(E_{a,b}) \otimes E_{a,b}, \tag{2.20}$$

which may be viewed as an $n \times n$ block matrix, where the $(a, b)^{th}$ block is $\Phi(E_{a,b})$. The function $J : \mathrm{T}(\mathbb{C}^n, \mathcal{Y}) \to \mathrm{L}(\mathcal{Y} \otimes \mathbb{C}^n)$ is linear due to the linearity of its argument, and as $\{E_{a,b}\}_{a,b=1}^n$ is a basis for $\mathrm{L}(\mathbb{C}^n)$, J is also a bijection. Thus, J is a vector space isomorphism, and is thus commonly referred to as the *Choi isomorphism*.

The prevalence of the use of this map in finite dimensional quantum information is that many naturally motivated properties of Φ are naturally captured by corresponding properties of the matrix $J(\Phi)$. For example:

- Φ is Hermiticity preserving if and only if $J(\Phi)$ is Hermitian.
- Φ is completely positive if and only if $J(\Phi)$ is positive semidefinite [12].
- Φ is trace-preserving if and only if $\mathrm{Tr}_{\mathcal{Y}}(J(\Phi)) = \mathbb{1}_n$.
- Φ is an entanglement breaking channel[2] if and only if $\frac{1}{n}J(\Phi)$ is a separable density matrix.

Two results of this thesis are new natural correspondences, which we state informally here:

- Φ is a complete trace-norm isometry if and only if $\frac{1}{n}J(\Phi)$ is maximally entangled with respect to the negativity.
- Φ is a reversible quantum channel if and only if $\frac{1}{n}J(\Phi)$ is maximally entangled with respect to any weak entanglement measure.

2.4 Norms on $\mathrm{L}(\mathcal{X}, \mathcal{Y})$ and $\mathrm{T}(\mathcal{X}, \mathcal{Y})$

We will make use of three norms on matrices. For $A \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$, define:

- The operator norm:

$$\|A\| = \sup\{\|Ax\| : x \in \mathbb{C}^n, \|x\| = 1\}. \tag{2.21}$$

- The Hilbert-Schmidt norm:

$$\|A\|_2 = \sqrt{\langle A, A\rangle} = \sqrt{\mathrm{Tr}(A^* A)}. \tag{2.22}$$

- The trace-norm:

$$\|A\|_1 = \mathrm{Tr}(\sqrt{A^* A}). \tag{2.23}$$

In terms of the singular values in decreasing order $\sigma_1, \ldots, \sigma_r$ of a matrix A, it holds that $\|A\| = \sigma_1$, $\|A\|_2 = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$, and $\|A\|_1 = \sigma_1 + \cdots + \sigma_r$. We may also write these norms in terms of optimizations:

[2] That is, Φ is a channel, and for any density matrix $\rho \in \mathrm{D}(\mathbb{C}^n \otimes \mathcal{Z})$, $(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\rho)$ is separable.

- $\|A\| = \max\{|\langle X, A\rangle| : X \in \mathrm{L}(\mathcal{X}, \mathcal{Y}), \|X\|_1 \leq 1\}$,
- $\|A\|_2 = \max\{|\langle X, A\rangle| : X \in \mathrm{L}(\mathcal{X}, \mathcal{Y}), \|X\|_2 \leq 1\}$, and
- $\|A\|_1 = \max\{|\langle X, A\rangle| : X \in \mathrm{L}(\mathcal{X}, \mathcal{Y}), \|X\| \leq 1\}$.

Note that the operator norm is characterized as an optimization of inner products over matrices in the trace-norm unit ball, and similarly the trace-norm is characterized as an optimization over matrices in the operator norm unit ball. This fact is referred to as the *duality* of these norms. When $m \geq n$, we may also further simplify the optimization for $\|A\|_1$ as

$$\|A\|_1 = \max\{|\langle U, A\rangle| : U \in \mathrm{U}(\mathcal{X}, \mathcal{Y})\}. \tag{2.24}$$

A simple fact we will make use of is that, for $A \in \mathrm{L}(\mathcal{X})$, it holds that $\|A\|_1 = \mathrm{Tr}(A)$ if and only if $A \geq 0$.

We will also use the following norms on $\mathrm{T}(\mathcal{X}, \mathcal{Y})$. For $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, define:

- The induced trace-norm:

$$\|\Phi\|_1 = \max\{\|\Phi(X)\|_1 : X \in \mathrm{L}(\mathcal{X}), \|X\|_1 = 1\}. \tag{2.25}$$

- The induced "Hermitian" trace-norm:

$$\|\Phi\|_{1,H} = \max\{\|\Phi(H)\|_1 : H \in \mathrm{Herm}(\mathcal{X}), \|H\|_1 = 1\}. \tag{2.26}$$

- The induced operator norm:

$$\|\Phi\| = \max\{\|\Phi(X)\| : X \in \mathrm{L}(\mathcal{X}), \|X\| = 1\}. \tag{2.27}$$

- The induced "Hermitian" operator norm:

$$\|\Phi\|_H = \max\{\|\Phi(H)\| : H \in \mathrm{Herm}(\mathcal{X}), \|H\| = 1\}. \tag{2.28}$$

Note that all of the above norms are just usual operator norms of the Φ for various input spaces and norms. We will also use completely bounded versions of these norms, the completely bounded trace norm:

$$|||\Phi|||_1 = \sup\{\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_1 : k \in \mathbb{N}\}, \tag{2.29}$$

and the completely bounded norm:

$$\|\Phi\|_{\mathrm{cb}} = \sup\{\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\| : k \in \mathbb{N}\} \tag{2.30}$$

Due to the duality of the operator and trace-norm, for general $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ it holds that

$$\|\Phi\|_1 = \|\Phi^*\|, \text{ and } |||\Phi|||_1 = \|\Phi^*\|_{\mathrm{cb}}. \tag{2.31}$$

Similarly, if $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is Hermiticity preserving, it also holds that

$$\|\Phi\|_{1,H} = \|\Phi^*\|_H. \tag{2.32}$$

These relations enable interconversion of facts from one norm to the other.

We will assume various facts about these norms:

- For any $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, it holds that $|||\Phi|||_1 = \|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\|_1$ [69, Theorem 3.46], and equivalently, $\|\Phi\|_{\mathrm{cb}} = \|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\|$ [46, Proposition 8.11].
- For $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ Hermiticity preserving, it holds that

$$|||\Phi|||_1 = \|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\|_{1,H} \tag{2.33}$$

  [69, Theorem 3.51], equivalently $\|\Phi\|_{\mathrm{cb}} = \|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\|_H$.
- For $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, it holds that

$$\begin{aligned} \|\Phi\|_{1,H} &= \max\{\|\Phi(\rho)\|_1 : \rho \in \mathrm{D}(\mathcal{X})\} && (2.34)\\ &= \max\{\|\Phi(uu^*)\|_1 : u \in \mathcal{X}, \|u\| = 1\}. && (2.35) \end{aligned}$$

- For $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ positive, it holds that $\|\Phi\|_1 = \|\Phi\|_{1,H}$ [69, Theorem 3.39]. Equivalently, $\|\Phi\| = \|\Phi(\mathbb{1}_{\mathcal{X}})\|$ [46, Corollary 2.9]. From this it may be further deduced that, for completely positive $\Phi \in \mathrm{CP}(\mathcal{X}, \mathcal{Y})$ it holds that $|||\Phi|||_1 = \|\Phi\|_1 = \|\Phi\|_{1,H}$, and equivalently $\|\Phi\|_{\mathrm{cb}} = \|\Phi\| = \|\Phi(\mathbb{1}_{\mathcal{X}})\|$.

Lastly, we draw attention to a special case of Wittstock's decomposition theorem [46, Theorem 8.5]. It states that, for a Hermiticity preserving map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, there exists a completely positive map $\Psi \in \mathrm{CP}(\mathcal{X}, \mathcal{Y})$ for which both $\Psi + \Phi$ and $\Psi - \Phi$ are both completely positive, and $\|\Psi\|_{\mathrm{cb}} \leq \|\Phi\|_{\mathrm{cb}}$ (or, equivalently, we may alternatively conclude the existence of such a Ψ with $|||\Psi|||_1 \leq |||\Phi|||_1$). In this thesis we consider norms of Hermiticity preserving maps, as well as decompositions of them into completely positive maps satisfying certain norm relations, and hence, even though we will not directly use Wittstock's decomposition theorem, some of the results and concepts bare similarity to it.

Chapter 3

Single-shot quantum state and channel discrimination

As described in the introduction, single-shot quantum channel discrimination is the task of determining, given only a single use, which of two known quantum channels is acting on a system. This problem has been studied for a long time in various settings due to its simplicity, as well as its relationship to commonly used mathematical objects in the communities of quantum information and operator algebras.

In this chapter, we give the standard formalization of single-shot quantum channel discrimination and present the background of foundational theorems. As it is our focus of study, we focus on certain aspects of the use of entanglement in this setting.

To begin, we introduce single-shot quantum state discrimination, which is essentially a primitive in the channel discrimination version.

3.1 Single-shot quantum state discrimination

A single-shot quantum state discrimination game is a single-player game specified by a pair of density matrices $\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})$, and a probability $\lambda \in [0, 1]$. The game proceeds as follows:

1. The referee samples a bit $\alpha \in \{0, 1\}$ according to $p(0) = \lambda$, $p(1) = 1 - \lambda$.
2. The player is given a single copy of ρ_α.

3. The player must guess α by measuring ρ_α, with the goal of optimizing the probability of guessing correctly.

All that the player can do in this situation is to choose a two outcome measurement $\mu(0), \mu(1) \in \mathrm{Pos}(\mathcal{X})$, where the player guesses the state was ρ_0 if they get the 0 outcome, and ρ_1 if they get the 1 outcome.[1] Given such a measurement, and denoting the outcome the player gets as $\beta \in \{0, 1\}$, the probability that they correctly guess α is then given by

$$\begin{aligned} & P(\lambda = 0 \text{ and } \beta = 0) + P(\lambda = 1 \text{ and } \beta = 1) && (3.1)\\ &\qquad = P(\alpha = 0)P(\beta = 0|\alpha = 0) + P(\alpha = 1)P(\beta = 1|\alpha = 1) && (3.2)\\ &\qquad = \lambda\langle \mu(0), \rho_0\rangle + (1 - \lambda)\langle \mu(1), \rho_1\rangle. && (3.3) \end{aligned}$$

Hence, the optimal performance of the player is given by the maximization of the above expression over all two-outcome measurements. The Holevo-Helstrom theorem [22, 24] gives a closed form expression for this value.

Theorem 3.1 (Holevo-Helstrom theorem)**.** *Let $\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})$ be density matrices, and let $\lambda \in [0, 1]$. For any measurement $\mu : \{0, 1\} \to \mathrm{Pos}(\mathcal{X})$, it holds that*

$$\lambda\langle \mu(0), \rho_0\rangle + (1 - \lambda)\langle \mu(1), \rho_1\rangle \leq \frac{1}{2} + \frac{1}{2}\left\| \lambda\rho_0 - (1 - \lambda)\rho_1 \right\|_1. \qquad (3.4)$$

Moreover, there exists a choice of projective measurement for which equality is achieved.[2]

For a proof of the above theorem, see [69, Theorem 3.4].

3.2 Single-shot quantum channel discrimination

The task of single-shot quantum channel discrimination is defined similarly to state discrimination. It is formulated as a single-player game parameterized

[1] As the player must respond with a 1 or 0, any post-processing of the post-measurement state (if one chooses a formulation of measurement in which the system is not destroyed) or of measurement results with possibly more than two outcomes may simply be combined into a single measurement with only two outcomes.

[2] It will not be relevant for our purposes, but a choice of projective measurement achieving the bound is setting $\mu(0)$ and $\mu(1)$ to be, respectively, the projections onto the positive and non-positive parts of the matrix $\lambda\rho_0 - (1 - \lambda)\rho_1$.

by a pair of channels $\Gamma_0, \Gamma_1 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ to be discriminated, and a probability $\lambda \in [0,1]$ serving as a free parameter. The triple $(\lambda, \Gamma_0, \Gamma_1)$ is known to the player, and the game proceeds as follows:

1. The referee samples a bit $\alpha \in \{0,1\}$ according to $p(0) = \lambda$, $p(1) = 1 - \lambda$.
2. The player is given a single use of Γ_α (i.e. they must probe it using a single input state).
3. The player guesses α (after making a single measurement on the output), with the goal of maximizing the probability that they guess correctly.

Player strategies may be assumed to consist only of a choice of input state ρ, and a two-outcome measurement μ.

We will consider the scenario in which the player potentially has access to an auxiliary system $\mathcal{Z}$, and their strategies consist of preparing (possibly entangled) states $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Z})$, passing the $\mathcal{X}$ system to the referee while keeping $\mathcal{Z}$, then attempting to discriminate the outputs $(\Gamma_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\rho)$ and $(\Gamma_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\rho)$ (see Figure 3.1).

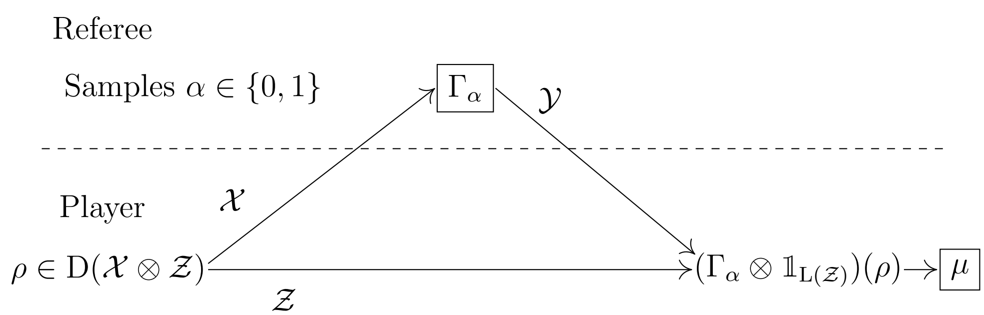


Figure 3.1: Diagrammatic representation of a player strategy in single-shot quantum channel discrimination. The dashed line separates actions taken by the player and those taken by the referee.

Given the Holevo-Helstrom theorem, if the player chooses to use the input state $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Z})$, then after optimizing over measurements, their success probability in the game is

$$\frac{1}{2} + \frac{1}{2}\|\lambda(\Gamma_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\rho) - (1-\lambda)(\Gamma_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\rho)\|_1. \tag{3.5}$$

Hence, we have the following theorem.

Theorem 3.2 (Holevo-Helstrom for channels)**.** *Let $\Gamma_0, \Gamma_1 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ be quantum channels and let $\lambda \in [0,1]$. For the quantum channel discrimination game specified by $(\lambda, \Gamma_0, \Gamma_1)$, the following statements hold.*

1. *If the player has access to an auxiliary system $\mathcal{Z}$ and is able to prepare any state in $\mathrm{D}(\mathcal{X} \otimes \mathcal{Z})$ and perform any measurement, their optimal probability of success in the game is*

$$\frac{1}{2} + \frac{1}{2}\left\|(\lambda\Gamma_0 - (1-\lambda)\Gamma_1) \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right\|_{1,H}. \tag{3.6}$$

2. *If the player has no restrictions on the strategy used (i.e. they can use any auxiliary system, any density matrix, and any measurement), then their optimal probability of success in the game is*

$$\frac{1}{2} + \frac{1}{2}|||\lambda\Gamma_0 - (1-\lambda)\Gamma_1|||_1. \tag{3.7}$$

Moreover, there exists a unit vector $u \in \mathcal{X} \otimes \mathcal{X}$ for which

$$|||\lambda\Gamma_0 - (1-\lambda)\Gamma_1|||_1 = \left\|\lambda(\Gamma_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(uu^*) - (1-\lambda)(\Gamma_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(uu^*)\right\|_1, \tag{3.8}$$

i.e. the player can always achieve optimal performance using an auxiliary system the same size as the input of the channels.

Proof. The first statement follows from Equation (3.5) and the general fact that, for any linear map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, it holds that

$$\begin{aligned} \|\Phi\|_{1,H} &= \max\{\|\Phi(H)\|_1 : H \in \mathrm{Herm}(\mathcal{X}), \|H\|_1 = 1\} && (3.9)\\ &= \max\{\|\Phi(\rho)\|_1 : \rho \in \mathrm{D}(\mathcal{X})\} && (3.10)\\ &= \max\{\|\Phi(uu^*)\|_1 : u \in \mathcal{X}, \|u\| = 1\}. && (3.11) \end{aligned}$$

The second statement follows from the following fact, given as Theorem 3.51 in [69]: For a Hermiticity preserving map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, it holds that

$$|||\Phi|||_1 = \left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\right\|_{1,H}. \tag{3.12}$$

□

Hence, the norm $\|\cdot\|_{1,H}$ and the completely bounded trace-norm $|||\cdot|||_1$ have operational interpretation in terms of the above game. Furthermore, the potential increase in the value of

$$\|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_{1,H} \tag{3.13}$$

with k has operational interpretation in terms of the usefulness of entanglement in this setting.

Chapter 4

Motivation and outline of results

In this thesis we are motivated by questions regarding the role of entanglement in single-shot quantum channel discrimination. In this chapter we give explicit formalizations of the questions investigated in the thesis, and give an outline of the results.

To start, we present two well-known examples that have many special properties that help to guide meaningful questions.

4.1 Special example: matrix transposition

Matrix transposition has long served as a classic example of a map for which the norms $\|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_1$ grow with k. In general, it holds that:

$$\|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_1 \leq k\|\Psi\|_1, \tag{4.1}$$

and hence the rate of growth of these norms is bounded [46, Exercise 3.10]. In fact, transposition is known to saturate the above inequality, and so is also the canonical example of the most extreme version of this behaviour. Specifically, letting T_n denote transposition on $\mathrm{L}(\mathbb{C}^n)$, for $k \leq n$ it holds that

$$\|T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_1 = k = k\|T_n\|_1. \tag{4.2}$$

This fact was originally given in [63], but we will reprove it here in our own notation for later use.

In quantum information, the transpose provides one of the most-used tests of whether or not a density matrix is entangled. Specifically, the positive partial transpose test states that, for a density matrix $\rho \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^m)$, if the matrix

$$(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(\rho) \tag{4.3}$$

is not positive semi-definite, then ρ is necessarily entangled. It is easy to see that this statement generalizes to *any* positive map (not just the transpose), and in fact it is a classic result that the existence of a positive map $\Lambda \in \mathrm{T}(\mathbb{C}^n)$ for which $(\Lambda \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(\rho)$ is not positive semidefinite is in fact necessary for ρ to be entangled [25]. In general, it is not sufficient to check whether the matrix in Equation (4.3) is positive semi-definite to conclude that ρ is separable, but it is known that if this condition is satisfied, then the distillable entanglement of ρ is zero [26]. Hence, the transpose seems to be an unusually good map for detecting entanglement.

In [66], Vidal and Werner defined an entanglement measure known as the *negativity*, which quantifies how much $(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(\rho)$ fails to remain positive semidefinite. For a density matrix $\rho \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^m)$ the *negativity* is defined as

$$\rho \mapsto \|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(\rho)\|_1. \tag{4.4}$$

Note that this differs from the definition in [66] by multiplicative and additive scalars, and so referring to the above quantity as the negativity is an abuse of terminology. Nevertheless, we will always work directly with the above form, which we refer to loosely as the negativity.[1]

We now prove various relevant facts about the behaviour of transposition.

As proven in [66, Proposition 8], when evaluated on a rank-1 matrix, the negativity takes a simple form.

Proposition 4.1 (Vidal and Werner)**.** *For $A, B \in \mathrm{L}(\mathbb{C}^m, \mathbb{C}^n)$ it holds that*

$$\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(\mathrm{vec}(A)\mathrm{vec}(B)^*)\|_1 = \|A\|_1 \|B\|_1. \tag{4.5}$$

[1] As an interesting aside, in [43] it is shown that the proof bounding distillable entanglement in terms of entanglement negativity depends only on the property that all tensor powers of the transpose are positive. They pose the still-open problem of whether or not the transpose is essentially the unique positive (but not completely positive) map with this property.

Proof. Using properties of vectorization, we have

$$\begin{aligned}
&(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(\mathrm{vec}(A)\mathrm{vec}(B)^*) &(4.6)\\
&\qquad = (T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})((A \otimes \mathbb{1}_m)\mathrm{vec}(\mathbb{1}_m)\mathrm{vec}(\mathbb{1}_m)^*(B^* \otimes \mathbb{1}_m)) &(4.7)\\
&\qquad = (\overline{B} \otimes \mathbb{1}_m)W_{\mathbb{C}^m,\mathbb{C}^m}(A^{\mathsf{T}} \otimes \mathbb{1}_m) &(4.8)\\
&\qquad = (\overline{B} \otimes A^{\mathsf{T}})W_{\mathbb{C}^n,\mathbb{C}^m}, &(4.9)
\end{aligned}$$

and hence

$$\begin{aligned}
\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(\mathrm{vec}(A)\mathrm{vec}(B)^*)\|_1 &= \|(\overline{B} \otimes A^{\mathsf{T}})W_{\mathbb{C}^n,\mathbb{C}^m}\|_1 &(4.10)\\
&= \|A\|_1\|B\|_1. &(4.11)
\end{aligned}$$

□

Note that [66, Proposition 8] is proven only for the case $A = B$. From this we may deduce the form of rank-1 matrices that maximize the norm of the linear map $\mathrm{T}_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}$.

Proposition 4.2. *For unit vectors $u, v \in \mathbb{C}^n \otimes \mathbb{C}^m$, it holds that*

$$\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(uv^*)\|_1 \leq \min(n, m), \tag{4.12}$$

with equality if and only if both u and v are maximally entangled. In particular this implies

$$\|T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\|_{1,H} = \|T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\|_1 = \min(n, m). \tag{4.13}$$

Proof. For unit vectors $u, v \in \mathbb{C}^n \otimes \mathbb{C}^m$, let $A, B \in \mathrm{L}(\mathbb{C}^m \otimes \mathbb{C}^n)$ be the matrices satisfying $u = \mathrm{vec}(A)$ and $v = \mathrm{vec}(B)$. By Proposition 4.1,

$$\begin{aligned}
\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(\mathrm{vec}(A)\mathrm{vec}(B)^*)\|_1 &= \|A\|_1\|B\|_1\\
&\leq \min(n, m)\|A\|_2\|B\|_2 = \min(n, m),
\end{aligned} \tag{4.14}$$

where the inequality follows from the inequality $\|A\|_1 \leq \sqrt{\min(n, m)}\|A\|_2$, which holds with equality if and only if either A or A^* is a scalar multiple of an isometry. Hence, we have the inequality in Equation (4.12), with equality holding if and and only if u and v are maximally entangled.

Equation (4.13) follows as the induced 1-norm can be written as an optimization over matrices of the form uv^* for unit vectors $u, v \in \mathbb{C}^n \otimes \mathbb{C}^m$. □

The equality condition for Equation (4.12), when $u = v$, is the well known fact that the only pure states which maximize negativity are maximally entangled.

4.2 Special example: Werner-Holevo channels

The Werner-Holevo channels, defined in [70], provide a well-known example in quantum channel discrimination that has several special properties with regard to the usefulness of entanglement.

Definition 4.3. For an integer $n \geq 2$, the *Werner-Holevo channels* are denoted $\Phi_n^{(0)}, \Phi_n^{(1)} \in \mathrm{C}(\mathbb{C}^n)$, and are defined to act on all $X \in \mathrm{L}(\mathbb{C}^n)$ as

$$\Phi_n^{(0)}(X) = \frac{1}{n+1}(\mathrm{Tr}(X)\mathbb{1}_n + X^{\mathsf{T}}), \text{ and } \Phi_n^{(1)}(X) = \frac{1}{n-1}(\mathrm{Tr}(X)\mathbb{1}_n - X^{\mathsf{T}}). \tag{4.15}$$

That the Werner-Holevo channels are in fact channels is straightforward to verify: that they are trace preserving can be checked by inspection, and complete positivity may be checked by looking at their Choi matrices. Explicitly,

$$J(\Phi_n^{(0)}) = \frac{1}{n+1}(\mathbb{1}_n \otimes \mathbb{1}_n + W_{\mathbb{C}^n,\mathbb{C}^n}) \tag{4.16}$$

and

$$J(\Phi_n^{(0)}) = \frac{1}{n-1}(\mathbb{1}_n \otimes \mathbb{1}_n - W_{\mathbb{C}^n,\mathbb{C}^n}), \tag{4.17}$$

from which it follows that $J(\Phi_n^{(0)}) \geq 0$ and $J(\Phi_n^{(1)}) \geq 0$, as $W_{\mathbb{C}^n,\mathbb{C}^n}$ is Hermitian and $\|W_{\mathbb{C}^n,\mathbb{C}^n}\| = 1$.

The Werner-Holevo channels satisfy a particular relation which makes their relevance as an example in single-shot quantum channel discrimination immediately clear. For the probability $\lambda_n = \frac{n+1}{2n}$, it holds that

$$\lambda_n \Phi_n^{(0)} - (1 - \lambda_n)\Phi_n^{(1)} = \frac{1}{n} T_n. \tag{4.18}$$

Hence, the operational properties of the game specified by $(\lambda_n, \Phi_n^{(0)}, \Phi_n^{(1)})$ directly correspond to norm properties of the transpose.

Proposition 4.4. *For a positive integer $n \geq 2$, let $\Phi_n^{(0)}, \Phi_n^{(1)} \in \mathrm{C}(\mathbb{C}^n)$ be the Werner-Holevo channels, and let $\lambda_n = \frac{n+1}{2n}$. For the single-shot channel* *discrimination game specified by the triple* $(\lambda_n, \Phi_n^{(0)}, \Phi_n^{(1)})$*, the optimal success probability achievable by a player using an auxiliary system of dimension* k *is*

$$\frac{1}{2} + \frac{\min(k, n)}{2n}. \tag{4.19}$$

Proof. By Equation (4.18) along with the Holevo-Helstrom Theorem for channels (Theorem 3.2), the optimal success probability of a player using an auxiliary system of dimension k is given by

$$\frac{1}{2} + \frac{1}{2n} \| T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)} \|_{1,H}, \tag{4.20}$$

and by Proposition 4.2 this evaluates to the desired expression. □

Observe the following facts about the above game:

- Optimal (and perfect) performance can be achieved if and only if the auxiliary system is at least as large as the input dimension of the channels. Hence, this example demonstrates that sometimes having an auxiliary system of dimension equal to the input is *necessary* for optimal discrimination.

- If the player uses no entanglement (corresponding to using an auxiliary system of dimension $k = 1$), the player cannot do better than the *trivial strategy* of simply guessing that $\Phi_n^{(0)}$ was chosen by the referee, without bothering to optimize the input or measurement. That is, without entanglement, the optimal success probability is exactly $\lambda_n = \frac{n+1}{2n} = \frac{1}{2} + \frac{1}{2n}$, which is exactly the success probability achievable if the player always guesses $\Phi_n^{(0)}$. In a sense, this means that nothing can be learned about which channel acted in this game if no entanglement is available.

- There is a large gap between the optimal performances with and without entanglement: the game can be won with certainty using entanglement, but without entanglement one cannot do better than the trivial strategy, and the value of the trivial strategy is small. In particular, it satisfies the norm relation

$$1 = \| \lambda_n \Phi_n^{(0)} - (1 - \lambda_n) \Phi_n^{(1)} \|_1 = n \| \lambda_n \Phi_n^{(0)} - (1 - \lambda_n) \Phi_n^{(1)} \|_{1,H}. \tag{4.21}$$

4.3 Questions on entanglement in single-shot quantum channel discrimination

As described in the introduction, the work in this thesis is motivated by the following two over-arching questions regarding the advantage provided by entanglement in single-shot quantum channel discrimination.

1. How much entanglement is in general necessary to achieve an optimal strategy?
2. What is the largest possible advantage provided by entanglement?

A quick observation is that, for these questions to be interesting, they must be asked in relation to some restriction on dimension of the channels in the game. In particular, the Werner-Holevo channel example shows that if no restriction is placed on the dimensions of the channels, then the answer to both questions is "arbitrary".

We consider the following formalizations of the above questions.

Question 1. How large of an auxiliary system is necessary to perform optimally in all channel discrimination games with input dimension n and output dimension m? Mathematically, for what k does it hold that

$$\|(\lambda\Gamma_0 - (1-\lambda)\Gamma_1) \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_{1,H} = |||\lambda\Gamma_0 - (1-\lambda)\Gamma_1|||_1 \tag{4.22}$$

for all channels $\Gamma_0, \Gamma_1 \in \mathrm{C}(\mathbb{C}^n, \mathbb{C}^m)$ and probabilities $\lambda \in [0,1]$?

Question 2. For what quantum channel discrimination games with input dimension n and output dimension m is the advantage provided by entanglement maximal? Mathematically: Characterize the channels $\Gamma_0, \Gamma_1 \in \mathrm{C}(\mathbb{C}^n, \mathbb{C}^m)$ and probabilities $\lambda \in [0,1]$ for which the gap between the norms

$$\|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\|_{1,H} \text{ and } |||\lambda\Gamma_0 - (1-\lambda)\Gamma_1|||_1 \tag{4.23}$$

is as large as possible.

We note that, for Question 1, we are using the dimension of the auxiliary system as a proxy for entanglement. It is conceivable that there may exist channel discrimination games in which the states achieving an optimal strategy require high auxiliary dimension, but not a lot of entanglement. While this is a possibility, the games we consider in relation to this question all require maximal entanglement to achieve an optimal strategy, and as such, the dimension of the auxiliary system directly corresponds to the amount of entanglement required.

4.4 Outline

The following chapters give mathematical results that will be used later for addressing the above questions.

- In Chapter 5, the structure of maximal entanglement is considered. In particular, we prove that the entanglement negativity achieves its maximal value on an arbitrary matrix if and only if it has a simple structure containing the maximally entangled state. On the set of density matrices, we prove that essentially the same result holds for a wide class of entanglement measures. In both of these contexts we generalize these results to a multi-party setting.
- In Chapter 6, we prove various characterizations of complete trace-norm isometries and reversible quantum channels. The main contribution is a characterization of complete trace-norm isometries in terms of certain algebraic relations. We also show that a linear map is a complete trace-norm isometry if and only if its Choi matrix is maximally entangled.

The results directly addressing channel discrimination appear in the following chapters.

- In Chapter 7 we address Question 1. The Holevo-Helstrom theorem for channels (Theorem 3.2) states that an auxiliary system of dimension $k = n$ is always sufficient, and the Werner-Holevo channel game presented in the preceding section gives that this is sometimes necessary when $m \geq n$. Thus, it is natural to ask if it is generally possible to use an auxiliary system with dimension smaller than the input dimension when $m < n$. In Chapter 7 we show that this is not the case by exhibiting a family of channel discrimination games in which the output dimension can be made arbitrarily small compared to the input, but for which it is still always necessary to use an auxiliary system as large as the input to optimally discriminate the channels. The intuition and proof of this result depend on the structure of maximally entangled states given in Chapter 5.

- In Chapter 8 we make progress on Question 2. We prove that the Werner-Holevo channel discrimination game is essentially the unique game with input dimension n satisfying the norm relation

$$1 = |||\lambda\Gamma_0 - (1-\lambda)\Gamma_1|||_1 = n\|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\|_1. \tag{4.24}$$

  Note that the above is the maximum possible gap between the two norms given that the input dimension of the channels is n. The above norm relation says that the game can be won with certainty using arbitrary entanglement, but that it is hard to win the game without entanglement, with the bound on performance without entanglement given by $\|\cdot\|_1$ being as small as possible given that the game can be won with certainty. The main part of proving this uniqueness is proving that transposition is essentially the unique linear map satisfying $\|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_1 = k\|\Psi\|_1$. The proof of this fact depends on the algebraic characterization of complete trace-norm isometries given in Chapter 6.

Chapter 5

The structure of maximal entanglement

In this chapter we consider the structure of *maximal entanglement*. For pure states, there is a natural notion of what maximal entanglement should be, given in Definition 2.1. For arbitrary mixed states, it is not immediately clear what a natural notion of maximal entanglement should be. One way to define "maximal entanglement" for mixed states is to consider entanglement measures, which are mathematical functions that attempt to quantify entanglement. Relative to an entanglement measure, we will say that a state is *maximally entangled* if it achieves the maximum possible value of this function.

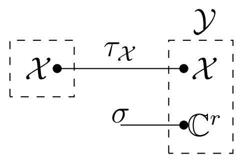
Here, we consider two settings. The first is the negativity as applied to *arbitrary* matrices. That is, we consider the function

$$\mathrm{L}(\mathbb{C}^n \otimes \mathbb{C}^m) \ni X \mapsto \|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(X)\|_1, \tag{5.1}$$

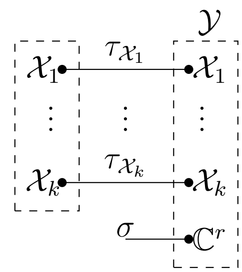
and characterize the set of (normalized) matrices achieving the maximal value of this function. This provides a notion, and characterization of, maximal entanglement for arbitrary matrices.

The second setting is a consideration of a broad class of functions on bipartite density matrices that we call *weak entanglement measures*. We similarly characterize the set of density matrices achieving the maximal value for these functions, and note that most well-known entanglement measures fall into this class.

In both contexts, we also extend the characterization of maximal entanglement to a multi-party setting, in which one system is maximally entangled with several others. The general structure proven in this setting is given in Figure 5.1 below.



(a) Structure of maximal entanglement



(b) Extension to multiple parties

Figure 5.1: Figure (a) illustrates the structure of maximal entanglement for two parties, $\mathcal{X}$ and $\mathcal{Y}$. When $\dim(\mathcal{X}) \leq \dim(\mathcal{Y})$, a state is maximally entangled if and only if $\mathcal{Y}$ roughly factorizes as $\mathcal{X} \otimes \mathbb{C}^r$, and the state looks like the canonical maximally entangled state $\tau_{\mathcal{X}}$ between the two $\mathcal{X}$ systems, plus some possibly noisy density matrix σ on the remainder $\mathbb{C}^r$.
Figure (b) illustrates the generalization to multiple parties. If a system $\mathcal{Y}$ is, on an individual basis, maximally entangled with the systems $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_k$, then $\mathcal{Y}$ roughly factorizes as $\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \otimes \mathbb{C}^r$, and the state looks like the canonical maximally entangled state between each pair of copies of $\mathcal{X}_i$, possible with some noisy density matrix ρ on $\mathbb{C}^r$ left over.

The results in this chapter are joint work with John Watrous, and appear in [50].

5.1 Matrices with maximal negativity

In Proposition 4.2 it was shown that for unit vectors $u, v \in \mathbb{C}^n \otimes \mathbb{C}^m$, it holds that

$$\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(uv^*)\|_1 = n \tag{5.2}$$

(the maximum possible value) if and only if u and v are both maximally entangled. Our first goal of this section is to generalize this to a characterization of the (not necessarily rank-1) matrices $X \in \mathrm{L}(\mathbb{C}^n \otimes \mathbb{C}^m)$ satisfying $\|X\|_1 = 1$ and $\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(X)\|_1 = n$.

For this we will first prove a technical fact about equality conditions in the trace-norm triangle inequality for a set of Hilbert-Schmidt orthogonal matrices.

The proof of this will require two facts. Firstly, for $A \in \mathrm{L}(\mathbb{C}^n)$, it holds that

$$\|A\|_1 = \max\{|\langle U, A\rangle| : U \in \mathrm{U}(\mathbb{C}^n)\}, \tag{5.3}$$

and the second is that $\mathrm{Tr}(A) = \|A\|_1$ if and only if $A \geq 0$.

Proposition 5.1. *Let $\{A_i\}_{i=1}^r \subset \mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$ be a Hilbert-Schmidt orthogonal set. It holds that*

$$\left\| \sum_{i=1}^r A_i \right\|_1 = \sum_{i=1}^r \|A_i\|_1, \tag{5.4}$$

if and only if $A_i A_j^* = 0$ and $A_i^* A_j = 0$ for all $i \neq j$.

Proof. First, assuming that $A_i A_j^* = 0$ and $A_i^* A_j = 0$ for $i \neq j$, Equation (5.4) may be verified directly using the definition of the trace-norm.

Next, assume that Equation (5.4) holds, and consider first the case that $n = m$, and $B, C \in \mathrm{L}(\mathbb{C}^n)$ are Hilbert-Schmidt orthogonal matrices for satisfying $\|B + C\|_1 = \|B\|_1 + \|C\|_1$. Let $U \in \mathrm{U}(\mathbb{C}^n)$ be a unitary satisfying

$$\langle U, B + C\rangle = \|B + C\|_1. \tag{5.5}$$

It follows that $\langle U, B\rangle = \|B\|_1$ and $\langle U, C\rangle = \|C\|_1$, and therefore $U^*B = B^*U$ and $U^*C = C^*U$ are both positive semidefinite matrices. We have

$$\langle B^*U, U^*C\rangle = \langle U^*B, C^*U\rangle = \langle B, C\rangle = 0, \tag{5.6}$$

and therefore $(B^*U)(U^*C) = 0$ and $(U^*B)(C^*U) = 0$, as Hilbert-Schmidt-orthogonal positive semidefinite matrices have product equal to zero. It follows that $B^*C = 0$ and $BC^* = 0$.

Now choose $i, j \in \{1, \ldots, r\}$ with $i \neq j$. The equality (5.4) implies that $\|A_i + A_j\|_1 = \|A_i\|_1 + \|A_j\|_1$. Defining $B, C \in \mathrm{L}(\mathbb{C}^n \oplus \mathbb{C}^m)$ as

$$B = \begin{pmatrix} 0 & 0 \\ A_i & 0 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 0 & 0 \\ A_j & 0 \end{pmatrix}, \tag{5.7}$$

we find that B and C are Hilbert-Schmidt orthogonal matrices satisfying

$$\|B + C\|_1 = \|B\|_1 + \|C\|_1, \tag{5.8}$$

and therefore $B^*C = 0$ and $BC^* = 0$ from the argument above. This implies that $A_i A_j^* = 0$ and $A_i^* A_j = 0$ as required. □

With this we may characterize the matrices with maximal entanglement negativity.

Theorem 5.2. *Let $X \in \mathrm{L}(\mathbb{C}^n \otimes \mathbb{C}^m)$ with $\|X\|_1 \leq 1$. The following are equivalent.*

1. $\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(X)\|_1 = n$.
2. *It holds that $m \geq n$, and there exists a positive integer $r \leq m/n$, a density matrix $\sigma \in \mathrm{D}(\mathbb{C}^r)$, and isometries $U, V \in \mathrm{U}(\mathbb{C}^n \otimes \mathbb{C}^r, \mathbb{C}^m)$ for which*

$$X = (\mathbb{1}_n \otimes U)(\tau_n \otimes \sigma)(\mathbb{1}_n \otimes V^*), \tag{5.9}$$

where $\tau_n \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^n)$ is the canonical maximally entangled state.

When $X \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^m)$ the above equivalence holds with $V = U$.

Proof. The fact that statement 2 implies statement 1 follows by a direct computation together with Proposition 4.2.

Now suppose that statement 1 holds, and observe that Proposition 4.2 immediately implies $m \geq n$. Let

$$X = \sum_{i=1}^{r} s_i x_i y_i^* \tag{5.10}$$

be a singular value decomposition of X, where $r = \mathrm{rank}(X)$. By Proposition 4.2 all of the x_i and y_i must be maximally entangled, as the triangle inequality would otherwise allow one to conclude that

$$\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(X)\|_1 < n. \tag{5.11}$$

Hence, for each i there exist isometries $A_i, B_i \in \mathrm{U}(\mathbb{C}^n, \mathbb{C}^m)$ for which

$$x_i = \frac{1}{\sqrt{n}}\mathrm{vec}(A_i^{\mathsf{T}}) \quad \text{and} \quad y_i = \frac{1}{\sqrt{n}}\mathrm{vec}(B_i^{\mathsf{T}}). \tag{5.12}$$

Now, note that

$$(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(X) = \frac{1}{n} W_{\mathbb{C}^n,\mathbb{C}^m} \sum_{i=1}^{r} s_i A_i \otimes B_i^*, \tag{5.13}$$

so that

$$n = \|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(X)\|_1 = \frac{1}{n}\Big\| \sum_{i=1}^{r} s_i A_i \otimes B_i^* \Big\|_1 \leq \frac{1}{n}\sum_{i=1}^{r} s_i \|A_i \otimes B_i^*\|_1 = n, \tag{5.14}$$

where the the last equality follows from the A_i and B_i being isometries, and therefore

$$\|A_i \otimes B_i^*\|_1 = n^2 \tag{5.15}$$

for every i. Hence, we have equality in the triangle inequality for these operators (which are Hilbert-Schmidt orthogonal as they arise from a singular value decomposition), and so Proposition 5.1 implies

$$(A_i \otimes B_i^*)^*(A_j \otimes B_j^*) = A_i^* A_j \otimes B_i B_j^* = 0, \tag{5.16}$$

$$(A_i \otimes B_i^*)(A_j \otimes B_j^*)^* = A_i A_j^* \otimes B_i^* B_j = 0, \tag{5.17}$$

for all $i \neq j$. As these are isometries, $B_i B_j^* \neq 0$, so the first expression above gives $A_i^* A_j = 0$, and likewise the second implies $B_i^* B_j = 0$ for all $i \neq j$. Hence the A_i (and respectively the B_i) embed $\mathbb{C}^n$ into r mutually orthogonal n-dimensional subspaces of $\mathbb{C}^m$, giving $rn \leq m$.

Lastly, to get the particular form of X, define $U, V \in \mathrm{U}(\mathbb{C}^n \otimes \mathbb{C}^r, \mathbb{C}^m)$ as

$$U = \sum_{i=1}^{r} A_i \otimes e_i^* \quad \text{and} \quad V = \sum_{i=1}^{r} B_i \otimes e_i^*, \tag{5.18}$$

where the fact that U and V are isometries follows from $A_i^* A_j = 0 = B_i^* B_j$ for $i \neq j$. Defining

$$\sigma = \sum_{i=1}^{r} s_i E_{ii} \in \mathrm{D}(\mathbb{C}^r), \tag{5.19}$$

we see that

$$X = \frac{1}{n}\sum_{i=1}^{r} s_i \mathrm{vec}(A_i^{\mathsf{T}}) \mathrm{vec}(B_i^{\mathsf{T}})^* \tag{5.20}$$

$$= (\mathbb{1}_n \otimes U)\Big(\sum_{i=1}^{r} \frac{s_i}{n} \mathrm{vec}(\mathbb{1}_n)\mathrm{vec}(\mathbb{1}_n)^* \otimes E_{ii}\Big)(\mathbb{1}_n \otimes V^*) \tag{5.21}$$

$$= (\mathbb{1}_n \otimes U)(\tau_n \otimes \sigma)(\mathbb{1}_n \otimes V^*), \tag{5.22}$$

as required.

When $X \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^m)$, in the above $B_i = A_i$, and hence $V = U$. $\square$

Remark 5.3. Later in this thesis we will make use of the additional special case of the above theorem when X is Hermitian. In this case the second statement may be rewritten as: $m \geq n$, and there exists a positive integer $r \leq m/n$, a Hermitian matrix $H \in \mathrm{Herm}(\mathbb{C}^r)$ with $\|H\|_1 = 1$, and an isometry $U \in \mathrm{U}(\mathbb{C}^n \otimes \mathbb{C}^r, \mathbb{C}^m)$ for which

$$X = (\mathbb{1}_n \otimes U)(\tau_n \otimes H)(\mathbb{1}_n \otimes U^*). \tag{5.23}$$

The only change necessary to the proof is to take a spectral decomposition of X, rather than a singular value decomposition. The rest of the proof follows as before.

5.1.1 Generalization to multiple parties

Next we generalize Theorem 5.2 to a multiparty setting. Before stating the theorem we introduce some notation. Given a k-fold tensor product of complex Euclidean spaces $\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k$, the *reduction to the i^{th} system* is denoted $R_i \in \mathrm{C}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k, \mathcal{X}_i)$, and acts as

$$R_i(X_1 \otimes \cdots \otimes X_k) = \Big(\prod_{j \neq i} \mathrm{Tr}(X_j)\Big) X_i \tag{5.24}$$

for all $X_1 \in \mathrm{L}(\mathcal{X}_1), \ldots, X_k \in \mathrm{L}(\mathcal{X}_k)$. We will also write the transpose on $\mathrm{L}(\mathcal{X})$ as $T_{\mathcal{X}}$. With this notation we may state the theorem of this section.

Theorem 5.4. *Let $\mathcal{X}_1 = \mathbb{C}^{n_1}, \ldots, \mathcal{X}_k = \mathbb{C}^{n_k}$, $\mathcal{Y} = \mathbb{C}^m$,*

$$N = \prod_{i=1}^{k} n_i = \dim(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k), \tag{5.25}$$

and let $X \in \mathrm{L}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \otimes \mathcal{Y})$ with $\|X\|_1 = 1$. The following are equivalent:

1. *$\|(T_{\mathcal{X}_i} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})((R_i \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(X))\|_1 = n_i$, for all $1 \leq i \leq k$.*
2. *$\|(T_{\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(X)\|_1 = N$.*
3. *It holds that $m \geq N$, and there exists a positive integer $r \leq m/N$, a density matrix $\sigma \in \mathrm{D}(\mathbb{C}^r)$, and isometries $U, V \in \mathrm{U}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \otimes \mathbb{C}^r, \mathcal{Y})$*

for which

$$X = (\mathbb{1}_{\mathcal{X}_1\otimes\cdots\otimes\mathcal{X}_k} \otimes U)(\tau_{\mathcal{X}_1\otimes\cdots\otimes\mathcal{X}_k} \otimes \sigma)(\mathbb{1}_{\mathcal{X}_1\otimes\cdots\otimes\mathcal{X}_k} \otimes V^*), \tag{5.26}$$

where $\tau_{\mathcal{X}_1\otimes\cdots\otimes\mathcal{X}_k} \in \mathrm{D}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \otimes \mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k)$ *is the canonical maximally entangled state.*

If $X \in \mathrm{D}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \otimes \mathcal{Y})$ *the above equivalence holds with* $V = U$*.*

Informally, the above theorem states that, when $\dim(\mathcal{Y}) \geq N$, a matrix $X \in \mathrm{L}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \otimes \mathcal{Y})$ is maximally entangled between the systems

$$\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \tag{5.27}$$

and $\mathcal{Y}$ if and only if each reduction of X, $R_{\mathcal{X}_i\otimes\mathcal{Y}}(X) \in \mathrm{L}(\mathcal{X}_i \otimes \mathcal{Y})$ is maximally entangled between $\mathcal{X}_i$ and $\mathcal{Y}$.

To prove the theorem we require two lemmas.

Lemma 5.5. *Let* $X \in \mathrm{L}(\mathcal{X} \otimes \mathcal{Y})$ *with* $\|X\|_1 = 1$*. If* $\mathrm{Tr}_{\mathcal{Y}}(X) = uv^*$ *for some unit vectors* $u, v \in \mathcal{X}$*, then there exists* $\sigma \in \mathrm{D}(\mathcal{Y})$ *for which* $X = uv^* \otimes \sigma$*.*

Proof. First consider the case in which X is positive semidefinite, and therefore a density matrix by the condition $\|X\|_1 = 1$. The partial trace is a positive map, from which it follows that $v = u$. Define a projection $\Pi = \mathbb{1}_{\mathcal{X}} - uu^*$, and observe that $\langle \Pi \otimes \mathbb{1}_{\mathcal{Y}}, X\rangle = \langle \Pi, \mathrm{Tr}_{\mathcal{Y}}(X)\rangle = 0$. As X and $\Pi \otimes \mathbb{1}_{\mathcal{Y}}$ are both positive semidefinite, it follows that $(\Pi \otimes \mathbb{1}_{\mathcal{Y}})X = X(\Pi \otimes \mathbb{1}_{\mathcal{Y}}) = 0$, and therefore

$$X = (uu^* \otimes \mathbb{1}_{\mathcal{Y}} + \Pi \otimes \mathbb{1}_{\mathcal{Y}})X(uu^* \otimes \mathbb{1}_{\mathcal{Y}} + \Pi \otimes \mathbb{1}_{\mathcal{Y}}) \tag{5.28}$$
$$= (uu^* \otimes \mathbb{1}_{\mathcal{Y}})X(uu^* \otimes \mathbb{1}_{\mathcal{Y}}) \tag{5.29}$$
$$= uu^* \otimes \sigma, \tag{5.30}$$

where $\sigma = (u^* \otimes \mathbb{1}_{\mathcal{Y}})X(u \otimes \mathbb{1}_{\mathcal{Y}}) \in \mathrm{D}(\mathcal{Y})$.

For the general case, let $U \in \mathrm{U}(\mathcal{Y})$ be a unitary satisfying $Uu = v$. It follows that

$$\|(U \otimes \mathbb{1}_{\mathcal{Y}})X\|_1 = 1 = \mathrm{Tr}((U \otimes \mathbb{1}_{\mathcal{Y}})X), \tag{5.31}$$

and therefore $(U\otimes\mathbb{1}_{\mathcal{Y}})X$ is positive semidefinite. Applying the positive semidefinite case to $(U\otimes\mathbb{1}_{\mathcal{Y}})X$ yields $(U\otimes\mathbb{1}_{\mathcal{Y}})X = vv^*\otimes\sigma$ for some choice of $\sigma \in \mathrm{D}(\mathcal{Y})$, and therefore $X = uv^* \otimes \sigma$, which completes the proof. □

Lemma 5.6. *Let $X \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$, and let $\Pi_1 \in \mathrm{L}(\mathcal{Y})$ and $\Pi_2 \in \mathrm{L}(\mathcal{X})$ be orthogonal projections. If*

$$\|\Pi_1 X \Pi_2\|_1 = \|X\|_1, \tag{5.32}$$

then it holds that $\Pi_1 X \Pi_2 = X$.

Proof. Let $X = \sum_{i=1}^{r} s_i u_i v_i^*$ be a singular value decomposition of X. Then, we have that

$$\begin{aligned}\sum_{i=1}^{r} s_i = \|X\|_1 = \|\Pi_1 X \Pi_2\|_1 &= \Big\| \sum_{i=1}^{r} s_i \Pi_1 u_i v_i^* \Pi_2 \Big\|_1 \\ &\leq \sum_{i=1}^{r} s_i \|\Pi_1 u_i v_i^* \Pi_2\|_1 \leq \sum_{i=1}^{r} s_i.\end{aligned} \tag{5.33}$$

Hence, all inequalities are equalities, which implies

$$1 = \|\Pi_1 u_i v_i^* \Pi_2\|_1 = \|\Pi_1 u_i\| \|\Pi_2 v_i\| \tag{5.34}$$

for all $1 \leq i \leq r$. It follows that $\Pi_1 u_i = u_i$ and $\Pi_2 v_i = v_i$ for all i, and hence $\Pi_1 X \Pi_2 = X$. □

Before proving the theorem we introduce an implicit permutation notation. At points in the proof we will be working with matrices that act on a tensor product space, where the ordering of the tensor factors for which it is convenient to specify the matrix is not the same as the ordering used in the context that the matrix appears. This primarily occurs for matrices of product form. For example, given $A \in \mathrm{L}(\mathcal{X} \otimes \mathcal{Z})$, and $B \in \mathrm{L}(\mathcal{Y})$, the matrix $A \otimes B \in \mathrm{L}(\mathcal{X} \otimes \mathcal{Z} \otimes \mathcal{Y})$ has a simple form, but if our spaces are naturally ordered as $\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z}$, then we must write

$$(\mathbb{1}_{\mathcal{X}} \otimes W_{\mathcal{Z},\mathcal{Y}})(A \otimes B)(\mathbb{1}_{\mathcal{X}} \otimes W_{\mathcal{Z},\mathcal{Y}}^*) \tag{5.35}$$

to specify it as a matrix in $\mathrm{L}(\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z})$, which can become clunky.

To avoid this, we introduce the following notation. For some finite list of complex Euclidean spaces $\mathcal{Z}_1, \ldots, \mathcal{Z}_k$, a permutation

$$\sigma : \{1, \ldots, k\} \to \{1, \ldots, k\}, \tag{5.36}$$

and a matrix $X \in \mathrm{L}(\mathcal{Z}_1 \otimes \cdots \otimes \mathcal{Z}_k)$, we write

$$\underbrace{X}_{\in \mathrm{L}(\mathcal{Z}_{\sigma(1)} \otimes \cdots \otimes \mathcal{Z}_{\sigma(k)})} = P X P^*, \tag{5.37}$$

where $P \in \mathrm{U}(\mathcal{Z}_1 \otimes \cdots \otimes \mathcal{Z}_k, \mathcal{Z}_{\sigma(1)} \otimes \cdots \otimes \mathcal{Z}_{\sigma(k)})$ is the isometry which permutes the subsystems as given in the definition. For the example in the preceding paragraph, this notation gives

$$\underbrace{A \otimes B}_{\in \mathrm{L}(\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z})} = (\mathbb{1}_{\mathcal{X}} \otimes W_{\mathcal{Z},\mathcal{Y}})(A \otimes B)(\mathbb{1}_{\mathcal{X}} \otimes W^*_{\mathcal{Z},\mathcal{Y}}). \tag{5.38}$$

Note as well that for complex Euclidean spaces $\mathcal{A}$ and $\mathcal{B}$, it holds that

$$\tau_{\mathcal{A} \otimes \mathcal{B}} = \underbrace{\tau_{\mathcal{A}} \otimes \tau_{\mathcal{B}}}_{\in \mathrm{L}(\mathcal{A} \otimes \mathcal{B} \otimes \mathcal{A} \otimes \mathcal{B})}. \tag{5.39}$$

In the above there is a potential ambiguity as multiple copies of the same space appear, so it is not necessarily well defined. In this case however, the matrix is invariant under swapping the order of these copies, and so there is no real ambiguity.

Proof of Theorem 5.4. The equivalence of statements 2 and 3 is the content of Theorem 5.2, and from this we also retrieve the statement that if X is a density matrix, then we can take $V = U$ in statement 3. That statement 3 implies statement 1 follows by a direct computation, along with the observation in Equation (5.39). When $k = 1$, statements 1 and 2 are the same, so in this case there is nothing to prove. When $k = 2$ we will show that statement 1 implies statement 3 (in which case we will have the full equivalence for $k = 2$), then use induction to directly show that statement 1 is equivalent to statement 2 for $k > 2$.

For statement 1 implies statement 3 in the $k = 2$ case, to simplify notation we denote $\mathcal{A} = \mathcal{X}_1$, $\mathcal{B} = \mathcal{X}_2$, $a = n_1$, and $b = n_2$, and hence $N = ab$. We will use Lemmas 5.5 and 5.6 to deduce the required form of X from the structure that Theorem 5.2 gives for the reductions $\mathrm{Tr}_{\mathcal{A}}(X)$ and $\mathrm{Tr}_{\mathcal{B}}(X)$. By Theorem 5.2 it follows from $\left\| \big(T_{\mathcal{A}} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(\mathrm{Tr}_{\mathcal{B}}(X)) \right\|_1 = a$ that $a \leq m$, and there exists $s \in \{1, \ldots, \lfloor m/a \rfloor\}$, $\nu \in \mathrm{D}(\mathbb{C}^s)$, and isometries $A, B \in \mathrm{U}(\mathcal{A} \otimes \mathbb{C}^s, \mathcal{Y})$ for which

$$\mathrm{Tr}_{\mathcal{B}}(X) = (\mathbb{1}_{\mathcal{A}} \otimes A)(\tau_{\mathcal{A}} \otimes \nu)(\mathbb{1}_{\mathcal{A}} \otimes B^*). \tag{5.40}$$

This implies that

$$\mathrm{Tr}_{\mathcal{B}\otimes\mathbb{C}^s}((\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes A^*)X(\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes B)) = \tau_{\mathcal{A}}. \tag{5.41}$$

Note that

$$1 = \|\tau_{\mathcal{A}}\|_1 = \|\mathrm{Tr}_{\mathcal{B}\otimes\mathbb{C}^s}((\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes A^*)X(\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes B))\|_1 \tag{5.42}$$

$$\leq \|(\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes A^*)X(\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes B)\|_1 \leq \|X\|_1 = 1, \tag{5.43}$$

giving $\|(\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes A^*)X(\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes B)\|_1 = 1$, and so Lemma 5.5 implies that there exists $\eta \in \mathrm{D}(\mathcal{B}\otimes\mathbb{C}^s)$ for which

$$(\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes A^*)X(\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes B) = \underbrace{\tau_{\mathcal{A}} \otimes \eta}_{\in \mathrm{L}(\mathcal{A}\otimes\mathcal{B}\otimes\mathcal{A}\otimes\mathbb{C}^s)}, \tag{5.44}$$

and hence

$$(\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes AA^*)X(\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes BB^*) = (\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes A)\underbrace{(\tau_{\mathcal{A}} \otimes \eta)}_{\in \mathrm{L}(\mathcal{A}\otimes\mathcal{B}\otimes\mathcal{A}\otimes\mathbb{C}^s)}(\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes B^*). \tag{5.45}$$

As the above matrix has trace norm 1, and $\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes AA^*$ and $\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes BB^*$ are both orthogonal projections, Lemma 5.6 implies

$$X = (\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes A)\underbrace{(\tau_{\mathcal{A}} \otimes \eta)}_{\in \mathrm{L}(\mathcal{A}\otimes\mathcal{B}\otimes\mathcal{A}\otimes\mathbb{C}^s)}(\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes B^*). \tag{5.46}$$

Next, it holds that

$$\big\|\big(T_{\mathcal{B}} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^s)}\big)(\eta)\big\|_1 = \big\|\big(T_{\mathcal{B}} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(\mathrm{Tr}_{\mathcal{A}}(X))\big\|_1 = b, \tag{5.47}$$

and so again by Theorem 5.2, $b \leq s$, and there exists $r \in \{1, \ldots, \lfloor s/b \rfloor\}$, $\sigma \in \mathrm{D}(\mathbb{C}^r)$, and an isometry $S \in \mathrm{U}(\mathcal{B}\otimes\mathbb{C}^r, \mathbb{C}^s)$ for which

$$\eta = (\mathbb{1}_{\mathcal{B}} \otimes S)(\tau_{\mathcal{B}} \otimes \sigma)(\mathbb{1}_{\mathcal{B}} \otimes S^*). \tag{5.48}$$

Hence, letting $U = A(\mathbb{1}_{\mathcal{A}} \otimes S)$ and $V = B(\mathbb{1}_{\mathcal{A}} \otimes S)$ we get that

$$X = (\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes A) \underbrace{[\tau_{\mathcal{A}} \otimes (\mathbb{1}_{\mathcal{B}} \otimes S)(\tau_{\mathcal{B}} \otimes \sigma)(\mathbb{1}_{\mathcal{B}} \otimes S^*)]}_{\in \mathrm{L}(\mathcal{A}\otimes\mathcal{B}\otimes\mathcal{A}\otimes\mathbb{C}^s)} (\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes B^*) \tag{5.49}$$

$$= (\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes U) \underbrace{(\tau_{\mathcal{A}} \otimes \tau_{\mathcal{B}} \otimes \sigma)}_{\in \mathrm{L}(\mathcal{A}\otimes\mathcal{B}\otimes\mathcal{A}\otimes\mathcal{B}\otimes\mathbb{C}^r)} (\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes V^*) \tag{5.50}$$

$$= (\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes U)(\tau_{\mathcal{A}\otimes\mathcal{B}} \otimes \sigma)(\mathbb{1}_{\mathcal{A}\otimes\mathcal{B}} \otimes V^*), \tag{5.51}$$

and $ab \leq as \leq m$, and $r \leq s/b \leq m/ab$, as required.

Lastly, we show that statement 1 is equivalent to statement 2 for all k by induction. So, assuming the equivalence holds for some $k \geq 2$, we show it holds for $k+1$. Note that

$$\left\| \left(T_{\mathcal{X}_i} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\right)\left(\left(R_i \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\right)(X)\right)\right\|_1 = n_i \tag{5.52}$$

for all $1 \leq i \leq k$, by the induction hypothesis, is equivalent to

$$\left\| \left(T_{\mathcal{X}_1\otimes\cdots\otimes\mathcal{X}_k} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\right)(\mathrm{Tr}_{\mathcal{X}_{k+1}}(X))\right\|_1 = \prod_{i=1}^{k} n_i, \tag{5.53}$$

which, together with $\left\| \left(T_{\mathcal{X}_{k+1}} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\right)\left(\left(R_{k+1} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\right)(X)\right)\right\|_1 = n_{k+1}$, again by the induction hypothesis, is equivalent to

$$\left\| \left(T_{\mathcal{X}_1\otimes\cdots\otimes\mathcal{X}_{k+1}} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\right)(X)\right\|_1 = \prod_{i=1}^{k+1} n_i, \tag{5.54}$$

as required. □

5.2 Weak entanglement measures

In the previous section we looked at the structure of bipartite matrices that maximize negativity, and proved a multipartite version of this result. In this section, we show that, when restricting attention to density matrices, the same results hold for a broad class of entanglement measures that we call *weak entanglement measures.*

Definition 5.7. A *weak entanglement measure* is a family of functions

$$\{\mathrm{E}_{n,m} : n, m \in \mathbb{N},\ 1 \leq n \leq m\}, \tag{5.55}$$

each of which takes the form

$$\mathrm{E}_{n,m} : \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^m) \to \mathbb{R}, \tag{5.56}$$

for which the following properties hold:

1. There exists a function $g : \mathbb{N} \to \mathbb{R}$ for which

$$\max_{\rho \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^m)} \mathrm{E}_{n,m}(\rho) = g(n). \tag{5.57}$$

   That is, we assume that the maximum exists and that it is a function only of the minimum of the two dimensions. We call g the *maximum function* for the family $\{\mathrm{E}_{n,m}\}$.

2. For any unit vector $u \in \mathrm{S}(\mathbb{C}^n \otimes \mathbb{C}^m)$, it holds that $\mathrm{E}_{n,m}(uu^*) = g(n)$ if and only if u is maximally entangled (in the sense given in Definition 2.1).

3. The measure is monotonically decreasing under quantum channels acting on the second subsystem. That is, for all density matrices $\rho \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^m)$ and channels $\Phi \in \mathrm{C}(\mathbb{C}^m, \mathbb{C}^k)$ for $k \geq n$, it holds that

$$\mathrm{E}_{n,k}((\mathbb{1}_{\mathrm{L}(\mathbb{C}^n)} \otimes \Phi)(\rho)) \leq \mathrm{E}_{n,m}(\rho). \tag{5.58}$$

4. Each function $\mathrm{E}_{n,m}$ is *pure state convex*: for any set

$$\{u_1, \ldots, u_N\} \subset \mathrm{S}(\mathbb{C}^n \otimes \mathbb{C}^m) \tag{5.59}$$

   and probability vector $(p_1, \ldots, p_N)$, it holds that

$$\mathrm{E}_{n,m}\left(\sum_{i=1}^{N} p_i u_i u_i^*\right) \leq \sum_{i=1}^{N} p_i \, \mathrm{E}_{n,m}\big(u_i u_i^*\big). \tag{5.60}$$

A few comments on this definition are in order. First, pure state convexity may seem an odd axiom (as opposed to general convexity), but there may exist entanglement measures that are pure state convex and not generally convex. (For example, distillable entanglement is known to be pure-state convex [15, Lemma 25], but may not be generally convex [59].) Second, it is generally desired that entanglement measures satisfy stronger versions of the third condition (e.g., monotonicity with respect to any LOCC channel between both subsystems). Furthermore entanglement measures usually treat the two subsystems symmetrically, and Property 3 is asymmetric in that it only applies to the second subsystem. In our proof the subsystems are treated asymmetrically, and we only need monotonicity to hold with respect to the second system (and hence this result can be applied to functions like the coherent information).

The set of weak entanglement measures includes negativity [66], coherent information [58], squashed entanglement [13, 65], entanglement of formation, and distillable entanglement. See [4, Table 1] for a list of commonly used entanglement measures and the properties that they are known to satisfy.

In order to prove the theorem that follows we will make use of the following simple lemma.

Lemma 5.8. *Let $U, V \in \mathrm{U}(\mathcal{X}, \mathcal{Y})$ be isometries. If it holds that:*

- *U and V are Hilbert-Schmidt orthogonal (i.e. $\langle U, V\rangle = 0$), and*
- *$\alpha U + \beta V$ is proportional to an isometry for all choices of $\alpha, \beta \in \mathbb{C}$,*

then $U^* V = 0$ (i.e., U and V map $\mathcal{X}$ into orthogonal subspaces of $\mathcal{Y}$).

Proof. It suffices to consider the pairs $(\alpha, \beta) = (1, 1)$ and $(\alpha, \beta) = (1, i)$. As $U + V$ and $U + iV$ are proportional to isometries, the following matrices must be proportional to the identity:

$$\big(U + V\big)^*\big(U + V\big) = 2\mathbb{1} + (U^*V + V^*U), \tag{5.61}$$

$$\big(U + iV\big)^*\big(U + iV\big) = 2\mathbb{1} + i(U^*V - V^*U). \tag{5.62}$$

Hence, both $U^*V + V^*U$ and $U^*V - V^*U$ must be proportional to the identity. As U^*V and V^*U are traceless, we conclude that

$$U^*V + V^*U = 0 \quad \text{and} \quad U^*V - V^*U = 0, \tag{5.63}$$

which implies $U^*V = 0$ as required. □

Theorem 5.9. *Let $n \le m$ be positive integers, and let $\rho \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^m)$. The following statements are equivalent:*

1. *For every weak entanglement measure $\{\mathrm{E}_{s,t}\}$ with maximum function g it holds that $\mathrm{E}_{n,m}(\rho) = g(n)$.*
2. *Statement 1 holds for any weak entanglement measure.*

3. *There exists a positive integer $r \leq m/n$, a density matrix $\sigma \in \mathrm{D}(\mathbb{C}^r)$, and an isometry $U \in \mathrm{U}(\mathbb{C}^n \otimes \mathbb{C}^r, \mathbb{C}^m)$ for which*

$$\rho = (\mathbb{1}_n \otimes U)(\tau_n \otimes \sigma)(\mathbb{1}_n \otimes U^*). \tag{5.64}$$

Proof. Statement 1 trivially implies statement 2 (as the set of weak entanglement measures is nonempty).

Now assume statement 2 holds: $\mathrm{E}_{n,m}(\rho) = g(n)$ for some weak entanglement measure $\{\mathrm{E}_{s,t}\}$ with maximum function g. By the pure-state convexity axiom (Property 4), for any pure-state decomposition

$$\rho = \sum_{i=1}^{N} p_i v_i v_i^* \tag{5.65}$$

(for $p_1, \ldots, p_N$ positive) it holds that

$$g(n) = \mathrm{E}_{n,m}(\rho) \leq \sum_{i=1}^{N} p_i \, \mathrm{E}_{n,m}(v_i v_i^*) \tag{5.66}$$

and $\mathrm{E}_{n,m}(v_i v_i^*) \leq g(n)$, implying that $\mathrm{E}_{n,m}(v_i v_i^*) = g(n)$, for all $i = 1, \ldots, N$. Hence, by Property 2, every pure state decomposition of ρ necessarily consists only of maximally entangled states. This is equivalent to the statement that every unit vector $v \in \mathrm{Im}(\rho)$ contained in the image of ρ is maximally entangled.

Now consider a spectral decomposition

$$\rho = \sum_{i=1}^{r} p_i v_i v_i^* \tag{5.67}$$

of ρ, where $r = \mathrm{rank}(\rho)$ and we have restricted the sum to range only over indices corresponding to positive eigenvalues of ρ. By the argument above, one has that each v_i is maximally entangled, so there exists an orthogonal collection of isometries $\{V_1, \ldots, V_r\} \subset \mathrm{U}(\mathcal{X}, \mathcal{Y})$ for which

$$v_i = \frac{1}{\sqrt{n}} \mathrm{vec}(V_i^{\mathsf{T}}) \tag{5.68}$$

for each $i \in \{1, \ldots, r\}$. For each pair $i \neq j$ we find that

$$\mathrm{vec}\big(\alpha V_i^{\mathsf{T}} + \beta V_j^{\mathsf{T}}\big) \in \mathrm{Im}(\rho), \tag{5.69}$$

and therefore $\alpha V_i + \beta V_j$ is proportional to an isometry for all $\alpha, \beta \in \mathbb{C}$. By Lemma 5.8 it holds that $V_i^* V_j = 0$, and hence $rn \leq m$.

Along the same lines as in Theorem 5.2, define $U \in \mathrm{U}(\mathcal{X} \otimes \mathbb{C}^r, \mathcal{Y})$ and $\sigma \in \mathrm{D}(\mathbb{C}^r)$ as

$$U = \sum_{i=1}^{r} V_i \otimes e_i^* \quad \text{and} \quad \sigma = \sum_{i=1}^{r} p_i E_{ii}, \tag{5.70}$$

where the fact that U is an isometry follows from $V_i^* V_j = 0$ for $i \neq j$. It follows by direct multiplication that

$$\rho = (\mathbb{1}_{\mathcal{X}} \otimes U)(\tau_{\mathcal{X}} \otimes \sigma)(\mathbb{1}_{\mathcal{X}} \otimes U)^*, \tag{5.71}$$

and therefore statement 2 implies statement 3.

Finally, assume that statement 3 holds, let $\{\mathrm{E}_{s,t}\}$ be any weak entanglement measure with maximum function g, and define a channel $\Phi \in \mathrm{C}(\mathcal{Y}, \mathcal{X})$ as follows:

$$\Phi(X) = \mathrm{Tr}_{\mathbb{C}^r}(U^* Y U) + \langle \mathbb{1}_{\mathcal{Y}} - UU^*, Y \rangle \eta, \tag{5.72}$$

for all $Y \in \mathrm{L}(\mathcal{Y})$ and any fixed choice of a density operator $\eta \in \mathrm{D}(\mathcal{X})$. It holds that $(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Phi)(\rho) = \tau_{\mathcal{X}}$, so by Property 3 one has

$$g(n) = \mathrm{E}_{n,n}(\tau_{\mathcal{X}}) = \mathrm{E}_{n,n}((\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Phi)(\rho)) \leq \mathrm{E}_{n,m}(\rho) \leq g(n). \tag{5.73}$$

It follows that $\mathrm{E}_{n,m}(\rho) = g(n)$, and so statement 3 implies statement 1. □

Using the above characterization we can arrive at a density matrix version of Theorem 5.4 that holds for any weak entanglement measure.

Corollary 5.10. *Let $\mathcal{X}_1 = \mathbb{C}^{n_1}, \ldots, \mathcal{X}_k = \mathbb{C}^{n_k}$ and $\mathcal{Y} = \mathbb{C}^m$ for positive integers $n_1, \ldots, n_k$ and m satisfying $n = \prod_{i=1}^k n_i \leq m$, let $\rho \in \mathrm{D}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \otimes \mathcal{Y})$ be a density matrix, and let $\{E_{s,t}\}$ be any weak entanglement measure with maximum function g. The following statements are equivalent:*

1. *It holds that*
$$\mathrm{E}_{n_i,m}((R_i \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(\rho)) = g(n_i) \tag{5.74}$$
for all $i = 1, \ldots, k$.

2. *It holds that*
$$\mathrm{E}_{n,m}(\rho) = g(n). \tag{5.75}$$

3. There exists a positive integer $r \leq n/m$, a density matrix $\sigma \in \mathrm{D}(\mathbb{C}^r)$, and an isometry

$$U \in \mathrm{U}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \otimes \mathbb{C}^r, \mathcal{Y}) \tag{5.76}$$

for which

$$\rho = (\mathbb{1}_{\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k} \otimes U)(\tau_{\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k} \otimes \sigma)(\mathbb{1}_{\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k} \otimes U^*). \tag{5.77}$$

Proof. The equivalence of the above statements was shown for the negativity in Theorem 5.4, and Theorem 5.9 gives that statements 1 and 2 hold for the negativity if and only if they hold for all weak entanglement measures. $\square$

Chapter 6

Complete trace-norm isometries and reversible quantum channels

The goal of this chapter is to provide various characterizations of complete trace-norm isometries and reversible quantum channels, which we define below.

Definition 6.1. Let $V \subset \mathrm{L}(\mathcal{X})$ be a subspace, and let $\Phi : V \to \mathrm{L}(\mathcal{Y})$ be a linear map. We say that:

- Φ is a *trace-norm isometry* if $\|\Phi(X)\|_1 = \|X\|_1$ for all $X \in V$.
- Φ is a *k-trace-norm isometry*, or that Φ is *k-trace-norm isometric*, if $\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}$ is a trace-norm isometry on $V \otimes \mathrm{L}(\mathbb{C}^k)$.
- Φ is a *complete trace-norm isometry* if it is a k-trace-norm isometry for all positive integers $k \geq 1$.

Definition 6.2. A quantum channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ is called *reversible* if there exists a quantum channel $\Psi \in \mathrm{C}(\mathcal{Y}, \mathcal{X})$ for which $\Psi\Phi = \mathbb{1}_{\mathrm{L}(\mathcal{X})}$.

Complete trace-norm isometries are linear maps that have no effect on the trace norm when applied to their input space, even when considering block matrices. A classic example of a trace-norm isometry that is not a complete trace-norm isometry is matrix transposition: For all $A \in \mathrm{L}(\mathbb{C}^n)$ it holds that $\|T_n(A)\|_1 = \|A\|_1$, but

$$\begin{aligned}\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(\mathrm{vec}(\mathbb{1}_n)\mathrm{vec}(\mathbb{1}_n)^*)\|_1 &= \|W_{\mathbb{C}^n,\mathbb{C}^n}\|_1 \\ &= n^2 > n = \|\mathrm{vec}(\mathbb{1}_n)\mathrm{vec}(\mathbb{1}_n)^*\|_1.\end{aligned} \tag{6.1}$$

Our main purpose for studying complete trace-norm isometries, and their positive counterpart reversible quantum channels, is that they appear naturally when characterizing linear maps whose multiplicity maps have maximal norm in Chapter 8.

In Section 6.1 we provide characterizations of complete trace-norm isometries, in Section 6.2 we give several characterizations of reversible quantum channels, and in Section 6.3 we give norm characterizations for when a channel admits error-correcting codes. The results of Section 6.1 appear in [49], and the results of Section 6.2 appear in [50].

6.1 Complete trace-norm isometries

The purpose of this section is to give various characterizations of complete trace-norm isometries in $\mathrm{T}(\mathbb{C}^n, \mathbb{C}^m)$. Note that the structure of surjective operator norm isometries (and hence surjective complete operator norm isometries) between C^*-algebras is well-known [33]. Furthermore, in the matrix algebra case, a characterization of (not necessarily surjective) operator norm isometries in $\mathrm{T}(\mathbb{C}^n, \mathbb{C}^m)$ has been given for the case $m \leq 2n-1$ [8]. However, the dual/adjoint of a trace-norm isometry need not be an operator norm isometry, and so it is not possible to import those results here.

We give the various characterizations in the theorem below. Remarks and some background on what is already known, as well as some intermediate results, are given before its proof.

Theorem 6.3. *For a linear map $\Phi \in \mathrm{T}(\mathbb{C}^n, \mathbb{C}^m)$ the following are equivalent.*

1. *Φ is a complete trace-norm isometry.*
2. *Φ is a 2-trace-norm isometry.*
3. *There exists $X \in \mathrm{L}(\mathbb{C}^n) \setminus \{0\}$ for which $\|\Phi(X)\|_1 = \|X\|_1$, and the following implications hold for $A, B, C, D \in \mathrm{L}(\mathbb{C}^n)$:*
   - *$A^*B = C^*D \implies \Phi(A)^*\Phi(B) = \Phi(C)^*\Phi(D)$, and*
   - *$AB^* = CD^* \implies \Phi(A)\Phi(B)^* = \Phi(C)\Phi(D)^*$.*
4. *There exists $X \in \mathrm{L}(\mathbb{C}^n) \setminus \{0\}$ for which $\|\Phi(X)\|_1 = \|X\|_1$, and the following implications hold for $A, B \in \mathrm{L}(\mathbb{C}^n)$:*

- *$A^*B = 0 \implies \Phi(A)^*\Phi(B) = 0$, and*
- *$AB^* = 0 \implies \Phi(A)\Phi(B)^* = 0$.*

5. *There exists $X \in \mathrm{L}(\mathbb{C}^n) \setminus \{0\}$ for which $\|\Phi(X)\|_1 = \|X\|_1$, and the following implications hold for rank-1 $A, B \in \mathrm{L}(\mathbb{C}^n)$:*
    - *$A^*B = 0$ and $A^*A = B^*B \implies \Phi(A)^*\Phi(B) = 0$, and*
    - *$AB^* = 0$ and $AA^* = BB^* \implies \Phi(A)\Phi(B)^* = 0$,*

6. *$\|J(\Phi)\|_1 = n$ and $\|J(\Phi T_n)\|_1 = n^2$.*

7. *It holds that $m \geq n$, and there exists a positive integer $r \leq m/n$, a density matrix $\sigma \in \mathrm{D}(\mathbb{C}^r)$, and isometries $U, V \in \mathrm{U}(\mathbb{C}^n \otimes \mathbb{C}^r, \mathbb{C}^m)$ for which*
$$\Phi(X) = U(X \otimes \sigma)V^* \tag{6.2}$$
for all $X \in \mathrm{L}(\mathbb{C}^n)$.

8. *$\|\!|\Phi|\!\|_1 = 1$, and Φ has a left inverse $\Psi \in \mathrm{T}(\mathbb{C}^m, \mathbb{C}^n)$ with $\|\!|\Psi|\!\|_1 = 1$.*

If, in addition, Φ is positive, then statement 7 holds with $V = U$, making Φ a quantum channel, and Ψ may also be taken to be a quantum channel in statement 8 (and hence, Φ is a reversible quantum channel).

Before continuing some comments on the theorem are in order. In statement 6, the norm $\|J(\Phi T_n)\|_1$ appears, but this specific location of the transpose is an arbitrary notational choice. Using the definition of the Choi matrix and properties of the transpose, it may be verified that

$$\begin{aligned}\|J(\Phi T_n)\|_1 &= \|J(T_m \Phi)\|_1 \\ &= \|(T_m \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(J(\Phi))\|_1 = \|(\mathbb{1}_{\mathrm{L}(\mathbb{C}^m)} \otimes T_n)(J(\Phi))\|_1\end{aligned} \tag{6.3}$$

for any linear map $\Phi \in \mathrm{T}(\mathbb{C}^n, \mathbb{C}^m)$. With this interpretation, the characterization given in statement 6 says that Φ is a complete trace-norm isometry if and only if its Choi matrix is maximally entangled (and has a particular normalization).

Statements 3, 4, and 5 concern the map Φ preserving certain kinds of multiplication. The intuition for these statements comes from the explicit structure given in statement 7. However, in our proof, we show how they follow directly from the assumptions of Φ being either a complete or 2-trace-norm isometry.

The benefit of these alternative proofs, which we give separately in Proposition 6.5 below, is that they work in more generality (i.e. in the proposition we only use that the domain is a subspace $V \subset \mathrm{L}(\mathbb{C}^n)$), and so may be of independent interest. We also note that statement 5 may seem oddly specific, but it is included for being specially suited for use in the results in Chapter 8.

Lastly, while we give a complete proof of the above theorem, several equivalences may be deduced from [40], whose title "Isometries for Ky Fan Norms between Matrix Spaces" is self-explanatory of its content. In particular, as a special case of the results therein, an explicit structural characterization of (not necessarily complete) trace-norm isometries in $\mathrm{T}(\mathbb{C}^n, \mathbb{C}^m)$ is given. From this, the explicit structure of complete trace-norm isometries may be deduced by refining this structure, and indeed, this refinement only requires the additional assumption that the map is a 2-trace-norm isometry. Thus, the equivalence of statements 1, 2, and 7 may be viewed as a special case of the main theorem in [40]. Furthermore, the general technique of the proofs we give are in line with those of [40], and with linear norm preserver problems more generally [39]: translating between norm relations and algebraic relations for matrices. (See [7] for a survey of results on isometries of matrix spaces for unitarily invariant norms.)

With this last comment, we begin the proof of Theorem 6.3 with the following equivalence between a trace-norm relation for a 2×2 block-matrix, and statements about how the blocks multiply.

Proposition 6.4. *For matrices $A, B, C, D \in \mathrm{L}(\mathcal{X})$, it holds that*

$$\left\| \begin{pmatrix} A & B \\ C & D \end{pmatrix} \right\|_1 = \|A\|_1 + \|B\|_1 + \|C\|_1 + \|D\|_1, \tag{6.4}$$

if and only if

$$A^*B = AC^* = D^*C = DB^* = 0. \tag{6.5}$$

Proof. Apply Proposition 5.1 to the set

$$\{A \otimes E_{1,1}, B \otimes E_{1,2}, C \otimes E_{2,1}, D \otimes E_{2,2}\} \subset \mathrm{L}(\mathcal{X} \otimes \mathbb{C}^2). \tag{6.6}$$

□

Next, we prove a proposition containing some of the implications required for Theorem 6.3, but in more generality. We note that the proof takes inspiration from multiplicative domain proofs for unital and completely positive linear maps on C*-algebras (see [11] and [46, Theorem 3.18]).

Proposition 6.5. *Let $V \subset \mathrm{L}(\mathcal{X})$ be a subspace, and let $\Phi : V \to \mathrm{L}(\mathcal{Y})$ be linear.*

1. *If Φ is a 2-trace-norm isometry, then for $A, B \in V$ the following implications hold:*
   - $A^*B = 0 \implies \Phi(A)^*\Phi(B) = 0$, *and*
   - $AB^* = 0 \implies \Phi(A)\Phi(B)^* = 0$.
2. *If Φ is a complete trace-norm isometry, then for $A, B, C, D \in V$ the following implications hold:*
   - $A^*B = C^*D \implies \Phi(A)^*\Phi(B) = \Phi(C)^*\Phi(D)$, *and*
   - $AB^* = CD^* \implies \Phi(A)\Phi(B)^* = \Phi(C)\Phi(D)^*$.

Proof. First, assume Φ is a 2-trace-norm isometry and let $A, B \in V$. Assuming $A^*B = 0$, we have

$$\begin{aligned}\left\| \begin{pmatrix} \Phi(A) & \Phi(B) \\ 0 & 0 \end{pmatrix} \right\|_1 &= \left\| \begin{pmatrix} A & B \\ 0 & 0 \end{pmatrix} \right\|_1 \\ &= \|A\|_1 + \|B\|_1 = \|\Phi(A)\|_1 + \|\Phi(B)\|_1,\end{aligned} \tag{6.7}$$

where the second equality is by Proposition 6.4. Hence, also by Proposition 6.4, equality between the first and last expressions implies that $\Phi(A)^*\Phi(B) = 0$. Similarly, if $AB^* = 0$, then

$$\begin{aligned}\left\| \begin{pmatrix} \Phi(A) & 0 \\ \Phi(B) & 0 \end{pmatrix} \right\|_1 &= \left\| \begin{pmatrix} A & 0 \\ B & 0 \end{pmatrix} \right\|_1 \\ &= \|A\|_1 + \|B\|_1 = \|\Phi(A)\|_1 + \|\Phi(B)\|_1,\end{aligned} \tag{6.8}$$

and so $\Phi(A)\Phi(B)^* = 0$.

Next, assume Φ is a complete trace-norm isometry, and let $A, B, C, D \in V$. If $A^*B = C^*D$, then

$$\begin{pmatrix} A & 0 \\ -C & 0 \end{pmatrix}^* \begin{pmatrix} B & 0 \\ D & 0 \end{pmatrix} = 0. \tag{6.9}$$

Under the assumption that Φ is completely trace-norm isometric, $\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^2)}$ is a 2-trace-norm isometry, and so, by the 2-trace-norm isometry case, it holds that

$$\begin{pmatrix} \Phi(A) & 0 \\ -\Phi(C) & 0 \end{pmatrix}^* \begin{pmatrix} \Phi(B) & 0 \\ \Phi(D) & 0 \end{pmatrix} = 0, \tag{6.10}$$

giving $\Phi(A)^*\Phi(B) = \Phi(C)^*\Phi(D)$. Similarly, if $AB^* = CD^*$, then

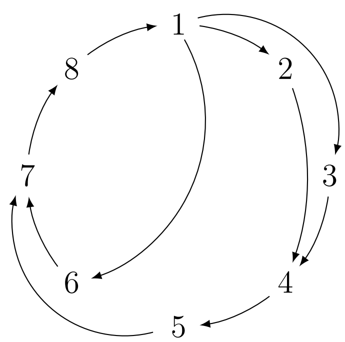
$$\begin{pmatrix} A & -C \\ 0 & 0 \end{pmatrix} \begin{pmatrix} B & D \\ 0 & 0 \end{pmatrix}^* = 0, \tag{6.11}$$

and again the 2-trace-norm isometry case implies that

$$\begin{pmatrix} \Phi(A) & -\Phi(C) \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Phi(B) & \Phi(D) \\ 0 & 0 \end{pmatrix}^* = 0, \tag{6.12}$$

giving $\Phi(A)\Phi(B)^* = \Phi(C)\Phi(D)^*$. □

Proof of Theorem 6.3. We prove the implications in the diagram below.



The implications appear in the order: 1 ⇒ 2, 3 ⇒ 4 ⇒ 5, 1 ⇒ 6, 2 ⇒ 4, 1 ⇒ 3, 8 ⇒ 1, 6 ⇒ 7, 7 ⇒ 8, and 5 ⇒ 7. All implications except 5 ⇒ 7, which is technically involved, follow essentially immediately from facts already given. The modified statements for the special case when Φ is positive are given before the proof of 5 ⇒ 7.

The implications that are immediate due to subsequent statements being logically weaker are 1 ⇒ 2 and 3 ⇒ 4 ⇒ 5. The implication 1 ⇒ 6 follows from the norm relations

$$\|J(\mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})\|_1 = n \text{ and } \|J(T_n)\|_1 = n^2, \tag{6.13}$$

and Φ being a complete trace-norm isometry. The implications 2 ⇒ 4 and 1 ⇒ 3 are both the content of Proposition 6.5, and the implication 8 ⇒ 1 is straightforward to verify.

For 6 ⇒ 7, the norm values of statement 6 imply by Theorem 5.2 that there exists a positive integer $r \leq m/n$, a density matrix $\sigma \in \mathrm{D}(\mathbb{C}^r)$, and isometries

$U, V \in \mathrm{U}(\mathbb{C}^r \otimes \mathbb{C}^n, \mathbb{C}^m)$ for which

$$\frac{1}{n} J(\Phi) = (U \otimes \mathbb{1}_n)(\sigma \otimes \tau_n)(V^* \otimes \mathbb{1}_n). \tag{6.14}$$

This is equivalent to the required form for Φ.

For 7 $\Rightarrow$ 8, define $\Psi(Y) = \mathrm{Tr}_{\mathbb{C}^r}(U^* Y V)$ for $Y \in \mathrm{L}(\mathbb{C}^m)$. Using the fact that the trace-norm is non-increasing under partial trace, it may be verified that Ψ has the required properties.

For the special case when Φ is positive, return to the proof of the implication 6 $\Rightarrow$ 7. As Φ is Hermiticity preserving, by Remark 5.3 and Theorem 5.2, there exists a Hermitian $H \in \mathrm{Herm}(\mathbb{C}^r)$ with $\|H\|_1 = 1$, and an isometry $U \in \mathrm{U}(\mathbb{C}^n \otimes \mathbb{C}^r, \mathbb{C}^m)$ for which $\Phi(X) = U(X \otimes H)U^*$ for all $X \in \mathrm{L}(\mathbb{C}^n)$. That Φ is positive implies $H \geq 0$, making H a density matrix, and giving Φ the required form. To see that Ψ may also be taken to be a quantum channel in statement 8, we define Ψ as as in the proof of 7 $\Rightarrow$ 8 with a slight modification. Fix a density matrix $\eta \in \mathrm{D}(\mathbb{C}^n)$, and set $\Psi(Y) = \mathrm{Tr}_{\mathbb{C}^r}(U^* Y U) + \mathrm{Tr}((\mathbb{1}_m - UU^*)Y)\eta$ for all $Y \in \mathrm{L}(\mathbb{C}^m)$. It is routine to verify that Ψ is a quantum channel and that $\Psi\Phi = \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)}$.

Lastly, we show 5 $\Rightarrow$ 7. We will use the assumption in statement 5 to build further facts about how outputs of Φ on rank-1 matrices multiply, which we break into a series of claims.

Claim 1. For unit vectors $x_1, x_2, y \in \mathbb{C}^n$ with $\langle x_1, x_2 \rangle = 0$, it holds that

$$\Phi(x_1 y^*)^* \Phi(x_1 y^*) = \Phi(x_2 y^*)^* \Phi(x_2 y^*) \tag{6.15}$$

and

$$\Phi(y x_1^*) \Phi(y x_1^*)^* = \Phi(y x_2^*) \Phi(y x_2^*)^*. \tag{6.16}$$

To see the first equality, note that $x_1 + x_2 \perp x_1 - x_2$, and so

$$0 = \Phi((x_1 + x_2)y^*)^* \Phi((x_1 - x_2)y^*) \tag{6.17}$$

$$= \Phi(x_1 y^*)^* \Phi(x_1 y^*) - \Phi(x_1 y^*)^* \Phi(x_2 y^*) + \Phi(x_2 y^*)^* \Phi(x_1 y^*) - \Phi(x_2 y^*)^* \Phi(x_2 y^*) \tag{6.18}$$

$$= \Phi(x_1 y^*)^* \Phi(x_1 y^*) - \Phi(x_2 y^*)^* \Phi(x_2 y^*), \tag{6.19}$$

where the second and third term in Equation (6.18) are 0 by application of statement 5. This gives the desired equality, and it follows similarly that $\Phi(y x_1^*) \Phi(y x_1^*)^* = \Phi(y x_2^*) \Phi(y x_2^*)^*$.

Claim 2. For any $x_1, x_2, y_1, y_2 \in \mathbb{C}^n$ with $\langle x_1, x_2 \rangle = \langle y_1, y_2 \rangle = 0$, it holds that

$$\Phi(x_1 y_1^*)^* \Phi(x_2 y_2^*) = 0 \text{ and } \Phi(x_1 y_1^*) \Phi(x_2 y_2^*)^* = 0. \tag{6.20}$$

For the first equality, assuming without loss of generality that x_1, x_2, y_1, and y_2 have unit length, Claim 1 gives that $\Phi(x_1 y_1^*)\Phi(x_1 y_1^*)^* = \Phi(x_1 y_2^*)\Phi(x_1 y_2^*)^*$, and statement 5 gives that $\Phi(x_1 y_2^*)^* \Phi(x_2 y_2^*) = 0$. These equalities imply the range relations

$$\mathrm{ran}(\Phi(x_1 y_1^*)) = \mathrm{ran}(\Phi(x_1 y_2^*)) \perp \mathrm{ran}(\Phi(x_2 y_2^*)), \tag{6.21}$$

giving $\Phi(x_1 y_1^*)^* \Phi(x_2 y_2^*) = 0$. It similarly holds that $\Phi(x_1 y_1^*) \Phi(x_2 y_2^*)^* = 0$.

Claim 3. For unit vectors $x_1, x_2, y_1, y_2 \in \mathbb{C}^n$ with $\langle x_1, x_2 \rangle = \langle y_1, y_2 \rangle = 0$, it holds that

$$\Phi(x_1 y_1^*)^* \Phi(x_1 y_2^*) = \Phi(x_2 y_1^*)^* \Phi(x_2 y_2^*) \tag{6.22}$$

and

$$\Phi(x_1 y_1^*) \Phi(x_2 y_1^*)^* = \Phi(x_1 y_2^*) \Phi(x_2 y_2^*)^*. \tag{6.23}$$

For the first equality, $x_1 + x_2 \perp x_1 - x_2$, so Claim 2 implies

$$\begin{aligned} 0 &= \Phi((x_1 + x_2) y_1^*)^* \Phi((x_1 - x_2) y_2^*) && (6.24)\\ &= \Phi(x_1 y_1^*)^* \Phi(x_1 y_2^*) - \Phi(x_1 y_1^*)^* \Phi(x_2 y_2^*) \\ &\qquad + \Phi(x_2 y_1^*)^* \Phi(x_1 y_2^*) - \Phi(x_2 y_1^*)^* \Phi(x_2 y_2^*) && (6.25)\\ &= \Phi(x_1 y_1^*)^* \Phi(x_1 y_2^*) - \Phi(x_2 y_1^*)^* \Phi(x_2 y_2^*), && (6.26) \end{aligned}$$

where we have used Claim 2 to determine that

$$\Phi(x_1 y_1^*)^* \Phi(x_2 y_2^*) = \Phi(x_2 y_1^*)^* \Phi(x_1 y_2^*) = 0. \tag{6.27}$$

It may be similarly reasoned that $\Phi(x_1 y_1^*) \Phi(x_2 y_1^*)^* = \Phi(x_1 y_2^*) \Phi(x_2 y_2^*)^*$.

We now use Claims 1 to 3 to construct the explicit structure of Φ by examining how it acts on elementary matrices. By Claim 1 it holds that

$$\Phi(E_{1,1})^* \Phi(E_{1,1}) = \Phi(E_{a,1})^* \Phi(E_{a,1}) \tag{6.28}$$

and

$$\Phi(E_{1,1}) \Phi(E_{1,1})^* = \Phi(E_{1,b}) \Phi(E_{1,b})^*, \tag{6.29}$$

for all $1 \leq a, b \leq n$, and hence there exists partial isometries

$$U_a : \mathrm{ran}(\Phi(E_{1,1})) \to \mathrm{ran}(\Phi(E_{a,1})) \tag{6.30}$$

and

$$V_b : \mathrm{ran}(\Phi(E_{1,1})^*) \to \mathrm{ran}(\Phi(E_{1,b})^*), \tag{6.31}$$

for which $\Phi(E_{a,1}) = U_a\Phi(E_{1,1})$ and $\Phi(E_{1,b}) = \Phi(E_{1,1})V_b^*$ (where U_1 and V_1 may be taken to be the orthogonal projections onto $\mathrm{ran}(\Phi(E_{1,1}))$ and $\mathrm{ran}(\Phi(E_{1,1})^*)$ respectively). Note as well that, statement 5 gives that

$$\Phi(E_{a,1})^*\Phi(E_{a',1}) = \Phi(E_{1,b})\Phi(E_{1,b'})^* = 0 \tag{6.32}$$

for $a \neq a'$ and $b \neq b'$, and hence the sets of partial isometries $\{U_a\}_{a=1}^n, \{V_b\}_{b=1}^n$ have mutually orthogonal ranges.

Next, we claim that for all $1 \leq a, b \leq n$ it holds that $\Phi(E_{a,b}) = U_a\Phi(E_{1,1})V_b^*$. In the previous claim this fact is established when at least one of a and b is 1, so we may assume that both $a, b \geq 2$. In this case, Claim 3 implies that

$$\Phi(E_{1,1})^*\Phi(E_{1,b}) = \Phi(E_{a,1})^*\Phi(E_{a,b}), \tag{6.33}$$

and so

$$\Phi(E_{1,1})^*\Phi(E_{1,1})V_b^* = \Phi(E_{1,1})^*U_a^*\Phi(E_{a,b}). \tag{6.34}$$

As $\mathrm{ran}(U_a^*) = \mathrm{ran}(\Phi(E_{1,1}))$ we may cancel the $\Phi(E_{1,1})^*$ from the left-side of the above equation to get that $\Phi(E_{1,1})V_b^* = U_a^*\Phi(E_{a,b})$ (alternatively, we may multiply on the left by the pseudo-inverse of $\Phi(E_{1,1})^*$). Finally, we have $\mathrm{ran}(U_a) = \mathrm{ran}(\Phi(E_{a,b}))$, as $\Phi(E_{a,b})^*\Phi(E_{a,b}) = \Phi(E_{a,1})^*\Phi(E_{a,1})$, and so

$$\Phi(E_{a,b}) = U_aU_a^*\Phi(E_{a,b}) = U_a\Phi(E_{1,1})V_b^*, \tag{6.35}$$

as required.

The last step is to show that the structure we have just deduced for Φ is the same as that in statement 7. Let $\Phi(E_{1,1}) = \sum_{i=1}^r s_i x_i y_i^*$ be a singular value decomposition. Define $\sigma = \sum_{i=1}^r s_i E_{i,i} \in \mathrm{L}(\mathbb{C}^r)$, which is clearly positive, and define matrices $U, V \in \mathrm{L}(\mathbb{C}^n \otimes \mathbb{C}^r, \mathbb{C}^m)$ to act as

$$U(e_a \otimes e_i) = U_a x_i, \text{ and } V(e_b \otimes e_j) = V_b y_j. \tag{6.36}$$

We may verify that these are in fact isometries:

$$\langle U(e_b \otimes e_j), U(e_a \otimes e_i)\rangle = \langle U_b x_j, U_a x_i\rangle = \delta_{a,b}\langle x_j, x_i\rangle = \delta_{a,b}\delta_{i,j}, \tag{6.37}$$

where we have used that U_a and U_b have orthogonal ranges for $a \neq b$. Hence, U is an isometry as it sends an orthonormal basis to an orthonormal set. The same proof shows that V is an isometry. Finally, we have that

$$\begin{aligned}\Phi(E_{a,b}) = U_a\Phi(E_{1,1})V_b^* &= \sum_{i=1}^{r} s_i U_a x_i y_i^* V_b^* \\ &= \sum_{i=1}^{r} s_i U(E_{a,b}\otimes E_{i,i})V^* = U(E_{a,b}\otimes\sigma)V^*.\end{aligned} \tag{6.38}$$

Thus, Φ has the desired form, and the last thing we need is that $\mathrm{Tr}(\sigma) = 1$. The final assumption is the existence of a non-zero $X \in \mathrm{L}(\mathbb{C}^n)$ satisfying $\|\Phi(X)\|_1 = \|X\|_1$. This gives

$$\|X\|_1 = \|\Phi(X)\|_1 = \|U(X\otimes\sigma)V^*\|_1 = \|X\|_1\mathrm{Tr}(\sigma), \tag{6.39}$$

and hence $\mathrm{Tr}(\sigma) = 1$ as desired. □

Remark 6.6. Consider an additional special case of Theorem 6.3 when Φ is Hermiticity preserving. As in the proof of the case when Φ is positive, there exists a positive integer $r \leq m/n$, a Hermitian $H \in \mathrm{Herm}(\mathbb{C}^r)$, and an isometry $U \in \mathrm{U}(\mathbb{C}^n\otimes\mathbb{C}^r,\mathbb{C}^m)$ for which $\Phi(X) = U(X\otimes H)U^*$ for all $X \in \mathrm{L}(\mathbb{C}^n)$. If $m < 2n$, then necessarily $r = 1$ and hence $H = \pm 1$. It follows that either Φ or $-\Phi$ is a reversible quantum channel. If $m \geq 2n$, then by considering the Hahn decomposition[1] of H, one may verify that this form is equivalent to the statement that there exists reversible quantum channels $\Phi_0, \Phi_1 \in \mathrm{C}(\mathbb{C}^n,\mathbb{C}^m)$ with orthogonal ranges and a number $r \in [0,1]$ for which

$$\Phi = r\Phi_0 - (1-r)\Phi_1. \tag{6.40}$$

6.2 Weak entanglement measures and reversible quantum channels

Next, we give various characterizations of reversible quantum channels. We apply Theorem 5.9 to show that a channel is reversible if and only if it preserves entanglement as measured by any weak entanglement measure. The structure given in Theorem 5.9 also allows us to re-derive a result from [45], where it

[1]The Hahn decomposition of a Hermitian matrix $H \in \mathrm{Herm}(\mathcal{X})$ is the unique decomposition of H as a difference $H = P - Q$ with $P, Q \geq 0$ and $PQ = 0$. For H Hermitian and $P, Q \geq 0$, it holds that $H = P - Q$ is the Hahn decomposition of H if and only if $\|H\|_1 = \|P\|_1 + \|Q\|_1$.

was shown that a channel is reversible if and only if it has a certain form. We also add in a couple of other conditions.

The statement of the theorem requires a couple of concepts from quantum information. First, for positive semidefinite matrices $P, Q \in \mathrm{Pos}(\mathcal{X})$, the *fidelity* is defined as

$$\mathrm{F}(P,Q) = \left\| \sqrt{P}\sqrt{Q} \right\|_1. \tag{6.41}$$

Second, for any pair of channels $\Phi \in \mathrm{C}(\mathcal{X},\mathcal{Y})$ and $\Psi \in \mathrm{C}(\mathcal{X},\mathcal{Z})$, it is said that Φ and Ψ are *complementary* if there exists an isometry $A \in \mathrm{U}(\mathcal{X}, \mathcal{Y}\otimes\mathcal{Z})$ such that

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}(AXA^*) \quad \text{and} \quad \Psi(X) = \mathrm{Tr}_{\mathcal{Y}}(AXA^*) \tag{6.42}$$

for all $X \in \mathrm{L}(\mathcal{X})$. We will also make use of a couple of simple facts, stated as lemmas as follows. (See, for instance, Corollary 3.23 and Proposition 2.29 in [69].)

Lemma 6.7. *For any $u, v \in \mathcal{X}\otimes\mathcal{Y}$ it holds that*

$$\mathrm{F}(\mathrm{Tr}_{\mathcal{Y}}(uu^*), \mathrm{Tr}_{\mathcal{Y}}(vv^*)) = \|\mathrm{Tr}_{\mathcal{X}}(uv^*)\|_1. \tag{6.43}$$

Lemma 6.8. *For $u \in \mathcal{X}\otimes\mathcal{Y}$ and $P \in \mathrm{Pos}(\mathcal{X}\otimes\mathcal{Z})$, if $\mathrm{Tr}_{\mathcal{Y}}(uu^*) = \mathrm{Tr}_{\mathcal{Z}}(P)$, then there exists a channel $\Psi \in \mathrm{C}(\mathcal{Y},\mathcal{Z})$ for which $(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Psi)(uu^*) = P$.*

Theorem 6.9. *Let $\mathcal{X} = \mathbb{C}^n$ and $\mathcal{Y} = \mathbb{C}^m$ for positive integers $n \le m$, let $\Phi \in \mathrm{C}(\mathcal{X},\mathcal{Y})$ be a channel, and let $\{\mathrm{E}_{s,t}\}$ be any weak entanglement measure with maximum function g. The following statements are equivalent:*

1. *Φ is reversible.*
2. *Φ preserves entanglement with respect to $\{\mathrm{E}_{s,t}\}$, meaning that for all positive integers $k \le n$ and all density matrices $\rho \in \mathrm{D}(\mathbb{C}^k \otimes \mathcal{X})$ it holds that*
$$\mathrm{E}_{k,m}\big((\mathbb{1}_{\mathrm{L}(\mathbb{C}^k)} \otimes \Phi)(\rho)\big) = \mathrm{E}_{k,n}(\rho). \tag{6.44}$$
3. *It holds that*
$$\mathrm{E}_{n,m}\big(\tfrac{1}{n}J(\Phi)\big) = g(n). \tag{6.45}$$
4. *There exists a positive integer $r \le m/n$, a density matrix $\sigma \in \mathrm{D}(\mathbb{C}^r)$, and an isometry $U \in \mathrm{U}(\mathcal{X}\otimes\mathbb{C}^r, \mathcal{Y})$ for which*
$$\Phi(X) = U(X \otimes \sigma)U^* \tag{6.46}$$
for all $X \in \mathrm{L}(\mathcal{X})$.

5. *It holds that*

$$\|\Phi(X)\|_1 = \|X\|_1 \tag{6.47}$$

for all $X \in \mathrm{L}(\mathcal{X})$*.*

6. *It holds that*

$$\mathrm{F}(\Phi(\rho), \Phi(\sigma)) = \mathrm{F}(\rho, \sigma) \tag{6.48}$$

for all $\rho, \sigma \in \mathrm{D}(\mathcal{X})$*.*

7. *If* $\Psi \in \mathrm{C}(\mathcal{X}, \mathcal{Z})$ *is complementary to* Φ*, then there exists a density operator* $\sigma \in \mathrm{D}(\mathcal{Z})$ *for which*

$$\Psi(X) = \mathrm{Tr}(X)\sigma \tag{6.49}$$

for all $X \in \mathrm{L}(\mathcal{X})$ *(i.e., all channels which are complementary to* Φ *are constant on* $\mathrm{D}(\mathcal{X})$*).*

Remark 6.10. We note that the equivalence of statements 1 and 4 is the content of [45, Theorem 2.1]. In the proof given therein, this equivalence follows from an argument similar to a key step of the proof of Theorem 5.9 (as well as Theorem 5.2). A similar argument has also been used to derive conditions under which an error map is correctable [36]. The equivalence of statements 4 and 6 follows from [42] for $\mathcal{Y} = \mathcal{X}$, but also for infinite dimensions. Similarly, the equivalence of statements 4 and 5 in infinite dimensions follows from [5]. Lastly, for the case of the coherent information, the equivalence of statements 1 and 3 is a special case of the result in [58, Section VI], in which it was shown that a channel is reversible on half of a bipartite pure state if and only if the data processing inequality is satisfied with equality.

Proof of Theorem 6.9. Assume that statement 1 holds, and let $\Psi \in \mathrm{C}(\mathcal{Y}, \mathcal{X})$ be a left-inverse of Φ. By the monotonicity of weak entanglement measures it holds that

$$\mathrm{E}_{k,n}(\rho) = \mathrm{E}_{k,n}\big((\mathbb{1}_{\mathrm{L}(\mathbb{C}^k)} \otimes \Psi\Phi)(\rho)\big) \leq \mathrm{E}_{k,m}\big((\mathbb{1}_{\mathrm{L}(\mathbb{C}^k)} \otimes \Phi)(\rho)\big) \leq \mathrm{E}_{k,n}(\rho) \tag{6.50}$$

for all choices of $k \leq n$ and $\rho \in \mathrm{D}(\mathbb{C}^k \otimes \mathcal{X})$. Hence, statement 1 implies statement 2.

Statement 2 immediately implies statement 3, as statement 3 is equivalent to the particular choice of $k = n$ and $\rho = \tau_{\mathcal{X}}$ in statement 2.

Next, under the assumption that statement 3 holds, one has that the Choi matrix of Φ is given by

$$J(\Phi) = (\mathbb{1}_{\mathcal{X}} \otimes U)(\operatorname{vec}(\mathbb{1}_{\mathcal{X}})\operatorname{vec}(\mathbb{1}_{\mathcal{X}})^* \otimes \sigma)(\mathbb{1}_{\mathcal{X}} \otimes U^*), \tag{6.51}$$

by Theorem 5.9. This is equivalent to

$$\Phi(X) = U(X \otimes \sigma)U^* \tag{6.52}$$

for all $X \in \mathrm{L}(\mathcal{X})$. It has therefore been proved that statement 3 implies statement 4.

By well-known properties of the trace norm and the fidelity function, one immediately finds that statement 4 implies both statements 5 and 6.

Now assume that statement 5 holds, and let $\Psi \in \mathrm{C}(\mathcal{X}, \mathcal{Z})$ be any complementary channel to Φ. For any two unit vectors $u, v \in \mathrm{S}(\mathcal{X})$, Lemma 6.7 implies that

$$\mathrm{F}(\Psi(uu^*), \Psi(vv^*)) = \|\Phi(uv^*)\|_1 = \|uv^*\|_1 = 1, \tag{6.53}$$

and therefore $\Psi(uu^*) = \Psi(vv^*)$. From this fact one concludes that Ψ is constant on $\mathrm{D}(\mathcal{X})$, i.e., there exists $\sigma \in \mathrm{D}(\mathcal{Z})$ for which $\Psi(X) = \operatorname{Tr}(X)\sigma$ for all $X \in \mathrm{L}(\mathcal{X})$. Statement 5 therefore implies statement 7.

Along somewhat similar lines, assume that statement 6 holds, and again let $\Psi \in \mathrm{C}(\mathcal{X}, \mathcal{Z})$ be any complementary channel to Φ. For any choice of orthogonal vectors $u, v \in \mathcal{X}$ it follows by Lemma 6.7 that

$$\|\Psi(uv^*)\|_1 = \mathrm{F}(\Phi(uu^*), \Phi(vv^*)) = \mathrm{F}(uu^*, vv^*) = 0, \tag{6.54}$$

and hence $\Psi(uv^*) = 0$. In particular, this implies that for $E_{ij} \in \mathrm{L}(\mathcal{X})$ with $i \neq j$ one has $\Psi(E_{ij}) = 0$. Furthermore, because

$$E_{ii} - E_{jj} = \frac{1}{2}[(e_i + e_j)(e_i - e_j)^* + (e_i - e_j)(e_i + e_j)^*] \tag{6.55}$$

and $(e_i + e_j) \perp (e_i - e_j)$, it follows that

$$\Psi(E_{ii}) - \Psi(E_{jj}) = \frac{1}{2}\Psi((e_i + e_j)(e_i - e_j)^*) - \frac{1}{2}\Psi((e_i - e_j)(e_i + e_j)^*) = 0. \tag{6.56}$$

That is, there exists $\sigma \in \mathrm{D}(\mathcal{Z})$ for which $\Psi(E_{ii}) = \sigma$ for all $1 \leq i \leq n$. Hence, we have

$$J(\Psi) = \sum_{i,j=1}^{n} E_{ij} \otimes \Psi(E_{ij}) = \mathbb{1}_{\mathcal{X}} \otimes \sigma, \tag{6.57}$$

which is equivalent to $\Psi(X) = \mathrm{Tr}(X)\sigma$ for all $X \in \mathrm{L}(\mathcal{X})$. Statement 6 therefore implies statement 7.

Finally, assume that statement 7 holds. Let $\Psi \in \mathrm{C}(\mathcal{X}, \mathcal{Z})$ be the complementary channel associated with any fixed Stinespring representation

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}(AXA^*) \quad \text{for} \quad A \in \mathrm{U}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z}). \tag{6.58}$$

Assuming that $\sigma \in \mathrm{D}(\mathcal{Z})$ satisfies $\Psi(X) = \mathrm{Tr}(X)\sigma$ for all $X \in \mathrm{L}(\mathcal{X})$, it holds that $J(\Psi) = \mathbb{1}_{\mathcal{X}} \otimes \sigma$, and hence

$$\mathrm{Tr}_{\mathcal{Y}}(\mathrm{vec}(A^{\mathsf{T}})\mathrm{vec}(A^{\mathsf{T}})^*) = \mathbb{1}_{\mathcal{X}} \otimes \sigma = \mathrm{Tr}_{\mathcal{X}}(\mathrm{vec}(\mathbb{1}_{\mathcal{X}})\mathrm{vec}(\mathbb{1}_{\mathcal{X}})^* \otimes \sigma). \tag{6.59}$$

By Lemma 6.8 there exists a channel $\Xi \in \mathrm{C}(\mathcal{Y}, \mathcal{X})$ for which

$$(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Xi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\mathrm{vec}(A^{\mathsf{T}})\mathrm{vec}(A^{\mathsf{T}})^*) = \mathrm{vec}(\mathbb{1}_{\mathcal{X}})\mathrm{vec}(\mathbb{1}_{\mathcal{X}})^* \otimes \sigma. \tag{6.60}$$

By tracing out $\mathcal{Z}$ we get

$$J(\Xi\Phi) = (\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Xi)(J(\Phi)) = \mathrm{vec}(\mathbb{1}_{\mathcal{X}})\mathrm{vec}(\mathbb{1}_{\mathcal{X}})^* = J(\mathbb{1}_{\mathrm{L}(\mathcal{X})}), \tag{6.61}$$

giving $\Xi\Phi = \mathbb{1}_{\mathrm{L}(\mathcal{X})}$. Statement 7 therefore implies statement 1, which completes the proof. □

6.3 Norm conditions for the existence of error-correcting codes

An important notion closely related to reversibility of quantum channels is that of quantum error-correction.

Definition 6.11. We say that a quantum channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ has a k-dimensional error correcting code if there exists a channel $\Psi \in \mathrm{C}(\mathbb{C}^k, \mathcal{X})$ for which $\Phi\Psi$ is reversible.

In other words, a channel has a k-dimensional error correcting code if it is reversible on some copy of $\mathrm{L}(\mathbb{C}^k)$ in the input space $\mathrm{L}(\mathcal{X})$.

Theorem 6.12. *For a quantum channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ the following are equivalent.*

1. *Φ has a k-dimensional error correcting code.*
2. *$\|\Phi \otimes T_k\|_1 = k$.*
3. *$\|\Phi \otimes T_k\|_{1,H} = k$.*

We note that the main difficulty in proving the above theorem is the inclusion of condition 2. For proving the above theorem we require the following generalization of unitary equivalence of purifications.

Proposition 6.13 (Unitary equivalence of purifications)**.** *For unit vectors $u, v, x, y \in \mathcal{X} \otimes \mathcal{Y}$, suppose that for*

$$X \equiv \mathrm{Tr}_{\mathcal{Y}}(uv^*) = \mathrm{Tr}_{\mathcal{Y}}(xy^*) \in \mathrm{L}(\mathcal{X}), \tag{6.62}$$

it holds that $\|X\|_1 = 1$. Then, there exists a unitary $U \in \mathrm{U}(\mathcal{Y})$ for which

$$u = (\mathbb{1}_{\mathcal{X}} \otimes U)x, \text{ and } v = (\mathbb{1}_{\mathcal{X}} \otimes U)y. \tag{6.63}$$

Proof. First consider the special case when $u = v$ and $x = y$, in which case $X \geq 0$. This is the standard set up for the unitary equivalence of purifications for density matrices. Letting $u = \mathrm{vec}(A)$ and $x = \mathrm{vec}(B)$ for $A, B \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$, we have

$$AA^* = \mathrm{Tr}_{\mathcal{Y}}(uu^*) = \mathrm{Tr}_{\mathcal{Y}}(xx^*) = BB^*. \tag{6.64}$$

Hence, there exists a unitary $U \in \mathrm{U}(\mathcal{Y})$ for which $A = BU$, and thus

$$u = \mathrm{vec}(A) = \mathrm{vec}(BU) = (\mathbb{1}_{\mathcal{X}} \otimes U^{\mathsf{T}})\mathrm{vec}(B) = (\mathbb{1}_{\mathcal{X}} \otimes U^{\mathsf{T}})x, \tag{6.65}$$

as required.

Now consider the general case. Let $V \in \mathrm{U}(\mathcal{X})$ be a unitary satisfying $\mathrm{Tr}(VX) = \|X\|_1 = 1$. We then have that

$$\mathrm{Tr}((V \otimes \mathbb{1}_{\mathcal{Y}})uv^*) = \mathrm{Tr}(VX) = 1 = \|uv^*\|_1, \tag{6.66}$$

and hence $(V \otimes \mathbb{1}_{\mathcal{Y}})uv^* \geq 0$, which holds if and only if $v = (V \otimes \mathbb{1}_{\mathcal{Y}})u$. This argument also holds for the pair of vectors x and y (for the same V), and hence $y = (V \otimes \mathbb{1}_{\mathcal{Y}})x$. Thus,

$$\mathrm{Tr}_{\mathcal{Y}}(xx^*) = \mathrm{Tr}_{\mathcal{Y}}(xy^*)V^* = \mathrm{Tr}_{\mathcal{Y}}(uv^*)V^* = \mathrm{Tr}_{\mathcal{Y}}(uu^*). \tag{6.67}$$

and by the special case, there is a unitary $U \in \mathrm{U}(\mathcal{Y})$ for which $u = (\mathbb{1}_{\mathcal{X}} \otimes U)x$, and similarly $v = (V \otimes \mathbb{1}_{\mathcal{Y}})u = (V \otimes U)x = (\mathbb{1}_{\mathcal{X}} \otimes U)y$. □

Lemma 6.14. *Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a quantum channel. If, for unit vectors $u, v \in \mathcal{X} \otimes \mathbb{C}^k$ it holds that $\|(\Phi \otimes T_k)(uv^*)\|_1 = k$, then it holds that*

$$\|(\Phi \otimes T_k)(uu^*)\|_1 = \|(\Phi \otimes T_k)(vv^*)\|_1 = k. \tag{6.68}$$

Proof. First, consider the special case that $\mathcal{X} = \mathcal{Z} \otimes \mathcal{Y}$, and $\Phi = \mathrm{Tr}_{\mathcal{Z}}$. The assumption that

$$\|(\mathrm{Tr}_{\mathcal{Z}} \otimes T_k)(uv^*)\|_1 = k \tag{6.69}$$

implies by Theorem 5.2 that $\dim(\mathcal{Y}) \geq k$, and there exists a positive integer $r \leq \dim(\mathcal{Y})/k$, a density matrix $\sigma \in \mathrm{D}(\mathbb{C}^r)$, and a pair of isometries

$$U, V \in \mathrm{U}(\mathbb{C}^r \otimes \mathbb{C}^k, \mathcal{Y}) \tag{6.70}$$

for which

$$(\mathrm{Tr}_{\mathcal{Z}} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)})(uv^*) = (U \otimes \mathbb{1}_k)(\sigma \otimes \tau_k)(V^* \otimes \mathbb{1}_k), \tag{6.71}$$

We may also assume $r \leq \dim(\mathcal{Z})$, as the rank of the above matrix cannot be greater than $\dim(\mathcal{Z})$.

Let $z \in \mathcal{Z} \otimes \mathbb{C}^r$ be a purification of σ, and define

$$x = \sqrt{\frac{1}{k}}(\mathbb{1}_{\mathcal{Z}} \otimes U \otimes \mathbb{1}_k)(z \otimes \mathrm{vec}(\mathbb{1}_k)) \tag{6.72}$$

and

$$y = \sqrt{\frac{1}{k}}(\mathbb{1}_{\mathcal{Z}} \otimes V \otimes \mathbb{1}_k)(z \otimes \mathrm{vec}(\mathbb{1}_k)). \tag{6.73}$$

Then, it holds that

$$\mathrm{Tr}_{\mathcal{Z}}(xy^*) = \mathrm{Tr}_{\mathcal{Z}}(uv^*), \tag{6.74}$$

and so by the unitary equivalence of purifications (Proposition 6.13), it holds there exists a unitary $R \in \mathrm{U}(\mathcal{Z})$ for which

$$u = (R \otimes \mathbb{1}_{\mathcal{Y}} \otimes \mathbb{1}_k)x \text{ and } v = (R \otimes \mathbb{1}_{\mathcal{Y}} \otimes \mathbb{1}_k)y. \tag{6.75}$$

By absorbing R into z, we may summarize with the conclusion that there exists a purification $z \in \mathcal{Z} \otimes \mathbb{C}^r$ of σ for which

$$u = \sqrt{\frac{1}{k}} (\mathbb{1}_{\mathcal{Z}} \otimes U \otimes \mathbb{1}_k)(z \otimes \text{vec}(\mathbb{1}_k)) \tag{6.76}$$

and

$$v = \sqrt{\frac{1}{k}} (\mathbb{1}_{\mathcal{Z}} \otimes V \otimes \mathbb{1}_k)(z \otimes \text{vec}(\mathbb{1}_k)). \tag{6.77}$$

From this we see that

$$\begin{aligned} \|(\text{Tr}_{\mathcal{Z}} \otimes T_k)(uu^*)\|_1 &= \|(\mathbb{1}_{\text{L}(\mathcal{Y})} \otimes T_k)((U \otimes \mathbb{1}_k)(\sigma \otimes \tau_k)(U^* \otimes \mathbb{1}_k))\|_1 &(6.78)\\ &= \|(\mathbb{1}_{\text{L}(\mathbb{C}^k)} \otimes T_k)(\tau_k)\|_1 &(6.79)\\ &= k. &(6.80)\end{aligned}$$

Similarly, $\|(\text{Tr}_{\mathcal{Z}} \otimes T_k)(vv^*)\|_1 = k$ as required.

For the general case, we use the Stinespring representation of Φ. That is, as Φ is a quantum channel, there exists a complex Euclidean space $\mathcal{Z}$ and an isometry $A \in \text{U}(\mathcal{X}, \mathcal{Z} \otimes \mathcal{Y})$ for which

$$\Phi(X) = \text{Tr}_{\mathcal{Z}}(AXA^*) \tag{6.81}$$

for all $X \in \text{L}(\mathcal{X})$. In terms of this representation, the hypothesis may be written as

$$\|(\text{Tr}_{\mathcal{Z}} \otimes T_k)[(A \otimes \mathbb{1}_k)uv^*(A \otimes \mathbb{1}_k)^*]\|_1 = \|(\Phi \otimes T_k)(uv^*)\|_1 = k, \tag{6.82}$$

and by applying the special case to the vectors $(A \otimes \mathbb{1}_k)u$ and $(A \otimes \mathbb{1}_k)v$, we arrive at

$$\|(\Phi \otimes T_k)(uu^*)\|_1 = \|(\text{Tr}_{\mathcal{Z}} \otimes T_k)[(A \otimes \mathbb{1}_k)uu^*(A \otimes \mathbb{1}_k)^*]\|_1 = k, \tag{6.83}$$

and similarly $\|(\Phi \otimes T_k)(vv^*)\|_1 = k$ as required. □

We can now prove the theorem of this section.

Proof of Theorem 6.12. We will prove $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 1$. For $1 \Rightarrow 2$, first note that

$$\begin{aligned} \|\Phi \otimes T_k\|_1 &= \|(\Phi \otimes \mathbb{1}_{\text{L}(\mathbb{C}^k)})(\mathbb{1}_{\text{L}(\mathcal{X})} \otimes T_k)\|_1 \\ &\leq \|\Phi \otimes \mathbb{1}_{\text{L}(\mathbb{C}^k)}\|_1 \|\mathbb{1}_{\text{L}(\mathcal{X})} \otimes T_k\|_1 \leq k, \end{aligned} \tag{6.84}$$

so we need only prove that $\|\Phi \otimes T_k\|_1 \geq k$. Let $\Psi \in \mathrm{C}(\mathbb{C}^k, \mathcal{X})$ be a channel for which $\Phi\Psi$ is reversible. Then, by Theorem 6.3

$$\|(\mathbb{1}_{\mathrm{L}(\mathcal{Y})} \otimes T_k)(J(\Phi\Psi))\|_1 = k^2, \tag{6.85}$$

and in particular $\|(\Phi \otimes T_k)(\rho)\|_1 = k$ for the density matrix $\rho = (\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)})(\tau_k)$, giving $\|\Phi \otimes T_k\|_1 \geq k$.

For $2 \Rightarrow 3$, let $u, v \in \mathcal{X} \otimes \mathbb{C}^k$ be a pair of unit vectors for which

$$\|(\Phi \otimes T_k)(uv^*)\|_1 = k. \tag{6.86}$$

Lemma 6.14 implies that $\|(\Phi \otimes T_k)(uu^*)\|_1 = \|(\Phi \otimes T_k)(vv^*)\|_1 = k$, giving $\|\Phi \otimes T_k\|_{1,H} = k$.

Finally, for $3 \Rightarrow 1$, let $u \in \mathcal{X} \otimes \mathbb{C}^k$ be a unit vector for which

$$\|(\Phi \otimes T_k)(uu^*)\|_1 = k. \tag{6.87}$$

It holds that

$$k = \|(\Phi \otimes T_k)(uu^*)\|_1 \leq \|(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes T_k)(uu^*)\|_1 \leq k. \tag{6.88}$$

Hence, $\|(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes T_k)(uu^*)\|_1 = k$, and so u is maximally entangled. In particular, this implies that there exists a channel $\Psi \in \mathrm{C}(\mathbb{C}^k, \mathcal{X})$ for which $\frac{1}{k}J(\Psi) = uu^*$. Thus

$$\begin{aligned}\|(\mathbb{1}_{\mathrm{L}(\mathcal{Y})} \otimes T_k)(J(\Phi\Psi))\|_1 &= \|(\Phi \otimes T_k)(J(\Psi))\|_1 \\ &= k\|(\Phi \otimes T_k)(uu^*)\|_1 = k^2,\end{aligned} \tag{6.89}$$

implying by Theorem 6.3 that $\Phi\Psi$ is reversible. □

Chapter 7

Auxiliary dimension in quantum channel discrimination[1]

In this chapter we address the question of how large an auxiliary system needs to be to achieve, for arbitrary quantum channels of fixed input and output dimensions, an optimal strategy. Specifically, given natural numbers $n, m \geq 1$, for what k does it hold that

$$|||\lambda\Gamma_0 - (1-\lambda)\Gamma_1|||_1 = \left\|(\lambda\Gamma_0 - (1-\lambda)\Gamma_1) \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_{1,H} \tag{7.1}$$

for all quantum channels $\Gamma_0, \Gamma_1 \in \mathrm{C}(\mathbb{C}^n, \mathbb{C}^m)$ and probabilities $\lambda \in [0,1]$?

It is well-known that the above holds when $k \geq n$, and the Werner-Holevo channels demonstrate that, when $m \geq n$, $k = n$ is sometimes necessary. Thus, here we ask whether the above might hold for $k < n$ when the output dimension m is (possibly much) smaller than the input dimension n.

We prove the following theorem, joint with John Watrous and appearing in [50], which demonstrates that this is not the case. We thank Gus Gutoski for posing the above question, and therefore motivating this theorem and the work in [50].

Theorem 7.1. *For every choice of positive integers $n \geq 2$ and $k \geq 1$ there exist channels*

$$\Gamma^{(0)}_{n,k}, \Gamma^{(1)}_{n,k} \in \mathrm{C}\big(\mathbb{C}^{n^k}, \mathbb{C}^{kn}\big) \tag{7.2}$$

[1] Note that in this chapter, and thesis generally, we use the term "auxiliary" rather than "ancilla", which is used in the paper these results are drawn from. The reason for the change is the historical usage of "ancilla", pointed out in [71].

such that for all real numbers $\lambda \in (0,1)$ it holds that

$$\left\| \lambda \Gamma^{(0)}_{n,k} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})} - (1-\lambda)\Gamma^{(1)}_{n,k} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})} \right\|_1 < \left\| \left\| \lambda \Gamma^{(0)}_{n,k} - (1-\lambda)\Gamma^{(1)}_{n,k} \right\| \right\|_1 = 1 \quad (7.3)$$

for every complex Euclidean space $\mathcal{Y}$ satisfying $\dim(\mathcal{Y}) < n^k$.

In other words, the theorem states that, even when the output dimension is made arbitrarily small as compared to the input, it is still sometimes necessary to use as much entanglement as is possible to achieve optimal discrimination. Note that more generally, this theorem provides a concrete proof of the fact that the completely bounded trace-norm is not generally achieved with an auxiliary system of dimension equal to that of the output of the map. An equivalent dual statement in terms of the completely bounded norm was proved by Haagerup in [20].

7.1 Proof of Theorem 7.1

The proof of the theorem is by construction. For each $n \geq 2$, let

$$\Phi^{(0)}_n, \Phi^{(1)}_n \in \mathrm{C}(\mathbb{C}^n) \quad (7.4)$$

denote the Werner-Holevo channels given in Definition 4.3, and for a sequence of complex Euclidean spaces $\mathcal{X}_1, \ldots, \mathcal{X}_k$, let $R_i \in \mathrm{C}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k, \mathcal{X}_i)$ denote the reduction to the i^{th} system, as in Chapter 5.

To define the channels appearing in Theorem 7.1, for each $n \geq 2$ and $k \geq 1$, let $\mathcal{X}_1 = \cdots = \mathcal{X}_k = \mathcal{X} = \mathbb{C}^n$. For $\alpha \in \{0,1\}$, we define the channels

$$\Gamma^{(\alpha)}_{n,k} \in \mathrm{C}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k, \mathbb{C}^k \otimes \mathcal{X}) \quad (7.5)$$

for all $X \in \mathrm{L}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k)$ as

$$\Gamma^{(\alpha)}_{n,k}(X) = \frac{1}{k} \sum_{i=1}^{k} E_{ii} \otimes \Phi^{(\alpha)}_n \big(R_i(X)\big), \quad (7.6)$$

where each R_i is regarded as a channel of the form $R_i \in \mathrm{C}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k, \mathcal{X})$. Operationally, these channels represent randomly trashing all but one of the input subsystems while keeping a classical record of which is kept, then applying one of the Werner-Holevo channels to the remaining system. This is represented diagrammatically in Figure 7.1. It holds that $\Gamma^{(\alpha)}_{n,1} \cong \Phi^{(\alpha)}_n$ under the association $\mathbb{C} \otimes \mathcal{X} \cong \mathcal{X}$, and hence the Werner-Holevo channels themselves are contained in this family.

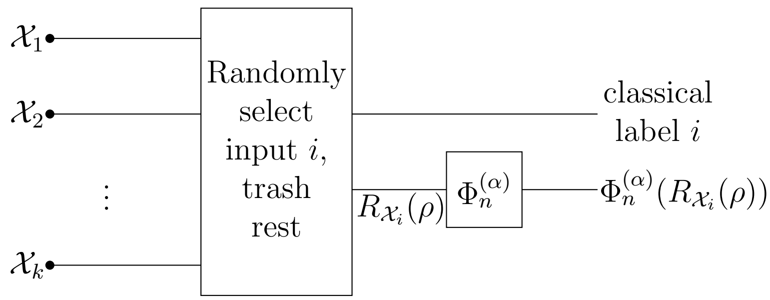


Figure 7.1: Diagrammatic representation of the action of the channel $\Gamma^{(\alpha)}_{n,k}$.

Similarly, define mappings

$$\Psi_{n,k} \in \mathrm{T}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k, \mathbb{C}^k \otimes \mathcal{X}) \tag{7.7}$$

for all $X \in \mathrm{L}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k)$ as

$$\Psi_{n,k}(X) = \sum_{i=1}^{k} E_{ii} \otimes T\big(R_i(X)\big). \tag{7.8}$$

For $\lambda_n = \frac{n+1}{2n}$ the following relations hold

$$\frac{1}{n} T = \lambda_n \Phi^{(0)}_n - (1-\lambda_n)\Phi^{(1)}_n, \tag{7.9}$$

$$\frac{1}{nk} \Psi_{n,k} = \lambda_n \Gamma^{(0)}_{n,k} - (1-\lambda_n)\Gamma^{(1)}_{n,k}. \tag{7.10}$$

The crux of proving Theorem 7.1 will be to prove that

$$\big\| \Psi_{n,k} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})} \big\|_1 < |||\Psi_{n,k}|||_1 = nk \tag{7.11}$$

whenever $\dim(\mathcal{Y}) < n^k$, which is equivalent to the desired norm relation of the theorem for the particular probability λ_n. The specific value λ_n is used to make many expressions easier to work with, and the extension of the result from a particular probability to arbitrary $\lambda \in (0,1)$ will be made by a simple argument.

Before proving the theorem we require a corollary of Theorem 5.4 and a lemma.

Corollary 7.2. *Let $\mathcal{X}, \mathcal{X}_1, \ldots, \mathcal{X}_k$ denote copies of $\mathbb{C}^n$, and let $\mathcal{Y} = \mathbb{C}^m$. For $X \in \mathrm{L}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \otimes \mathcal{Y})$ with $\|X\|_1 = 1$, the following are equivalent.*

1. $\left\|\left(\Psi_{n,k} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\right)(X)\right\|_1 = nk$.
2. $\left\|\left(T_{\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\right)(X)\right\|_1 = n^k$.
3. *It holds that $m \geq n^k$, and there is some $r \in \{1, \ldots, \lfloor m/n^k \rfloor\}$, $\sigma \in \mathrm{D}(\mathbb{C}^r)$, and*
$$U, V \in \mathrm{U}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \otimes \mathbb{C}^r, \mathcal{Y}) \tag{7.12}$$
for which
$$X = (\mathbb{1}_{\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k} \otimes U)(\tau_{\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k} \otimes \sigma)(\mathbb{1}_{\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k} \otimes V^*), \tag{7.13}$$
where $\tau_{\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k} \in \mathrm{D}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \otimes \mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k)$ is the canonical maximally entangled state.

When $X \in \mathrm{D}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \otimes \mathcal{Y})$ the above equivalence holds with $V = U$.

Proof. This is just a specialization of Theorem 5.4 to the task at hand. In particular, in this corollary all of the $\mathcal{X}$ systems are the same dimension, and the first of the equivalent conditions in Theorem 5.4 has been rewritten using the map $\Psi_{n,k}$. To see this, note that individually for each i
$$\|(T_{\mathcal{X}} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})((R_i \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(X))\|_1 \leq n, \tag{7.14}$$
and hence
$$\left\|\left(\Psi_{n,k} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\right)(X)\right\|_1 = \sum_{i=1}^{k} \|(T_{\mathcal{X}} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})((R_i \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(X))\|_1 = nk \tag{7.15}$$
if and only if
$$\|(T_{\mathcal{X}} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})((R_i \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(X))\|_1 = n, \tag{7.16}$$
for all i. □

Lemma 7.3. *Let $\Gamma_0, \Gamma_1 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ be quantum channels. The equation*
$$\|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\|_1 = 1 \tag{7.17}$$
holds for all $\lambda \in (0,1)$ if and only if it holds for a single $\lambda \in (0,1)$.

Proof. We begin the proof with an intermediate claim. Let $A, B \in \mathrm{L}(\mathcal{X})$ have $\|A\|_1 \leq 1$ and $\|B\|_1 \leq 1$. We claim that the equation

$$\|\lambda A - (1-\lambda)B\|_1 = 1 \tag{7.18}$$

holds for all $\lambda \in (0,1)$ if and only if it holds for a particular $\lambda \in (0,1)$. Thus, suppose $\|\alpha A - (1-\alpha)B\|_1 = 1$ for some particular $\alpha \in (0,1)$. Let $U \in \mathrm{U}(\mathcal{X})$ be a unitary for which $\langle U, \alpha A - (1-\alpha)B\rangle = 1$. Hence,

$$\begin{aligned} 1 = \alpha\langle U, A\rangle + (1-\alpha)\langle U, -B\rangle &\leq \alpha|\langle U, A\rangle| + (1-\alpha)|\langle U, -B\rangle| && (7.19)\\ &\leq \alpha\|A\|_1 + (1-\alpha)\|B\|_1 && (7.20)\\ &\leq 1. && (7.21)\end{aligned}$$

As equality holds, it necessarily follows that $\langle U, A\rangle = \langle U, -B\rangle = 1$. Hence, for any $\lambda \in (0,1)$, it holds that

$$\begin{aligned} 1 = \lambda\langle U, A\rangle + (1-\lambda)\langle U, -B\rangle &= \langle U, \lambda A - (1-\lambda)B\rangle \\ &\leq \|\lambda A - (1-\lambda)B\|_1 \leq 1, \end{aligned} \tag{7.22}$$

as required.

Now, returning to the statement of the lemma. If there exists $\alpha \in (0,1)$ for which $\|\alpha\Gamma_0 - (1-\alpha)\Gamma_1\|_1 = 1$, then there exists $X \in \mathrm{L}(\mathcal{X})$ with $\|X\|_1 = 1$ for which $\|\alpha\Gamma_0(X) - (1-\alpha)\Gamma_1(X)\|_1 = 1$. Noting that $\|\Gamma_0(X)\|_1 \leq 1$ and $\|\Gamma_1(X)\|_1 \leq 1$, we may next apply the preceding claim to $A = \Gamma_0(X)$ and $B = \Gamma_1(X)$ to conclude that $\|\lambda\Gamma_0(X) - (1-\lambda)\Gamma_1(X)\|_1 = 1$ for all $\lambda \in (0,1)$. Thus, as it generally holds that

$$\|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\|_1 \leq 1, \tag{7.23}$$

it follows that $\|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\|_1 = 1$ for arbitrary $\lambda \in (0,1)$. □

Proof of Theorem 7.1. Fix $n \geq 2$ and $k \geq 1$, and let $\mathcal{X}_1, \ldots, \mathcal{X}_k$, and $\mathcal{X}$ denote copies of $\mathbb{C}^n$. For our examples we identify

$$\mathbb{C}^{n^k} \cong \mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \quad \text{and} \quad \mathbb{C}^{kn} \cong \mathbb{C}^k \otimes \mathcal{X}. \tag{7.24}$$

Let $\Gamma^{(0)}_{n,k}, \Gamma^{(1)}_{n,k}, \Psi_{n,k} \in \mathrm{T}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k, \mathbb{C}^k \otimes \mathcal{X})$ be as defined in Equation (7.6). For a complex Euclidean space $\mathcal{Y}$ with $\dim(\mathcal{Y}) < n^k$, Lemma 7.3

implies that the equation

$$\left\| \lambda \Gamma_{n,k}^{(0)} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})} - (1-\lambda)\Gamma_{n,k}^{(1)} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})} \right\|_1 < \left|\!\left|\!\left| \lambda \Gamma_{n,k}^{(0)} - (1-\lambda)\Gamma_{n,k}^{(1)} \right|\!\right|\!\right|_1 = 1, \quad (7.25)$$

holds for all $\lambda \in (0,1)$ if and only if it holds for a particular $\lambda \in (0,1)$.

Thus, setting $\lambda_n = \frac{n+1}{2n}$, the theorem will be proved if we show that

$$\left\| \lambda_n \Gamma_{n,k}^{(0)} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})} - (1-\lambda_n)\Gamma_{n,k}^{(1)} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})} \right\|_1 < \left|\!\left|\!\left| \lambda_n \Gamma_{n,k}^{(0)} - (1-\lambda_n)\Gamma_{n,k}^{(1)} \right|\!\right|\!\right|_1 = 1 \quad (7.26)$$

whenever $\dim(\mathcal{Y}) < n^k$. The above is equivalent to showing that

$$\left\| \Psi_{n,k} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})} \right\|_1 < |\!|\!| \Psi_{n,k} |\!|\!|_1 = nk \quad (7.27)$$

whenever $\dim(\mathcal{Y}) < n^k$, where the linear map $\Psi_{n,k}$ is defined in Equation (7.8).

By Corollary 7.2, for the canonical maximally entangled state

$$\tau_{\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k} \in \mathrm{D}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \otimes \mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k) \quad (7.28)$$

it holds that

$$\left\| \left(\Psi_{n,k} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k)} \right) (\tau_{\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k}) \right\|_1 = nk, \quad (7.29)$$

and hence $|\!|\!| \Psi_{n,k} |\!|\!|_1 = nk$. Furthermore, for any complex Euclidean space $\mathcal{Y}$ with $\dim(\mathcal{Y}) < n^k$ and $X \in \mathrm{L}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_k \otimes \mathcal{Y})$ with $\|X\|_1 = 1$, Corollary 7.2 also implies that

$$\left\| \left(\Psi_{n,k} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})} \right)(X) \right\|_1 < nk, \quad (7.30)$$

and hence

$$\left\| \Psi_{n,k} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})} \right\|_1 < nk. \quad (7.31)$$

This completes the proof of Equation (7.27), and therefore completes the proof of the theorem. □

7.2 Further questions

In this chapter we have shown that there exists a family of channel discrimination problems for which a perfect discrimination requires an auxiliary system with dimension equal to that of the input, even when the output dimension is much smaller. Beyond this it would be nice to have a formula for, or even non-trivial bounds on, $\left\|\Psi_{n,k} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\right\|_1$ when $m < n^k$. To serve as a launching ground for future investigations, in Appendix B we have included numerically computed lower bounds for $\left\|\Psi_{n,2} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\right\|_1$ for $2 \le n \le 6$ and $n \le m \le n^2$, computed in MATLAB using QETLAB [31]. More generally, one could try to find non-trivial bounds on

$$\left\|(\lambda\Phi_0 - (1-\lambda)\Phi_1) \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_1 \tag{7.32}$$

for all $\Phi_0, \Phi_1 \in \mathrm{C}(\mathbb{C}^n, \mathbb{C}^m)$ in terms of n, m, k, and $|||\lambda\Phi_0 - (1-\lambda)\Phi_1|||_1$, though this is likely a much more difficult task.

Theorem 5.4 shows that for $m \ge n^k$ the matrices achieving the optimal value have a special form where the ancilla system factorizes into k copies of $\mathbb{C}^n$. This seems intuitively natural, as in the channel discrimination setting, discriminating these channels is like playing k separate Werner-Holevo channel discrimination games using a single resource system, where the referee randomly selects which game will be played and throws away the rest of the input systems. In this setting, Theorem 5.4 says that all optimal strategies are independent, in the sense that the only way of creating an optimal strategy is to stick together k-instances of optimal strategies for discriminating the Werner-Holevo channels. It is thus natural to conjecture that this would be true for $m < n^k$, however this is not the case. For the $k = 2$ case, we show in Proposition A.1 in Appendix A that such independent strategies have the optimal value $n + \lfloor m/n \rfloor$ when $n \le m < n^2$, however, lower bounds on the optimal value computed in Appendix B are well above this.

Chapter 8

On the maximal gap between the optimal entangled and unentangled strategies

In this chapter we prove results motivated by the question of how large the gap between the optimal performance of entangled and unentangled strategies can be (with respect to fixed input and output dimensions). Fully answering this question would involve characterizing the quantum channels

$$\Gamma_0, \Gamma_1 \in \mathrm{C}(\mathbb{C}^n, \mathbb{C}^m) \tag{8.1}$$

and probabilities $\lambda \in (0, 1)$ that attain the maximal gap between the norms

$$\|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\|_{1,H} \ \text{ and } \ \|\!|\lambda\Gamma_0 - (1-\lambda)\Gamma_1|\!\|_1. \tag{8.2}$$

More generally, one could ask: For a linear map $\Psi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, what is the maximal gap between $\|\Psi\|_{1,H}$ and $\|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_{1,H}$? This question may be asked with respect to more specific classes of maps, e.g. Hermiticity preserving maps, or the kind of maps appearing in quantum channel discrimination. It is not currently known what the answer to any of these questions is, but if we replace $\|\cdot\|_{1,H}$ with $\|\cdot\|_1$, the answer is known. For a linear map $\Psi \in \mathrm{T}(\mathbb{C}^n, \mathcal{Y})$, and $1 \leq k \leq n$, it holds that [46, Exercise 3.10]

$$\|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_1 \leq k\|\Psi\|_1, \tag{8.3}$$

and it is known that this bound is saturated by matrix transposition [63].

In this chapter we characterize the linear maps saturating the bound in Equation (8.3), and find that, essentially, the transpose is the unique map achieving this bound. We then leverage this result to prove a uniqueness result for the Werner-Holevo channel discrimination game. In particular, it is essentially the unique game satisfying the norm relation

$$1 = \|\!|\lambda\Gamma_0 - (1-\lambda)\Gamma_1|\!\|_1 = n\|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\|_1, \tag{8.4}$$

for quantum channels Γ_0, Γ_1 with input dimension n and $\lambda \in [0,1]$. This norm relation says that the game can be won with certainty using entanglement, but is hard to win without entanglement, with the bound on unentangled performance provided by $\|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\|_1$ being as small as possible given that the game can be won with certainty.

Finally, we examine the additional property of the Werner-Holevo channel discrimination game that unentangled states cannot be used to improve on the trivial strategy, and prove a purely operational uniqueness result for the Werner-Holevo channel discrimination game; of the games for which entanglement cannot improve on the trivial strategy, the game has the largest possible gap between the trivial strategy and arbitrary entangled strategies.

The results in Sections 8.1 and 8.2 are from [49], and the results of Section 8.3 are new.

8.1 Characterization of linear maps whose multiplicity maps have maximal norm

First, we prove that the transpose uniquely saturates the inquality

$$\|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_1 \leq k\|\Psi\|_1. \tag{8.5}$$

Recall that this inequality has been long known in more generality within the theory of C*-algebras [46, Exercise 3.10]. We first prove the inequality, then analyze equality conditions to arrive at the conclusion.

Theorem 8.1. *Let $\Phi \in \mathrm{T}(\mathbb{C}^n, \mathbb{C}^m)$ be linear with $\|\Phi\|_1 = 1$. It holds that*

$$\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_1 \leq k, \tag{8.6}$$

with equality if and only if $n, m \geq k$, and for any unit vectors $u, v \in \mathbb{C}^n \otimes \mathbb{C}^k$

satisfying

$$\|(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)})(uv^*)\|_1 = k \tag{8.7}$$

(of which at least one such pair must exist), the following statements hold:

1. *u and v are maximally entangled; i.e. there exist isometries*

$$U, V \in \mathrm{U}(\mathbb{C}^k, \mathbb{C}^n) \tag{8.8}$$

for which

$$u = \sqrt{\frac{1}{k}}\mathrm{vec}(U) \text{ and } v = \sqrt{\frac{1}{k}}\mathrm{vec}(V). \tag{8.9}$$

2. *There exists a complete trace-norm isometry*

$$\Psi : T_n(U\mathrm{L}(\mathbb{C}^k)V^*) \to \mathrm{L}(\mathbb{C}^m), \tag{8.10}$$

where

$$T_n(U\mathrm{L}(\mathbb{C}^k)V^*) = \{(UXV^*)^{\mathsf{T}} : X \in \mathrm{L}(\mathbb{C}^k)\}, \tag{8.11}$$

for which $\Phi(X) = \Psi(X^{\mathsf{T}})$ for all $X \in U\mathrm{L}(\mathbb{C}^k)V^*$.

Proof. Letting $u, v \in \mathbb{C}^n \otimes \mathbb{C}^k$ be unit vectors, we will first show that

$$\|(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)})(uv^*)\|_1 \leq k, \tag{8.12}$$

which will prove Equation (8.6). We may assume without loss of generality that u and v have decompositions of the form

$$u = \sum_{a=1}^{r} \alpha_a u_a \otimes e_a \quad \text{and} \quad v = \sum_{b=1}^{r} \beta_b v_b \otimes e_b, \tag{8.13}$$

for $r \leq \min(k, n)$, unit vectors $\alpha, \beta \in \mathbb{C}^r$ with positive entries, and orthonormal sets $\{u_a\}_{a=1}^r, \{v_b\}_{b=1}^r \subset \mathbb{C}^n$. We have

$$\|(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)})(uv^*)\|_1 = \left\| \sum_{a,b=1}^r \alpha_a \beta_b \Phi(u_a v_b^*) \otimes E_{a,b} \right\|_1 \tag{8.14}$$

$$\leq \sum_{a,b=1}^r \alpha_a \beta_b \|\Phi(u_a v_b^*)\|_1 \tag{8.15}$$

$$\leq \sum_{a,b=1}^r \alpha_a \beta_b \|u_a v_b^*\|_1 \tag{8.16}$$

$$= \sum_{a,b=1}^r \alpha_a \beta_b \tag{8.17}$$

$$= \langle 1_r, \alpha \rangle \langle 1_r, \beta \rangle \tag{8.18}$$

$$\leq \|1_r\|^2 \|\alpha\| \|\beta\| \tag{8.19}$$

$$= r \tag{8.20}$$

$$\leq k \tag{8.21}$$

where $1_r \in \mathbb{C}^r$ is the vector of all ones. Hence, it holds that $\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_1 \leq k$.

We now examine equality conditions. Suppose that $\|(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)})(uv^*)\|_1 = k$ for unit vectors $u, v \in \mathbb{C}^n \otimes \mathbb{C}^k$ with decompositions as in Equation (8.13). First, we may conclude that $r = k$, and hence $k \leq n$. Furthermore, equality in the application of Cauchy-Schwarz in Equation (8.19) implies that $\alpha = \beta = \sqrt{\frac{1}{k}} 1_k$, and so u and v are maximally entangled.

Thus, we may write $u = \sqrt{\frac{1}{k}}\mathrm{vec}(U)$ and $v = \sqrt{\frac{1}{k}}\mathrm{vec}(V)$ for isometries $U, V \in \mathrm{U}(\mathbb{C}^k, \mathbb{C}^n)$. It holds that

$$\left\| \sum_{a,b=1}^k \Phi(U E_{a,b} V^*) \otimes E_{a,b} \right\|_1 = k^2, \tag{8.22}$$

and this is equivalent to the more general fact that

$$\left\| \sum_{a,b=1}^k \Phi(x_a y_b^*) \otimes E_{a,b} \right\|_1 = k^2, \tag{8.23}$$

for any orthonormal bases $\{x_a\}_{a=1}^k \subset \mathrm{ran}(U)$ and $\{y_b\}_{b=1}^k \subset \mathrm{ran}(V)$. Since $\|\Phi\|_1 = 1$, the above implies that $\|\Phi(x_a y_b^*)\|_1 = 1$ for all $1 \leq a, b \leq k$. By looking at 2×2 block-sub-matrices, it also implies that, for any unit vectors $x_1, x_2 \in \operatorname{ran}(U)$ and $y_1, y_2 \in \operatorname{ran}(V)$ with $\langle x_1, x_2 \rangle = \langle y_1, y_2 \rangle = 0$, it holds that

$$\left\| \begin{pmatrix} \Phi(x_1 y_1^*) & \Phi(x_1 y_2^*) \\ \Phi(x_2 y_1^*) & \Phi(x_2 y_2^*) \end{pmatrix} \right\|_1 = 4. \tag{8.24}$$

As each block has trace-norm 1, the above 2×2 block matrix has trace-norm equal to the sum of the trace-norms of the blocks. Proposition 6.4 then implies the relations

$$\Phi(x_1 y_1^*)^* \Phi(x_1 y_2^*) = \Phi(x_1 y_1^*) \Phi(x_2 y_1^*)^* = 0 \tag{8.25}$$

for any $x_1, x_2 \in \operatorname{ran}(U)$ and $y_1, y_2 \in \operatorname{ran}(V)$ with $\langle x_1, x_2 \rangle = \langle y_1, y_2 \rangle = 0$.

This may be written in a more suggestive way: for $A, B \in U \mathrm{L}(\mathbb{C}^k) V^*$ rank-1, the following implications hold:

(i) $A^* B = 0$ and $A^* A = B^* B \implies \Phi(A) \Phi(B)^* = 0$, and

(ii) $AB^* = 0$ and $AA^* = BB^* \implies \Phi(A)^* \Phi(B) = 0$.

These implications are very similar to statement 5 in Theorem 6.3, but the adjoints appear in different locations. We may remedy this by defining

$$\Psi : T_n(U \mathrm{L}(\mathbb{C}^k) V^*) \to \mathrm{L}(\mathbb{C}^m) \tag{8.26}$$

as $\Psi = \Phi T_n$, where T_n is the transpose on $\mathrm{L}(\mathbb{C}^n)$. We claim that, for

$$A, B \in T_n(U \mathrm{L}(\mathbb{C}^k) V^*) \tag{8.27}$$

rank-1, the following implications hold:

(a) $A^* B = 0$ and $A^* A = B^* B \implies \Psi(A)^* \Psi(B) = 0$, and

(b) $AB^* = 0$ and $AA^* = BB^* \implies \Psi(A) \Psi(B)^* = 0$.

We will prove that (ii) $\Rightarrow$ (a), with (i) $\Rightarrow$ (b) being similar. Let

$$A^\intercal, B^\intercal \in T_n(U \mathrm{L}(\mathbb{C}^k) V^*) \tag{8.28}$$

be rank-1. The statements $(A^\intercal)^* B^\intercal = 0$ and $(A^\intercal)^* A^\intercal = (B^\intercal)^* B^\intercal$ are equivalent to $AB^* = 0$ and $AA^* = BB^*$, so (ii) implies that $\Phi(A)^* \Phi(B) = 0$, which is in turn equivalent to $\Psi(A^\intercal)^* \Psi(B^\intercal) = 0$.

Thus, by Theorem 6.3, $\Psi = \Phi T_n$ is a complete trace-norm isometry on $T_n(U\mathrm{L}(\mathbb{C}^k)V^*)$, as required.[1] □

As a corollary to Theorems 6.3 and 8.1, we provide two characterizations of the set of linear maps $\Phi \in \mathrm{T}(\mathbb{C}^n, \mathbb{C}^m)$ satisfying $|||\Phi|||_1 = n\|\Phi\|_1$.

Corollary 8.2. *For $\Phi \in \mathrm{T}(\mathbb{C}^n, \mathbb{C}^m)$ linear, the following are equivalent:*

1. $\|\Phi\|_1 = 1$ *and* $|||\Phi|||_1 = n$.
2. $\|J(\Phi)\|_1 = n^2$ *and* $\|J(\Phi T_n)\|_1 = n$.
3. *It holds that $m \geq n$, and there exists a complete trace-norm isometry $\Psi \in \mathrm{T}(\mathbb{C}^n, \mathbb{C}^m)$ for which $\Phi = \Psi T_n$.*

In the above, if Φ is Hermiticity preserving so is Ψ, and if Φ is positive then so is Ψ (and hence is a reversible quantum channel).

Proof. It is immediate that $3 \Rightarrow 1$ and $3 \Rightarrow 2$. That $1 \Rightarrow 3$ is given by Theorem 8.1, and that $2 \Rightarrow 3$ is given by Theorem 6.3. For the special cases, since $\Psi = \Phi T_n$, if Φ is Hermiticity preserving so is Ψ, as it is a composition of Hermiticity preserving maps. The same logic applies if Φ is positive; with Ψ being a reversible quantum channel following from the positive case of Theorem 6.3. □

8.2 A uniqueness result for the Werner-Holevo channels

For $n \geq 2$, recall the definition of the Werner-Holevo channels given in Definition 4.3, $\Phi_n^{(0)}, \Phi_n^{(1)} \in \mathrm{C}(\mathbb{C}^n)$, and $\lambda = \frac{n+1}{2n} \in (0,1)$. The channel discrimination specified by the triple $(\lambda_n, \Phi_n^{(0)}, \Phi_n^{(1)})$ satisfies the following norm relations:

$$
\begin{aligned}
1 = |||\lambda\Phi_n^{(0)} - (1-\lambda)\Phi_n^{(1)}|||_1 &= n\|\lambda\Phi_n^{(0)} - (1-\lambda)\Phi_n^{(1)}\|_1 && (8.29)\\
&= n\|\lambda\Phi_n^{(0)} - (1-\lambda)\Phi_n^{(1)}\|_{1,H} && (8.30)
\end{aligned}
$$

[1] Note that Theorem 6.3 as stated only applies to maps whose domain is all of $\mathrm{L}(\mathbb{C}^n)$. Here, the domain of Ψ is $T_n(U\mathrm{L}(\mathbb{C}^k)V^*) = \overline{V}\mathrm{L}(\mathbb{C}^k)U^{\intercal} \subset \mathrm{L}(\mathbb{C}^n)$, so technically we are applying Theorem 6.3 to conclude that the linear map $X \mapsto \Psi(V^{\intercal}X\overline{U})$ is a complete trace-norm isometry on $\mathrm{L}(\mathbb{C}^k)$. However, this is equivalent to Ψ being a complete trace-norm isometry on $T_n(U\mathrm{L}(\mathbb{C}^k)V^*)$.

Notice that, in particular, the gap between the norms

$$\|\lambda\Phi_n^{(0)} - (1-\lambda)\Phi_n^{(1)}\|_1, \text{ and } \|\!|\lambda\Phi_n^{(0)} - (1-\lambda)\Phi_n^{(1)}|\!\|_1 \tag{8.31}$$

is as large as it can possibly be, given that

$$\|\!|\lambda\Phi_n^{(0)} - (1-\lambda)\Phi_n^{(1)}|\!\|_1 \leq n\|\lambda\Phi_n^{(0)} - (1-\lambda)\Phi_n^{(1)}\|_1 \tag{8.32}$$

and

$$\|\lambda\Phi_n^{(0)} - (1-\lambda)\Phi_n^{(1)}\|_1 \leq \|\!|\lambda\Phi_n^{(0)} - (1-\lambda)\Phi_n^{(1)}|\!\|_1 \leq 1. \tag{8.33}$$

The following theorem gives that the channel discrimination game specified by $(\lambda_n, \Phi_n^{(0)}, \Phi_n^{(1)})$ is in some sense unique in satisfying this norm relation.

Theorem 8.3. *Let $\Gamma_0, \Gamma_1 \in \mathrm{C}(\mathbb{C}^n, \mathbb{C}^m)$ be quantum channels, $\lambda \in (0,1)$ be a probability, and let $\Phi_n^{(0)}, \Phi_n^{(1)} \in \mathrm{C}(\mathbb{C}^n)$ be the Werner-Holevo channels as given in Definition 4.3. It holds that*

$$1 = \|\!|\lambda\Gamma_0 - (1-\lambda)\Gamma_1|\!\|_1 = n\|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\|_1 \tag{8.34}$$

if and only if $m \geq n$ and:

- *For $m < 2n$, there exists a reversible quantum channel $\Psi \in \mathrm{C}(\mathbb{C}^n, \mathbb{C}^m)$ for which either*

$$(\lambda, \Gamma_0, \Gamma_1) = (\lambda_n, \Psi\Phi_n^{(0)}, \Psi\Phi_n^{(1)}) \tag{8.35}$$

  or

$$(\lambda, \Gamma_0, \Gamma_1) = (1-\lambda_n, \Psi\Phi_n^{(1)}, \Psi\Phi_n^{(0)}). \tag{8.36}$$

- *For $m \geq 2n$, there exists $r \in [0,1]$ and two reversible channels*

$$\Psi_0, \Psi_1 \in \mathrm{C}(\mathbb{C}^n, \mathbb{C}^m) \tag{8.37}$$

  with orthogonal ranges for which $\lambda = r\lambda_n + (1-r)(1-\lambda_n)$, and

$$\lambda\Gamma_0 = r\lambda_n\Psi_0\Phi_n^{(0)} + (1-r)(1-\lambda_n)\Psi_1\Phi_n^{(1)}, \tag{8.38}$$

  and

$$(1-\lambda)\Gamma_1 = r(1-\lambda_n)\Psi_0\Phi_n^{(1)} + (1-r)\lambda_n\Psi_1\Phi_n^{(0)}. \tag{8.39}$$

Remark 8.4. Theorem 8.3 may be interpreted as saying that the game specified by $(\lambda_n, \Phi_n^{(0)}, \Phi_n^{(1)})$ uniquely satisfies Equation (8.34) in the following sense: Any game $(\lambda, \Gamma_0, \Gamma_1)$, whose channels have domain $\mathrm{L}(\mathbb{C}^n)$ and satisfy Equation (8.34), is constructed out of, and is reducible by the player to, the game

$(\lambda_n, \Phi_n^{(0)}, \Phi_n^{(1)})$ in a way that perfectly preserves success probabilities. Indeed, mathematically, one can check that

$$\begin{aligned}\|\lambda(\Gamma_0 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)})(X) - (1-\lambda)(\Gamma_1 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)})(X)\|_1 \\ = \left\|\lambda_n\big(\Phi_n^{(0)} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\big)(X) - (1-\lambda_n)\big(\Phi_n^{(1)} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\big)(X)\right\|_1\end{aligned} \tag{8.40}$$

for all integers $k \geq 1$ and matrices $X \in \mathrm{L}(\mathbb{C}^n \otimes \mathbb{C}^k)$. Operationally, the construction/reduction of such games $(\lambda, \Gamma_0, \Gamma_1)$ in terms of $(\lambda_n, \Phi_n^{(0)}, \Phi_n^{(1)})$ goes as follows.

- For the case $m < 2n$, the construction and reduction are natural; up to a reversible quantum channel (which the player can undo) and a relabeling of the channels (which the player knows), the game $(\lambda, \Gamma_0, \Gamma_1)$ is exactly the game $\big(\lambda_n, \Phi_n^{(0)}, \Phi_n^{(1)}\big)$.
- For the case $m \geq 2n$, the relation between $(\lambda, \Gamma_0, \Gamma_1)$ and $\big(\lambda_n, \Phi_n^{(0)}, \Phi_n^{(1)}\big)$ is less clear, though it can be thought of as a convex combination of relabelings of the game $\big(\lambda_n, \Phi_n^{(0)}, \Phi_n^{(1)}\big)$, where the player is able to detect which labeling is being used. Specifically, with probability r, Γ_0 acts as $\Phi_n^{(0)}$ and Γ_1 acts as $\Phi_n^{(1)}$, and with probability $(1-r)$ the labels are reversed. As Ψ_0 and Ψ_1 have orthogonal ranges, the player is able to measure which labelling is being used without disturbance. Once this is done, the situation from the players perspective is now the same as in the case $m < 2n$, and they may act accordingly.

Before proving Theorem 8.3, we prove a lemma regarding the uniqueness of certain decompositions of Hermiticity preserving maps into differences of completely positive maps.

Lemma 8.5. *Let $\Phi \in \mathrm{T}(\mathbb{C}^n, \mathbb{C}^m)$ be Hermiticity preserving,*

$$\Psi_0, \Psi_1 \in \mathrm{CP}(\mathbb{C}^n, \mathbb{C}^m) \tag{8.41}$$

be completely positive and satisfy

$$\Phi = \Psi_0 - \Psi_1 \text{ and } |||\Phi|||_1 = |||\Psi_0|||_1 + |||\Psi_1|||_1, \tag{8.42}$$

and let $u \in \mathbb{C}^n \otimes \mathbb{C}^n$ be a unit vector satisfying $|||\Phi|||_1 = \|(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*)\|_1$. It follows that

$$|||\Psi_0|||_1 = \|(\Psi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*)\|_1 \text{ and } |||\Psi_1|||_1 = \left\|(\Psi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*)\right\|_1, \tag{8.43}$$

and for any other completely positive maps $\Psi_0', \Psi_1' \in \mathrm{CP}(\mathbb{C}^n, \mathbb{C}^m)$ *satisfying the conditions in Equation* (8.42)*,*

$$(\Psi_0' \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*) = (\Psi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*) \text{ and } (\Psi_1' \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*) = (\Psi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*). \quad (8.44)$$

Hence, if such a u *exists with full Schmidt-rank, the completely positive maps* Ψ_0, Ψ_1 *satisfying Equation* (8.42) *are unique (if they exist).*

Proof. We have

$$|||\Phi|||_1 = \|(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*)\|_1 \quad (8.45)$$
$$= \|(\Psi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*) - (\Psi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*)\|_1 \quad (8.46)$$
$$\leq \|(\Psi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*)\|_1 + \|(\Psi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*)\|_1 \quad (8.47)$$
$$\leq |||\Psi_0|||_1 + |||\Psi_1|||_1 \quad (8.48)$$
$$= |||\Phi|||_1. \quad (8.49)$$

Hence, all inequalities are equalities, and therefore

$$|||\Psi_0|||_1 = \|(\Psi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*)\|_1 \quad (8.50)$$

and

$$|||\Psi_1|||_1 = \|(\Psi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*)\|_1. \quad (8.51)$$

Next, as Ψ_0 and Ψ_1 are completely positive, it holds that both

$$(\Psi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*) \geq 0 \text{ and } (\Psi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*) \geq 0, \quad (8.52)$$

and so equality in Equation (8.47) implies that

$$(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*) = (\Psi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*) - (\Psi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*) \quad (8.53)$$

is the Hahn decomposition of $(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*)$. Thus, for any other completely positive maps $\Psi_0', \Psi_1' \in \mathrm{CP}(\mathbb{C}^n, \mathbb{C}^m)$ satisfying the hypotheses,

$$(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*) = (\Psi_0' \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*) - (\Psi_1' \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*) \quad (8.54)$$

is also the Hahn decomposition of $(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*)$. Equation (8.44) therefore follows by the uniqueness of the Hahn decomposition.

Finally, if $u \in \mathbb{C}^n \otimes \mathbb{C}^n$ has full Schmidt-rank, then a linear map

$$\Gamma \in \mathrm{T}(\mathbb{C}^n, \mathbb{C}^m) \tag{8.55}$$

is uniquely specified by the matrix $(\Gamma \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(uu^*)$, and so Equation (8.44) implies the uniqueness of the pair Ψ_0 and Ψ_1 (assuming such a pair exists). □

In the above lemma, the premise is the existence of a decomposition of a Hermiticity preserving map into a difference of completely positive maps satisfying a particular norm relation. We note that this is similar in flavour to the conclusion of Wittstock's decomposition theorem [46, Theorem 8.5].

Proof of Theorem 8.3. In both cases the "if" part is a matter of verifying Equation (8.40), where the case $m \geq 2n$ requires use of the fact that Ψ_0 and Ψ_1 have orthogonal ranges.

Thus, assume we have a channel discrimination triple $(\lambda, \Gamma_0, \Gamma_1)$ satisfying Equation (8.34). By Corollary 8.2, the norm relation implies

$$\lambda\Gamma_0 - (1-\lambda)\Gamma_1 = \frac{1}{n}\Psi T_n, \tag{8.56}$$

for $\Psi \in \mathrm{T}(\mathbb{C}^n, \mathbb{C}^m)$ a Hermiticity preserving complete trace-norm isometry. Remark 6.6 gives the following structure for Ψ:

- If $m < 2n$, either Ψ or $-\Psi$ is a reversible quantum channel.
- If $m \geq 2n$, there exists $r \in [0,1]$ and $\Psi_0, \Psi_1 \in \mathrm{C}(\mathbb{C}^n, \mathbb{C}^m)$ reversible quantum channels with orthogonal ranges for which $\Psi = r\Psi_0 - (1-r)\Psi_1$.

In what follows we will work with the form of Ψ in the case $m \geq 2n$, as the case $m < 2n$ can be subsumed by the case $r = 0$ or $r = 1$ when $m \geq 2n$, even though it is not possible for two reversible channels $\Psi_0, \Psi_1 : \mathrm{C}(\mathbb{C}^n, \mathbb{C}^m)$ to have orthogonal ranges when $m < 2n$.

Observe the following facts:

- $\frac{1}{n}\Psi T_n$ is Hermiticity preserving and decomposes as a difference of completely positive maps as given in Equation (8.56),
- $\left\|\left|\frac{1}{n}\Psi T_n\right|\right\|_1 = 1 = \left\|\left|\lambda\Gamma_0\right|\right\|_1 + \left\|\left|(1-\lambda)\Gamma_1\right|\right\|_1$, and
- $\left\|\left|\frac{1}{n}\Psi T_n\right|\right\|_1 = \left\|\frac{1}{n}(\Psi T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^n)})(\tau_n)\right\|_1$, where $\tau_n \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^n)$ is the canonical maximally entangled state.

When taken together these facts imply, by Lemma 8.5, that Equation (8.56) is the *unique* decomposition of $\frac{1}{n}\Psi T_n$ into a difference of CP maps with the above properties. In the remainder of the proof, we will exhibit a (seemingly) different decomposition of $\frac{1}{n}\Psi T_n$, verify that it also satisfies the assumptions of Lemma 8.5, then conclude that the two decompositions are necessarily the same.

Note that $\frac{1}{n}T_n = \lambda_n \Phi_0^{(n)} - (1-\lambda_n)\Phi_1^{(n)}$, and hence

$$\begin{aligned}\frac{1}{n}\Psi T_n &= (r\Psi_0 - (1-r)\Psi_1)(\lambda_n \Phi_0^{(n)} - (1-\lambda_n)\Phi_1^{(n)}) \\ &= \left[r\lambda_n \Psi_0 \Phi_0^{(n)} + (1-r)(1-\lambda_n)\Psi_1\Phi_1^{(n)}\right] \\ &\qquad\qquad - \left[(1-r)\lambda_n\Psi_1\Phi_0^{(n)} + r(1-\lambda_n)\Psi_0\Phi_1^{(n)}\right].\end{aligned} \tag{8.57}$$

The maps in the square brackets are completely positive, and satisfy

$$\begin{aligned}&\left|\left|\left| r\lambda_n \Psi_0 \Phi_0^{(n)} + (1-r)(1-\lambda_n)\Psi_1\Phi_1^{(n)}\right|\right|\right|_1 \\ &\qquad + \left|\left|\left|(1-r)\lambda_n\Psi_1\Phi_0^{(n)} + r(1-\lambda_n)\Psi_0\Phi_1^{(n)}\right|\right|\right|_1 \\ &\qquad\quad = r\lambda_n + (1-r)(1-\lambda_n) + (1-r)\lambda_n + r(1-\lambda_n) \\ &\qquad\quad = 1.\end{aligned} \tag{8.58}$$

Hence, by the uniqueness clause of Lemma 8.5, Equations (8.38) and (8.39) hold.

When $m < 2n$, either $r = 0$ or $r = 1$, in which case either

$$(\lambda, \Gamma_0, \Gamma_1) = \left(\lambda_n, \Psi_0\Phi_n^{(0)}, \Psi_0\Phi_n^{(1)}\right) \tag{8.59}$$

or

$$(\lambda, \Gamma_0, \Gamma_1) = \left(1-\lambda_n, \Psi_1\Phi_n^{(1)}, \Psi_1\Phi_n^{(0)}\right), \tag{8.60}$$

as required. □

8.3 The trivial strategy and the Werner-Holevo channels

For a channel discrimination game specified by channels $\Phi_0, \Phi_1 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ and $\lambda \in (0,1)$, the *trivial strategy* is to simply guess whichever channel had the higher probability of occurring. That is, the trivial strategy is to try to directly guess the bit α sampled by the referee, while ignoring the additional resource of a use of the channel Φ_α. This strategy succeeds with probability $\max(\lambda, 1-\lambda)$.

In this section we will consider the structure of games that have the property that, without entanglement, the trivial strategy cannot be improved upon. That is, games for which

$$\frac{1}{2} + \frac{1}{2}\|\lambda\Phi_0 - (1-\lambda)\Phi_1\|_{1,H} = \max(\lambda, 1-\lambda). \tag{8.61}$$

In some sense, if a game satisfies this property, then no information can be gained about which channel acted without using entanglement.

Before beginning we need a definition.

Definition 8.6. A linear map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is called *trace-scaling* if there exists $\alpha \in \mathbb{C}$ for which $\mathrm{Tr}(\Phi(X)) = \alpha\mathrm{Tr}(X)$ for all $X \in \mathrm{L}(\mathcal{X})$. We call α the trace-constant of Φ.

The Werner-Holevo channel discrimination game $(\lambda_n, \Phi_n^{(0)}, \Phi_n^{(1)})$ provides an example of a game where the trivial strategy cannot be improved on without using entanglement. To see this, we have $\lambda_n\Phi_n^{(0)} - (1-\lambda_n)\Phi_n^{(1)} = \frac{1}{n}T_n$, and so for any density matrix $\rho \in \mathrm{D}(\mathbb{C}^n)$, it holds that

$$\|\lambda_n\Phi_n^{(0)}(\rho) - (1-\lambda_n)\Phi_n^{(1)}(\rho)\|_1 = \frac{1}{n}\|\rho^{\mathsf{T}}\|_1 = \frac{1}{n} = 2\lambda_n - 1, \tag{8.62}$$

which leads to Equation (8.61). In particular, this follows from the difference $\lambda_n\Phi_n^{(0)} - (1-\lambda_n)\Phi_n^{(1)} = \frac{1}{n}T_n$ being a positive and trace-scaling map.

The construction of the Werner-Holevo channels essentially ensures the above property; for $\Psi \in \mathrm{T}(\mathbb{C}^n)$ being the completely-depolarizing channel (i.e. $\Psi(X) = \frac{1}{n}\mathrm{Tr}(X)\mathbb{1}_n$ for all $X \in \mathrm{L}(\mathbb{C}^n)$), it holds that

$$\Phi_n^{(0)} = \frac{\Psi + \frac{1}{n}T_n}{2\lambda_n}, \text{ and } \Phi_n^{(1)} = \frac{\Psi - \frac{1}{n}T_n}{2(1-\lambda_n)}. \tag{8.63}$$

When written this way, they seem explicitly constructed so that

$$\lambda_n\Phi_n^{(0)} - (1-\lambda_n)\Phi_n^{(1)} \tag{8.64}$$

is positive and trace-scaling, and one could imagine constructing in a similar way other examples for which the trivial strategy cannot be improved upon without entanglement. Let $\Psi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ be a channel, and $\Lambda \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a positive and trace scaling map with constant $\alpha \in (0,1)$, for which both $\Psi \pm \Lambda$ are completely positive. Set $\lambda = \frac{1+\alpha}{2} \in (0,1)$ and define

$$\Phi_0 = \frac{\Phi + \Psi}{2\lambda}, \text{ and } \Phi_1 = \frac{\Phi - \Psi}{2(1-\lambda)}. \tag{8.65}$$

It follows that $\lambda\Phi_0 - (1-\lambda)\Phi_1 = \Lambda$, and hence, by the same reasoning as for the Werner-Holevo channels, the game specified by the triple $(\lambda, \Phi_0, \Phi_1)$ has the property that unentangled states do not provide an advantage over the trivial strategy. Something similar to this construction is done in [47], in which it is shown that for any entangled state, there exists a channel discrimination problem for which it can be used to outperform all unentangled states.

The following proposition and corollary show that in fact, the above structure is generic for channel discrimination games for which unentangled states cannot be used to outperform the trivial strategy.

Proposition 8.7. *Let $\lambda \in [\frac{1}{2}, 1)$ and let $\Phi_0, \Phi_1 \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be positive and trace-preserving maps. The following are equivalent.*

1. *For all density matrices $\rho \in \mathrm{D}(\mathcal{X})$, it holds that*

$$\|\lambda\Phi_0(\rho) - (1-\lambda)\Phi_1(\rho)\|_1 = 2\lambda - 1. \tag{8.66}$$

2. *The map $\lambda\Phi_0 - (1-\lambda)\Phi_1$ is positive.*

3. $\|\lambda\Phi_0 - (1-\lambda)\Phi_1\|_1 = 2\lambda - 1$

4. $\|\lambda\Phi_0 - (1-\lambda)\Phi_1\|_{1,H} = 2\lambda - 1$

Proof. First note that for any density matrix $\rho \in \mathrm{D}(\mathcal{X})$,

$$\begin{aligned} \|\lambda\Phi_0(\rho) - (1-\lambda)\Phi_1(\rho)\|_1 &\geq \mathrm{Tr}(\lambda\Phi_0(\rho)_1 - (1-\lambda)\Phi_1(\rho)) && (8.67)\\ &= \lambda - (1-\lambda) && (8.68)\\ &= 2\lambda - 1. && (8.69) \end{aligned}$$

In the context of channel discrimination, this says that any density matrix cannot do worse than the trivial strategy (assuming the player does not purposefully choose a bad measurement).

For 1 $\implies$ 2, observe that the assumption implies equality in Equation (8.67) for all $\rho \in \mathrm{D}(\mathcal{X})$. As $\mathrm{Tr}(A) = \|A\|_1$ if and only if $A \geq 0$, this implies that

$$\lambda\Phi_0(\rho) - (1-\lambda)\Phi_1(\rho) \geq 0 \tag{8.70}$$

for all $\rho \in \mathrm{D}(\mathcal{X})$, and hence $\lambda\Phi_0 - (1-\lambda)\Phi_1$ is positive.

For 2 $\implies$ 3, as the map is positive, the norm is achieved on some density matrix $\rho \in \mathrm{D}(\mathcal{X})$, and hence

$$\|\lambda\Phi_0 - (1-\lambda)\Phi_1\|_1 = \mathrm{Tr}(\lambda\Phi_0(\rho) - (1-\lambda)\Phi_1(\rho)) = 2\lambda - 1, \tag{8.71}$$

as required.

For 3 $\implies$ 4, we have

$$2\lambda - 1 = \|\lambda\Phi_0 - (1-\lambda)\Phi_1\|_1 \geq \|\lambda\Phi_0 - (1-\lambda)\Phi_1\|_{1,H} \geq 2\lambda - 1, \tag{8.72}$$

where the first inequality is by definition, and the second is from Equation (8.67).

Similarly, 4 $\implies$ 1 follows from

$$2\lambda - 1 = \|\lambda\Phi_0 - (1-\lambda)\Phi_1\|_{1,H} \geq \|\lambda\Phi_0(\rho) - (1-\lambda)\Phi_1(\rho)\|_1 \geq 2\lambda - 1. \tag{8.73}$$

□

Corollary 8.8. *For a single-shot channel discrimination game specified by channels* $\Phi_0, \Phi_1 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ *and a probability* $\lambda \in [1/2, 1)$*, it holds that the trivial strategy cannot be improved upon without using entanglement if and only if there exists a quantum channel* $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ *and a positive and trace-scaling map* $\Psi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ *for which*

$$\Phi_0 = \frac{\Phi + \Psi}{2\lambda}, \text{ and } \Phi_1 = \frac{\Phi - \Psi}{2(1-\lambda)}. \tag{8.74}$$

Proof. Note that if Φ_0 and Φ_1 have the form in the above equation, then $\lambda\Phi_0 - (1-\lambda)\Phi_1 = \Psi$, which is positive, and so the preceding proposition implies the trivial strategy cannot be improved upon without entanglement.

Conversely, if the trivial strategy cannot be improved upon without entanglement, then setting $\Psi = \lambda\Phi_0 - (1-\lambda)\Phi_1$, the preceding proposition gives Ψ is positive, and its form as a weighted difference of channels implies it is trace-scaling. Defining $\Phi = \lambda\Phi_0 + (1-\lambda)\Phi_1$, Φ is clearly a channel, and Φ and Ψ satisfy the required relations. □

The above proposition and corollary provide some insight into the structure of any examples of channels $\Phi_0, \Phi_1 \in \mathrm{C}(\mathbb{C}^n, \mathcal{Y})$ and $\lambda \in (0,1)$ that could potentially satisfy

$$\left\|\left|\lambda\Phi_0 - (1-\lambda)\Phi_1\right\|\right|_1 > n\|\lambda\Phi_0 - (1-\lambda)\Phi_1\|_{1,H}. \tag{8.75}$$

In particular, for the above to hold we cannot have

$$\|\lambda\Phi_0 - (1-\lambda)\Phi_1\|_1 = \|\lambda\Phi_0 - (1-\lambda)\Phi_1\|_{1,H}, \tag{8.76}$$

i.e. the map $\lambda\Phi_0 - (1-\lambda)\Phi_1$ cannot be positive, and equivalently by the above proposition, it will actually be necessary that there exists an unentangled strategy that does better than the trivial strategy.

Proposition 8.7 enables a version of Theorem 8.3 that gives the uniqueness of the Werner-Holevo game in terms of purely operational statements.

Corollary 8.9. *Let $\Gamma_0, \Gamma_1 \in \mathrm{C}(\mathbb{C}^n, \mathbb{C}^m)$ be quantum channels, and $\lambda \in [\frac{1}{2}, 1)$. The following are equivalent.*

1. *In the channel discrimination game specified by $(\lambda, \Gamma_0, \Gamma_1)$, the trivial strategy cannot be improved upon without using entanglement. Furthermore, of the games with this property and input dimension n, it has the maximal performance gap between the trivial strategy and arbitrary entangled strategies.*

2. *$m \geq n$, $\lambda = \lambda_n = \frac{n+1}{2n}$ and there exists a reversible quantum channel $\Psi \in \mathrm{C}(\mathbb{C}^n, \mathbb{C}^m)$ for which $\Gamma_0 = \Psi\Phi_n^{(0)}$ and $\Gamma_1 = \Psi\Phi_n^{(1)}$.*

Proof. For 1 $\implies$ 2, as the trivial strategy cannot be improved on without entanglement, Proposition 8.7 implies the norm relation

$$\|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\|_1 = \|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\|_{1,H} = 2\lambda - 1. \tag{8.77}$$

Hence, as it generally holds for games of input dimension n that

$$\left\|\left|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\right\|\right|_1 \leq n\|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\|_1, \tag{8.78}$$

the maximum achievable gap between these norms (and hence the maximum gap between the optimal value of entangled and unentangled strategies) occurs when

$$1 = \left\|\left|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\right\|\right|_1 = n\|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\|_1. \tag{8.79}$$

Hence, statement 1 is equivalent to both Equations (8.77) and (8.79) holding.

First, Equations (8.77) and (8.79) directly imply that $2\lambda - 1 = \frac{1}{n}$, and hence $\lambda = \lambda_n$. By Equation (8.79) Theorem 8.3 implies that $m \geq n$, and by virtue of the fact that we already know $\lambda = \lambda_n$, it further implies the required structure of Γ_0 and Γ_1 in statement 2.

For $2 \implies 1$, we have that $\frac{1}{n}\Psi T_n = \lambda\Gamma_0 - (1-\lambda)\Gamma_1$. Thus, Equations (8.77) and (8.79) both hold, and this has already been argued to be equivalent to statement 1. □

8.4 Further work and questions

The results in this chapter were motivated by the problem of characterizing the maximal gap between

$$\|\lambda\Gamma_0 - (1-\lambda)\Gamma_1\|_{1,H}, \text{ and } \|\!|\lambda\Gamma_0 - (1-\lambda)\Gamma_1|\!\|_1 \tag{8.80}$$

for quantum channels $\Gamma_0, \Gamma_1 \in \mathrm{C}(\mathbb{C}^n, \mathcal{Y})$ and $\lambda \in [0,1]$, and this question remains open. The main difficulty is proving a bound on $\|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_{1,H}$ in terms of $\|\Psi\|_{1,H}$. For example, does it generally hold that

$$\|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_{1,H} \leq k\|\Psi\|_{1,H}? \tag{8.81}$$

What if Ψ is assumed to be Hermiticity preserving? Some partial progress on this question and some related questions are discussed in the next chapter. It seems natural to conjecture that the above bound holds, and that transposition will uniquely saturate it.

Chapter 9

Partial results and conjectures

In this chapter we present partial results and conjectures for two continuations of the work in this thesis. We consider the following:

- In Section 9.1 we consider the question of bounding $\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_{1,H}$ in terms of $\|\Phi\|_{1,H}$, and show that a particular naive approach via modification of the proof of $\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_1 \leq k\|\Phi\|_1$ given in Theorem 8.1 cannot lead to a tight bound. We also conjecture that if Φ is Hermiticity preserving, then $\|\Phi\|_1 \leq \sqrt{2}\|\Phi\|_{1,H}$, and show how to prove it conditional on the conjecture that, for any Hermiticity preserving map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ and Hermitian matrices $H, K \in \mathrm{Herm}(\mathcal{X})$, it holds that

$$\left\| \begin{pmatrix} \Phi(H) \\ \Phi(K) \end{pmatrix} \right\|_1 \leq \|\Phi\|_{1,H} \left\| \begin{pmatrix} H \\ K \end{pmatrix} \right\|_1. \tag{9.1}$$

  A parallel approach in terms of the operator norm is also given.

- In Section 9.2 we consider the problem of proving a robust, or approximate, version of Theorem 5.2. In Theorem 5.2 the structure of matrices having maximal negativity was characterized, and here we consider the structure of density matrices $\rho \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^m)$ satisfying

$$\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(\rho)\|_1 \geq n(1-\epsilon) \tag{9.2}$$

  for $\epsilon \in [0, 1]$.

9.1 Investigating growth of $\left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_{1,H}$

In Chapter 8, we characterized linear maps $\Phi \in \mathrm{T}(\mathbb{C}^n, \mathbb{C}^m)$ saturating the known inequality $\left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_1 \leq k\|\Phi\|_1$. However, for applications in quantum channel discrimination, it is necessary to bound $\left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_{1,H}$ by $\|\Phi\|_{1,H}$. In particular, we are interested in finding a bound of the form

$$\left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_{1,H} \leq f(k)\|\Phi\|_{1,H} \tag{9.3}$$

for a function $f : \mathbb{N} \to \mathbb{R}$.

We expect that the optimal form of f will vary across different classes of maps, for example, when Φ is arbitrary, or when Φ is Hermiticity preserving. In the context of quantum channel discrimination, it also makes sense to consider the further restricted class of maps of the form $\lambda\Gamma_0 - (1-\lambda)\Gamma_1$ for quantum channels $\Gamma_0, \Gamma_1 \in \mathrm{C}(\mathbb{C}^n, \mathbb{C}^m)$ and $\lambda \in [0,1]$. As the the following proposition shows, this class of maps is (up to real-valued scalars) exactly the set of Hermiticity preserving and trace-scaling maps. Note that a similar fact is proven as Lemma 2 in [47], in which Hermiticity preserving and trace-annihilating maps are characterized in a similar way.

Proposition 9.1. *For a linear map $\Psi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ the following are equivalent.*

1. *Ψ is Hermiticity preserving and trace-scaling.*
2. *$\Psi = \Psi_0 - \Psi_1$ for $\Psi_0, \Psi_1 \in \mathrm{CP}(\mathcal{X}, \mathcal{Y})$ completely positive and trace-scaling.*
3. *$\Psi = \alpha(\lambda\Gamma_0 - (1-\lambda)\Gamma_1)$ for some quantum channels $\Gamma_0, \Gamma_1 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$, $\lambda \in [0,1]$, and $\alpha \in \mathbb{R}$.*

Proof. 3 $\Rightarrow$ 1 is immediate. For 1 $\Rightarrow$ 2, suppose $\Psi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is Hermiticity preserving and trace-scaling, and let $\beta \in \mathbb{R}$ be the parameter for which $\mathrm{Tr}(\Psi(X)) = \beta\mathrm{Tr}(X)$ for all $X \in \mathrm{L}(\mathcal{X})$. As Ψ is Hermiticity preserving, it can be written as $\Psi = \Psi_0 - \Psi_1$, with $\Psi_0, \Psi_1 \in \mathrm{CP}(\mathcal{X}, \mathcal{Y})$ completely positive. Set $Q = \|\Psi_0^*(\mathbb{1}_\mathcal{Y})\|\mathbb{1}_\mathcal{X} - \Psi_0^*(\mathbb{1}_\mathcal{Y}) \geq 0$. Fixing any density matrix $\sigma \in \mathrm{D}(\mathcal{Y})$, observe that

$$\mathrm{Tr}(\Psi_0(X) + \mathrm{Tr}(QX)\sigma) = \|\Psi_0^*(\mathbb{1}_\mathcal{Y})\|\mathrm{Tr}(X) \tag{9.4}$$

and

$$\operatorname{Tr}(\Psi_1(X) + \operatorname{Tr}(QX)\sigma) = \operatorname{Tr}(\Psi_1(X) - \Psi_0(X)) + \|\Psi^*(\mathbb{1}_{\mathcal{Y}})\|\operatorname{Tr}(X) \tag{9.5}$$
$$= (\|\Psi^*(\mathbb{1}_{\mathcal{Y}})\| - \beta)\operatorname{Tr}(X). \tag{9.6}$$

Hence, defining $\Psi_0'(X) = \Psi_0(X) + \operatorname{Tr}(QX)\sigma$ and $\Psi_1'(X) = \Psi_1(X) + \operatorname{Tr}(QX)\sigma$, we have that $\Psi = \Psi_0' - \Psi_1'$ with both Ψ_0' and Ψ_1' being completely positive and trace-scaling.

Finally, for $2 \Rightarrow 3$, if either $\Psi_0 = 0$ or $\Psi_1 = 0$, the claim is immediate, so assume both are non-zero. As they are non-zero, completely positive, and trace-scaling, there exists $\alpha_0, \alpha_1 > 0$ for which $\Gamma_0 = \frac{1}{\alpha_0}\Psi_0$ and $\Gamma_1 = \frac{1}{\alpha_1}\Psi_1$ are quantum channels. Setting $\alpha = \alpha_0 + \alpha_1$ and $\lambda = \frac{\alpha_0}{\alpha_0+\alpha_1}$, we see that

$$\Psi = \Psi_0 - \Psi_1 = \alpha_0\Gamma_0 - \alpha_1\Gamma_1 = \alpha(\lambda\Gamma_0 - (1-\lambda)\Gamma_1) \tag{9.7}$$

as required. □

9.1.1 A naive approach

A naive approach to bounding $\left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_{1,H}$ in terms of $\|\Phi\|_{1,H}$ is to bound $\|\Phi\|_1$ in terms of $\|\Phi\|_{1,H}$. For example, if $\|\Phi\|_1 \leq c\|\Phi\|_{1,H}$ for $c \geq 1$, then by Theorem 8.1 we have the bound

$$\left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_{1,H} \leq \left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_1 \leq k\|\Phi\|_1 \leq ck\|\Phi\|_{1,H}. \tag{9.8}$$

The value of c may depend on additional assumptions on Φ. For example, if Φ is positive, then $\|\Phi\|_1 = \|\Phi\|_{1,H}$, and so in this case we arrive at the bound

$$\left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_{1,H} \leq k\|\Phi\|_{1,H}. \tag{9.9}$$

which, by Theorem 8.1, is achieved if and only if (up to a scalar) Φ acts as the transpose followed by a reversible quantum channel.

By naively modifying the proof of the inequality in Theorem 8.1, we may arrive at a better naive bound of $\left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_{1,H}$ in terms of $\|\Phi\|_{1,H}$.

Proposition 9.2. *For $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ and $c \geq 1$, if $\|\Phi\|_1 \leq c\|\Phi\|_{1,H}$, then it holds that*

$$\left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\right\|_{1,H} \leq (c(k-1)+1)\|\Phi\|_{1,H}. \tag{9.10}$$

Proof. Let $u \in \mathcal{X} \otimes \mathbb{C}^k$ be a unit vector, and let $u = \sum_{i=1}^{r} \alpha_i u_i \otimes e_i$ be a Schmidt-decomposition. We have

$$\begin{aligned}
\left\|(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)})(uu^*)\right\|_1 &= \left\| \sum_{i,j=1}^{r} \alpha_i \alpha_j \Phi(u_i u_j^*) \otimes E_{i,j} \right\|_1 && (9.11)\\
&\leq \sum_{i=1}^{r} \alpha_i^2 \|\Phi(u_i u_i^*)\|_1 + \sum_{1 \leq i \neq j \leq r} \alpha_i \alpha_j \left\|\Phi(u_i u_j^*)\right\|_1 && (9.12)\\
&\leq \|\Phi\|_{1,H} + \sum_{1 \leq i \neq j \leq r} \alpha_i \alpha_j \left\|\Phi(u_i u_j^*)\right\|_1 && (9.13)\\
&\leq \|\Phi\|_{1,H} \left(1 + c \sum_{1 \leq i \neq j \leq r} \alpha_i \alpha_j \right) && (9.14)\\
&\leq \|\Phi\|_{1,H}(1 + c\|J_r - \mathbb{1}_r\|) && (9.15)\\
&= (c(k-1)+1)\|\Phi\|_{1,H} && (9.16)
\end{aligned}$$

where J_r is the matrix of all ones, and we have used that

$$\|J_r - \mathbb{1}_r\| = r - 1 \leq k - 1. \tag{9.17}$$

□

Given the above proposition, it is natural then to try to find the the optimal c for various classes of maps in $\mathrm{T}(\mathcal{X}, \mathcal{Y})$.

9.1.2 The maximum gap between $\|\Phi\|_1$ and $\|\Phi\|_{1,H}$ for arbitrary Φ

It is straightforward to see that generally, for $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ it holds that $\|\Phi\|_1 \leq 2\|\Phi\|_{1,H}$. To see this, for $X \in \mathrm{L}(\mathcal{X})$ we have

$$\begin{aligned}
\|\Phi(X)\|_1 &= \frac{1}{2}\|\Phi(X + X^*) + \Phi(X - X^*)\|_1 && (9.18)\\
&\leq \frac{1}{2}(\|\Phi(X + X^*)\|_1 + \|\Phi(X - X^*)\|_1) && (9.19)\\
&\leq \frac{1}{2}\|\Phi\|_{1,H}(\|X + X^*\|_1 + \|X - X^*\|_1) && (9.20)\\
&\leq 2\|\Phi\|_{1,H}\|X\|_1. && (9.21)
\end{aligned}$$

As such, $\|\Phi\|_1 \leq 2\|\Phi\|_{1,H}$.

This bound may actually be viewed as a generalization of the fact that, for a matrix $A \in \mathrm{L}(\mathcal{X})$, $\|A\| \leq 2w(A)$, where $w(A)$ is the *numerical radius* of A, defined as

$$w(A) = \max\{|\langle u, Au\rangle| : u \in \mathcal{X}, \|u\| = 1\}. \tag{9.22}$$

Explicitly, for the linear functional $f : \mathrm{L}(\mathcal{X}) \to \mathbb{C}$ defined as $f(X) = \langle A, X\rangle$, it holds that

$$\begin{aligned} \|f\|_1 &= \max\{|f(X)| : X \in \mathrm{L}(\mathcal{X}), \|X\|_1 = 1\} && (9.23)\\ &= \max\{|\langle A, X\rangle| : X \in \mathrm{L}(\mathcal{X}), \|X\|_1 = 1\} && (9.24)\\ &= \|A\|, && (9.25) \end{aligned}$$

and

$$\begin{aligned} \|f\|_{1,H} &= \max\{|f(uu^*)| : u \in \mathcal{X}, \|u\| = 1\} && (9.26)\\ &= \max\{|\langle A, uu^*\rangle| : u \in \mathcal{X}, \|u\| = 1\} && (9.27)\\ &= w(A). && (9.28) \end{aligned}$$

The following example [67] shows that, without restricting to some subset of $\mathrm{T}(\mathcal{X}, \mathcal{Y})$, the bound $\|\Phi\|_1 \leq 2\|\Phi\|_{1,H}$ cannot be improved.

Example 9.3. The elementary matrix

$$E_{01} = e_0 e_1^* = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathrm{L}(\mathbb{C}^2). \tag{9.29}$$

satisfies $\|E_{01}\| = 1 = 2w(E_{01})$. Hence, the linear map $f : \mathrm{L}(\mathbb{C}^2) \to \mathbb{C}$ defined as $f(X) = \langle E_{01}, X\rangle$ for all $X \in \mathrm{L}(\mathbb{C}^2)$ satisfies $\|f\|_1 = 2\|f\|_{1,H}$.

To see this, first note that $\|E_{01}\| = 1$ is immediate, so we need only show that $w(E_{01}) = \frac{1}{2}$. For $u \in \mathbb{C}^2$ a unit vector, we have

$$|\langle u, E_{01}u\rangle| = |\langle e_0, u\rangle||\langle e_1, u\rangle| \leq \frac{|\langle e_0, u\rangle|^2 + |\langle e_1, u\rangle|^2}{2} = \frac{1}{2} \tag{9.30}$$

where we have applied the arithmetic geometric mean inequality. Equality holds if and only if $|\langle e_0, u\rangle|^2 = |\langle e_1, u\rangle|^2 = \frac{1}{2}$, and hence $w(E_{01}) = \frac{1}{2}$.

We conjecture that the above example is essentially unique in the following sense.

Conjecture 9.4. *Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be linear. If $\|\Phi\|_1 = 2\|\Phi\|_{1,H}$, then for any* *unit vectors $u, v \in \mathcal{X}$ for which $\|\Phi(uv^*)\|_1 = \|\Phi\|_1$ (at least one such pair must exist), it holds that:*

- *u and v are orthogonal.*
- *$\Phi(vu^*) = \Phi(uu^*) = \Phi(vv^*) = 0$.*

We believe the above conjecture is true, and the following theorem verifies most of its conclusions.

Theorem 9.5. *Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be linear. If $\|\Phi\|_1 = 2\|\Phi\|_{1,H}$, then for any unit vectors $u, v \in \mathbb{C}^n$ for which $\|\Phi(uv^*)\|_1 = \|\Phi\|_1$ (at least one such pair must exist), it holds that:*

- *u and v are orthogonal.*
- *$\Phi(vu^*) = \Phi(uu + vv^*) = 0$.*

Thus, assuming the conjecture is true, completing its proof only requires proving that one (and hence all) of the matrices

$$\Phi(uu^*), \Phi(vv^*), \text{ or } \Phi(uu^* - vv^*) \tag{9.31}$$

is zero. The proof of Theorem 9.5, along with the known completion of Conjecture 9.4 for the special case when $\mathcal{Y} = \mathbb{C}$, may be found in Appendix C.

Whether Conjecture 9.4 is true or not, Theorem 9.5 already implies that the naive approach in Proposition 9.2 cannot yield optimal bounds of

$$\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_{1,H} \tag{9.32}$$

in terms of $\|\Phi\|_{1,H}$. Specifically, we applied the bounds $\|\Phi(E_{ii})\|_1 \leq \|\Phi\|_{1,H}$ for all i and $\|\Phi(E_{ij})\|_1 \leq 2\|\Phi\|_{1,H}$ for all $i \neq j$. However, if $\|\Phi(E_{ij})\|_1 = 2\|\Phi\|_{1,H}$ then Theorem 9.5 already implies that $\|\Phi(E_{ji})\|_1 = 0$, and if the conjecture is true we will also be able to conclude that $\|\Phi(E_{ii})\|_1 = \|\Phi(E_{jj})\|_1 = 0$. Hence, finding an optimal (achievable) bound for $\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_{1,H}$ in terms of $\|\Phi\|_{1,H}$ will require somehow simultaneously bounding the norms of the matrices $\Phi(E_{ii}), \Phi(E_{jj}), \Phi(E_{ij})$, and $\Phi(E_{ji})$.

9.1.3 The maximum gap between $\|\Phi\|_1$ and $\|\Phi\|_{1,H}$ for Hermiticity preserving Φ

We now restrict attention to Hermiticity preserving maps. The map in the following example is a modified version of a map taken from Proposition 2 in [67].

Example 9.6. We will construct a Hermiticity preserving map $\Phi \in \mathrm{T}(\mathbb{C}^2)$ for which $\|\Phi\|_1 = \sqrt{2}\|\Phi\|_{1,H}$.

Let

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \text{and } Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{9.33}$$

denote the Pauli matrices. Define $\Phi \in \mathrm{T}(\mathbb{C}^2)$ to act on all $A \in \mathrm{L}(\mathbb{C}^2)$ as

$$\Phi(A) = \begin{pmatrix} \langle X, A\rangle & 0 \\ 0 & \langle Y, A\rangle \end{pmatrix}. \tag{9.34}$$

For any Hermitian matrix $H \in \mathrm{Herm}(\mathbb{C}^2)$, $\langle X, H\rangle$ and $\langle Y, H\rangle$ are both real-valued, and so Φ is Hermiticity preserving. We claim that

$$\|\Phi\|_1 = \sqrt{2}\|\Phi\|_{1,H} = 2. \tag{9.35}$$

For any $A \in \mathrm{L}(\mathbb{C}^2)$, we have $\|\Phi(A)\|_1 = |\langle X, A\rangle| + |\langle Y, A\rangle| \leq 2\|A\|_1$, and hence $\|\Phi\|_1 \leq 2$. To see that $\|\Phi\|_1 = 2$, consider the matrix

$$\frac{X + iY}{2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{9.36}$$

satisfying $\left\|\frac{1}{2}(X + iY)\right\|_1 = 1$. It holds that

$$\left\|\frac{1}{2}\Phi(X + iY)\right\|_1 = \left\|\frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & 2i \end{pmatrix}\right\|_1 = 2, \tag{9.37}$$

and hence $\|\Phi\|_1 = 2$.

Next any density matrix $\rho \in \mathrm{D}(\mathbb{C}^2)$ can be written as

$$\rho = \frac{1}{2}(\mathbb{1}_2 + r_1 X + r_2 Y + r_3 Z), \tag{9.38}$$

for $r = (r_1, r_2, r_3) \in \mathbb{R}^3$ with $\|r\| \leq 1$. Thus,

$$\|\Phi(\rho)\|_1 = |r_1| + |r_2| \leq \sqrt{2}, \tag{9.39}$$

with equality if and only if $r = \sqrt{\frac{1}{2}}(1, 1, 0)$. Hence,

$$\|\Phi\|_{1,H} = \max\{\|\Phi(\rho)\|_1 : \rho \in \mathrm{D}(\mathbb{C}^2)\} = \sqrt{2}. \tag{9.40}$$

We conjecture that $\sqrt{2}$ is the optimal value.

Conjecture 9.7. *For $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ linear and Hermiticity preserving, it holds that $\|\Phi\|_1 \leq \sqrt{2}\|\Phi\|_{1,H}$.*

Equivalently, as a linear map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is Hermiticity preserving if and only if Φ^* is, it holds that $\|\Phi\| \leq \sqrt{2}\|\Phi\|_H$.

It also seems natural to further conjecture that the phenomenon appearing in Example 9.6 is in some sense unique.

For the remainder of this section we present a potential path towards proving the conjecture. The approach taken is based on the intuition/guess that proving this will have something to do with matrix versions of the identity $|a + ib| = \sqrt{a^2 + b^2}$ for $a, b \in \mathbb{R}$. We give one in terms of the trace norm and one in terms of the operator norm. In both propositions, we make use of the fact that, for Hermitian H, K and $X = H + iK$, it holds that

$$XX^* + X^*X = 2(H^2 + K^2). \tag{9.41}$$

We thank Vern Paulsen for suggesting this to us.

Proposition 9.8. *For Hermitian $H, K \in \mathrm{Herm}(\mathcal{X})$, it holds that*

$$\sqrt{\frac{1}{2}}\mathrm{Tr}(\sqrt{H^2 + K^2}) \leq \|H + iK\|_1 \leq \mathrm{Tr}\big(\sqrt{H^2 + K^2}\big). \tag{9.42}$$

Additionally:

- *Equality holds in the first inequality if and only if $H^2 = K^2$ and*

$$HK = -KH \tag{9.43}$$

(i.e. $H + iK$ is nilpotent).

- *Equality holds in the second inequality if and only if H and K commute (i.e. $H + iK$ is normal).*

Proof. For the first inequality, we have

$$\mathrm{Tr}\big(\sqrt{H^2 + K^2}\big) = \left\| \begin{pmatrix} H \\ K \end{pmatrix} \right\|_1 = \sqrt{\frac{1}{2}} \left\| \begin{pmatrix} \mathbb{1}_n & i\mathbb{1}_n \\ \mathbb{1}_n & -i\mathbb{1}_n \end{pmatrix} \begin{pmatrix} H \\ K \end{pmatrix} \right\|_1 \tag{9.44}$$

$$= \sqrt{\frac{1}{2}} \left\| \begin{pmatrix} H + iK \\ H - iK \end{pmatrix} \right\|_1 \tag{9.45}$$

$$\leq \sqrt{\frac{1}{2}} (\|H + iK\|_1 + \|H - iK\|_1) \tag{9.46}$$

$$= \sqrt{2} \|H + iK\|_1, \tag{9.47}$$

where the inequality is the triangle inequality, the first equality is unitary invariance of the norm, and the final equality is $\|H+iK\|_1 = \|H-iK\|_1$. Equality holds if and only if equality holds in the triangle-inequality. Proposition 6.4 gives that this holds if and only if

$$(H + iK)(H - iK)^* = (H + iK)^2 = H^2 - K^2 + i(HK + KH) = 0, \tag{9.48}$$

which is equivalent to both $H^2 - K^2 = 0$ and $HK + KH = 0$, as required.

For the second inequality, denote $X = H + iK$ and observe that

$$X^*X + XX^* = 2(H^2 + K^2). \tag{9.49}$$

We have

$$\|H + iK\|_1 = \frac{1}{2}\mathrm{Tr}\big(\sqrt{X^*X}\big) + \frac{1}{2}\mathrm{Tr}\big(\sqrt{XX^*}\big) \tag{9.50}$$

$$\leq \mathrm{Tr}\left(\sqrt{\frac{X^*X + XX^*}{2}}\right) \tag{9.51}$$

$$= \mathrm{Tr}\big(\sqrt{H^2 + K^2}\big), \tag{9.52}$$

where the inequality follows from the operator concavity of $P \mapsto \mathrm{Tr}\big(\sqrt{P}\big)$. In fact, as this function is strictly operator concave (see Theorem 2.9 in [6]), equality holds if and only if $X^*X = XX^*$, i.e. $X = H + iK$ is normal, which is equivalent to H and K commuting. □

A similar proposition may be proved for the operator norm.

Proposition 9.9. *For Hermitian $H, K \in \mathrm{Herm}(\mathcal{X})$, it holds that*

$$\sqrt{\|H^2 + K^2\|} \leq \|H + iK\| \leq \sqrt{2}\sqrt{\|H^2 + K^2\|}. \tag{9.53}$$

Furthermore, equality can hold in both inequalities.

Proof. For the first inequality, as in the previous proposition, denote

$$X = H + iK \tag{9.54}$$

and observe $XX^* + X^*X = 2(H^2 + K^2)$. Hence

$$\begin{aligned} \|H + iK\|^2 &= \frac{1}{2}(\|XX^*\| + \|X^*X\|) && (9.55) \\ &\geq \frac{1}{2}\|XX^* + X^*X\| && (9.56) \\ &= \|H^2 + K^2\|, && (9.57) \end{aligned}$$

where in the first inequality we have used that $\|X\|^2 = \|X^*X\| = \|XX^*\|$, and the inequality is the triangle inequality. An example of matrices giving equality are $H = K = \mathbb{1}_{\mathcal{X}}$.

Similarly, for the second inequality we have

$$\|H + iK\|^2 = \|XX^*\| \leq \|XX^* + X^*X\| = 2\|H^2 + K^2\|. \tag{9.58}$$

An example of equality in this case is

$$H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \text{ and } K = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{9.59}$$

□

We now state another conjecture which, when coupled with either of the above propositions, would result in a proof of Conjecture 9.7.

Conjecture 9.10. *For $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ Hermiticity preserving, and Hermitian $H, K \in \mathrm{Herm}(\mathcal{X})$, it holds that*

$$\left\| \begin{pmatrix} \Phi(H) \\ \Phi(K) \end{pmatrix} \right\|_1 \leq \|\Phi\|_{1,H} \left\| \begin{pmatrix} H \\ K \end{pmatrix} \right\|_1. \tag{9.60}$$

Similarly, we also conjecture that the corresponding statement in terms of the *operator norm holds:*

$$\left\| \begin{pmatrix} \Phi(H) \\ \Phi(K) \end{pmatrix} \right\| \leq \|\Phi\|_H \left\| \begin{pmatrix} H \\ K \end{pmatrix} \right\|. \tag{9.61}$$

We note that the above conjecture may be viewed as a special case of the following more general conjecture.

Conjecture 9.11. *Let* $\Phi : \mathrm{Herm}(\mathcal{X}) \to \mathrm{Herm}(\mathcal{Y})$ *be a real linear map. For any* $(H_{ij}) \in M_n(\mathrm{Herm}(\mathcal{X}))$ *($n \times n$ block matrices whose entries are Hermitian matrices in* $\mathrm{Herm}(\mathcal{X})$*), it holds that*

$$\|(\Phi(H_{ij}))\|_1 \leq \|\Phi\|_{1,H} \|(H_{ij})\|_1. \tag{9.62}$$

We also conjecture the same inequality in the operator norm:

$$\|(\Phi(H_{ij}))\| \leq \|\Phi\|_H \|(H_{ij})\|. \tag{9.63}$$

We note that the above two conjectures hold for matrix transposition, which perhaps lends some evidence to their veracity. For example, in the case of Conjecture 9.10, given Hermitian $H, K \in \mathrm{Herm}(\mathcal{X})$, we have

$$\left\| \begin{pmatrix} H^{\mathsf{T}} \\ K^{\mathsf{T}} \end{pmatrix} \right\|_1 = \left\| \overline{\begin{pmatrix} H^{\mathsf{T}} \\ K^{\mathsf{T}} \end{pmatrix}} \right\|_1 = \left\| \begin{pmatrix} H^* \\ K^* \end{pmatrix} \right\|_1 = \left\| \begin{pmatrix} H \\ K \end{pmatrix} \right\|_1. \tag{9.64}$$

In fact, the above argument works for any matrix norm that remains unchanged under complex conjugation (e.g. unitarily invariant norms). This argument also clearly applies to any element of $M_n(\mathrm{Herm}(\mathcal{X}))$, and so when viewed as a real linear map on $\mathrm{Herm}(\mathcal{X})$, transposition is a complete isometry (with respect to any matrix norm invariant under complex conjugation).

Proof of Conjecture 9.7 assuming Conjecture 9.10. Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be Hermiticity preserving, and let $H, K \in \mathrm{Herm}(\mathcal{X})$ be arbitrary. It holds that

$$\begin{aligned}
\|\Phi(H+iK)\|_1 &= \|\Phi(H)+i\Phi(K)\|_1 && (9.65)\\
&\leq \mathrm{Tr}(\sqrt{\Phi(H)^2+\Phi(K)^2}) && (9.66)\\
&= \left\| \begin{pmatrix} \Phi(H) \\ \Phi(K) \end{pmatrix} \right\|_1 && (9.67)\\
&\leq \|\Phi\|_{1,H} \left\| \begin{pmatrix} H \\ K \end{pmatrix} \right\|_1 && (9.68)\\
&= \|\Phi\|_{1,H} \mathrm{Tr}(\sqrt{H^2+K^2}) && (9.69)\\
&\leq \|\Phi\|_{1,H} \sqrt{2} \|H+iK\|_1. && (9.70)
\end{aligned}$$

The first inequality is by Φ being Hermiticity preserving and Proposition 9.8, the second is by assumption of Conjecture 9.7, and the last is again by Proposition 9.8.

Hence, if Conjecture 9.7 holds, we arrive at $\|\Phi\|_1 \leq \sqrt{2}\|\Phi\|_{1,H}$. □

A similar proof is possible in terms of the operator norm using Proposition 9.9. It is interesting to compare properties of Example 9.6 with equality conditions for the inequalities in the above proof. Observe the following facts:

- Assuming the above proof is valid, equality holds in the last inequality if and only if $H^2 = K^2$ and H and K anti-commute. Note that this is exactly the case for the optimal choices of H and K in Example 9.6.
- Similarly, equality holds in the first inequality if and only if $\Phi(H)$ and $\Phi(K)$ commute, which again is the case in Example 9.6.

Thus, Example 9.6 and the above approach to proving Conjecture 9.7 are consistent, and moreover suggests that Example 9.6 is perhaps unique in some sense.

It is interesting that there are fairly natural paths forward in either the trace or operator norm. Any proof which may be possible in the operator norm may lead to a proof for arbitrary C*-algebras, though for finite dimensions, it seems that the trace-norm more easily leads to constraints on the form of any map achieving the extremal norm separations.

9.2 The structure of states with near maximal negativity

In this section we examine what is required to prove a robust version of Theorem 5.2 in the case of density matrices, and prove partial results in this direction. For a density matrix $\rho \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^m)$ with $\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(\rho)\|_1 = n$, the proof of Theorem 5.2 can roughly be broken into two steps:

1. Establishing that, for any decomposition $\rho = \sum_{i=1}^r p_i u_i u_i^*$ into pure states, all u_i are necessarily maximally entangled.

2. Given a set of Hilbert-Schmidt orthogonal isometries $\{U_i\}_{i=1}^r$ (which in this context come from a spectral decomposition of ρ), if every element of $\mathrm{span}\{U_i\}$ is proportional to an isometry, then the U_i have orthogonal ranges.

Thus, some robust versions of the above two arguments are likely required for proving a robust version of Theorem 5.2.[1]

Here we prove a couple of robustness results related to point 1 above. Logically, the proof of this point ultimately begins with the fact that, for a matrix $A \in \mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$ for $n \leq m$, it holds that $\|A\|_1 \leq \sqrt{n}\|A\|_2$, with equality if and only if A is a scalar multiple of an isometry. This inequality and its equality condition hold for a simple reason: there exists an isometry $U \in \mathrm{U}(\mathbb{C}^n, \mathbb{C}^m)$ for which $\|A\|_1 = \langle U, A\rangle$, and as such we may apply Cauchy-Schwarz to conclude that

$$\|A\|_1 = \langle U, A\rangle \leq \|U\|_2\|A\|_2 = \sqrt{n}\|A\|_2, \tag{9.71}$$

with equality if and only if A and U are linearly dependent, i.e. if and only if A is a scalar multiple of an isometry.

Thus, it is natural to begin by observing a robust version of the Cauchy-Schwarz inequality: For vectors $u, v \in \mathbb{C}^n$ and $\epsilon \in [0, 1]$, it holds that

$$|\langle u, v\rangle| \geq \|u\|\|v\|\sqrt{1-\epsilon} \tag{9.72}$$

[1] A robust version of this theorem will likely be similar in character to robust versions of "self-testing" (see, e.g. [44, 52, 51]). In this literature, the goal seems to be roughly to prove that the state winning a particular game with optimal probability is essentially unique, and furthermore, any state that wins with near optimal probability is necessarily close to the optimal state.

if and only if $\|P_u v - v\|^2 \leq \epsilon \|v\|^2$, where $P_u = \frac{1}{\|u\|^2} uu^*$, i.e. the orthogonal projection onto $\mathbb{C}u$. Informally, u and v almost satisfy Cauchy-Schwarz with equality if and only if they are almost linearly dependent. This may be seen directly from the proof of Cauchy-Schwarz, which starts with the observation that

$$\|v\|^2 = \|P_u v\|^2 + \|(\mathbb{1}_n - P_u)v\|^2 = \frac{1}{\|u\|^2} |\langle u, v\rangle|^2 + \|P_u v - v\|^2. \tag{9.73}$$

Rearranging this equality gives the desired approximate version of Cauchy-Schwarz.

This simple fact may be translated into a trace-norm condition for a matrix to be close to an isometry.

Proposition 9.12. *Let $n \leq m$, $A \in \mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$ with $\|A\|_2 = 1$, and let $\epsilon \in [0, 1]$. If $\|A\|_1 \geq \sqrt{n}\sqrt{1-\epsilon}$, then there exists an isometry $V \in \mathrm{U}(\mathbb{C}^n, \mathbb{C}^m)$ for which*

$$\left\| \sqrt{\frac{1}{n}} V - A \right\|_2 \leq 2\sqrt{\epsilon}. \tag{9.74}$$

Proof. Let $V \in \mathrm{U}(\mathbb{C}^n, \mathbb{C}^m)$ be an isometry for which

$$\langle V, A\rangle = \|A\|_1 \geq \sqrt{n}\sqrt{1-\epsilon} = \|V\|_2 \|A\|_2 \sqrt{1-\epsilon}. \tag{9.75}$$

Hence, these matrices satisfy approximate equality in Cauchy-Schwarz, and so

$$\left\| \frac{1}{n} \langle V, A\rangle V - A \right\|_2 \leq \sqrt{\epsilon}. \tag{9.76}$$

It follows that

$$\left\| \sqrt{\frac{1}{n}} V - A \right\|_2 \leq \left\| \sqrt{\frac{1}{n}} V - \frac{1}{n}\langle V, A\rangle V \right\|_2 + \left\| \frac{1}{n} \langle V, A\rangle V - A \right\|_2 \tag{9.77}$$

$$= (1 - \langle V, A\rangle / \sqrt{n}) + \left\| \frac{1}{n} \langle V, A\rangle V - A \right\|_2 \tag{9.78}$$

$$\leq 1 - \sqrt{1-\epsilon} + \sqrt{\epsilon} \tag{9.79}$$

$$\leq 2\sqrt{\epsilon} \tag{9.80}$$

where the equality holds as $\langle V, A\rangle \leq \|V\|_2 \|A\|_2 = \sqrt{n}$, the second inequality holds by $\langle V, A\rangle \geq \sqrt{n}\sqrt{1-\epsilon}$, and the last inequality holds as $1 - \sqrt{\epsilon} \leq \sqrt{1-\epsilon}$

for $\epsilon \in [0,1]$.[2] $\square$

We may apply the previous proposition to prove a statement about purifications of a density matrix with near maximal negativity.

Proposition 9.13. *Let $n \leq m$, and $\rho \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^m)$ be a density matrix. If*

$$\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(\rho)\|_1 \geq n(1-\epsilon), \tag{9.81}$$

then for any purification $u \in \mathbb{C}^n \otimes \mathbb{C}^m \otimes \mathbb{C}^r$ of ρ, there exists a unit vector $v \in \mathbb{C}^n \otimes \mathbb{C}^m \otimes \mathbb{C}^r$ that is maximally entangled between $\mathbb{C}^n$ and $\mathbb{C}^m \otimes \mathbb{C}^r$ for which $\|u - v\| \leq 2\sqrt{\epsilon}$.

Proof. It holds that

$$\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^r)})(uu^*)\|_1 \geq \|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(\rho)\|_1 \geq n(1-\epsilon). \tag{9.82}$$

Let $A \in \mathrm{L}(\mathbb{C}^n, \mathbb{C}^m \otimes \mathbb{C}^r)$ be such that $u = \mathrm{vec}(A^\intercal)$. By Proposition 4.1, it holds that

$$\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^r)})(uu^*)\|_1 = \|A\|_1^2. \tag{9.83}$$

Thus, it holds that $\|A\|_1 \geq \sqrt{n}\sqrt{1-\epsilon}$, and so Proposition 9.12 implies that there exists an isometry $V \in \mathrm{U}(\mathbb{C}^n, \mathbb{C}^m \otimes \mathbb{C}^r)$ for which $\|V/\sqrt{n} - A\|_2 \leq 2\sqrt{\epsilon}$. Hence, setting $v = \sqrt{\frac{1}{n}}\mathrm{vec}(V^\intercal)$, v is maximally entangled between $\mathbb{C}^n$ and $\mathbb{C}^m \otimes \mathbb{C}^r$, and we see that

$$\|v - u\| = \|V/\sqrt{n} - A\|_2 \leq 2\sqrt{\epsilon}, \tag{9.84}$$

as required. $\square$

Next, we prove a direct robust generalization of point 1 at the beginning of this section. In particular, point 1 says that if a density matrix has maximal negativity, then given any pure-state decomposition, all of the pure states must themselves be maximally entangled. Here we show that if a density matrix almost has maximal negativity, then given any pure state decomposition, the pure states in the decomposition are close to maximally entangled *on average*, where the weightings are given by the weightings of the pure states in the decomposition of the density matrix.

[2] To see this, note that the inequality is equivalent to the inequality $1 - 2\sqrt{\epsilon} + \epsilon \leq 1 - \epsilon$ (which can be arrived at by squaring both sides), and upon rearranging is equivalent to $\epsilon \leq \sqrt{\epsilon}$, which holds for $\epsilon \in [0,1]$.

In the proof we use that for $A, B \in \mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$

$$\|\mathrm{vec}(A)\mathrm{vec}(A)^* - \mathrm{vec}(B)\mathrm{vec}(B)^*\|_2 = \|A \otimes A^* - B \otimes B^*\|_2. \tag{9.85}$$

One way to see the above equality is that the entries of the matrices appearing on both sides of the equality are necessarily rearrangements of each other.

Proposition 9.14. *Let $n \leq m$ be positive integers, $\rho \in \mathrm{D}(\mathbb{C}^n \otimes \mathbb{C}^m)$ be a density matrix, and let $\epsilon \in [0,1]$. If it holds that*

$$\|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(\rho)\|_1 \geq n(1-\epsilon), \tag{9.86}$$

then for any decomposition $\rho = \sum_{i=1}^r p_i u_i u_i^*$, where $\{u_i\}_{i=1}^r \subset \mathbb{C}^n \otimes \mathbb{C}^m$ is a set of unit vectors, there exists a set of a maximally entangled unit vectors $\{v_i\}_{i=1}^r \subset \mathbb{C}^n \otimes \mathbb{C}^m$ for which it holds that

$$\sum_{i=1}^r p_i \|u_i u_i^* - v_i v_i^*\|_1^2 = 2\sum_{i=1}^r p_i \|u_i u_i^* - v_i v_i^*\|_2^2 \leq 16\epsilon. \tag{9.87}$$

Proof. First, note that, for any pair of unit vectors x and y, it holds that

$$\|xx^* - yy^*\|_1 = \sqrt{2}\|xx^* - yy^*\|_2, \tag{9.88}$$

and as such the equality in the conclusion is immediate, so we only need to prove the inequality.

Let $\rho = \sum_{i=1}^r p_i u_i u_i^*$ be a pure state decomposition, and let $u_i = \mathrm{vec}(A_i^{\intercal})$ for $A_i \in \mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$. It holds that

$$\begin{aligned} n(1-\epsilon) &\leq \|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(\rho)\|_1 && (9.89)\\ &\leq \sum_{i=1}^r p_i \|(T_n \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)})(u_i u_i^*)\|_1 && (9.90)\\ &= \sum_{i=1}^r p_i \|A_i\|_1^2 && (9.91)\\ &\leq n && (9.92)\end{aligned}$$

For each $1 \leq i \leq r$, let $V_i \in \mathrm{U}(\mathbb{C}^n, \mathbb{C}^m)$ be an isometry for which

$$\langle V_i, A_i \rangle = \|A_i\|_1. \tag{9.93}$$

Defining $V, A \in \mathrm{L}(\mathbb{C}^n \otimes \mathbb{C}^m \otimes \mathbb{C}^r, \mathbb{C}^m \otimes \mathbb{C}^n \otimes \mathbb{C}^r)$ as

$$V = \sum_{i=1}^{r} \sqrt{p_i} V_i \otimes V_i^* \otimes E_{ii}, \text{ and } A = \sum_{i=1}^{r} \sqrt{p_i} A_i \otimes A_i^* \otimes E_{ii}, \tag{9.94}$$

observe that

$$\langle V, A \rangle = \sum_{i=1}^{r} p_i \langle V_i \otimes V_i^*, A_i \otimes A_i^* \rangle \tag{9.95}$$

$$= \sum_{i=1}^{r} p_i \|A_i\|_1^2 \tag{9.96}$$

$$\geq n(1-\epsilon) \tag{9.97}$$

$$= n\sqrt{1-\delta}, \tag{9.98}$$

for $\delta = 2\epsilon - \epsilon^2 \in [0,1]$. Furthermore, noting that

$$\|V\|_2^2 = \sum_{i=1}^{r} p_i \|V_i\|_2^4 = n^2, \text{ and } \|A\|_2^2 = \sum_{i=1}^{r} p_i \|A_i\|_2^4 = 1, \tag{9.99}$$

we see that $\langle V, A \rangle \geq \|V\|_2 \|A\|_2 \sqrt{1-\delta}$, and hence

$$\left\| \frac{\langle V, A \rangle}{n^2} V - A \right\|_2^2 \leq \delta. \tag{9.100}$$

Thus, we have

$$\left\| \frac{1}{n} V - A \right\|_2 \leq \left\| \frac{1}{n} V - \frac{\langle V, A \rangle}{n^2} V \right\|_2 + \left\| \frac{\langle V, A \rangle}{n^2} V - A \right\|_2 \tag{9.101}$$

$$\leq 1 - \langle V, A \rangle / n + \sqrt{\delta} \tag{9.102}$$

$$\leq 1 - \sqrt{1-\delta} + \sqrt{\delta} \tag{9.103}$$

$$\leq 2\sqrt{\delta}, \tag{9.104}$$

where we have again used that $\sqrt{1-\delta} \geq 1 - \sqrt{\delta}$ for $\delta \in [0,1]$.

However, we may explicitly see that

$$\left\| \frac{1}{n} V - A \right\|_2^2 = \left\| \sum_{i=1}^{r} \sqrt{p_i} \Big(\frac{1}{n} V_i \otimes V_i^* - A_i \otimes A_i^* \Big) \otimes E_{ii} \right\|_2^2 \tag{9.105}$$

$$= \sum_{i=1}^{r} p_i \left\| \frac{1}{n} V_i \otimes V_i^* - A_i \otimes A_i^* \right\|_2^2 \tag{9.106}$$

$$= \sum_{i=1}^{r} p_i \| u_i u_i^* - v_i v_i^* \|_2^2, \tag{9.107}$$

where, for each i, $v_i = \sqrt{\frac{1}{n}} \mathrm{vec}(V_i^{\mathsf{T}})$ is maximally entangled. Hence, it holds that

$$\sum_{i=1}^{r} p_i \| u_i u_i^* - v_i v_i^* \|_2^2 \leq 4\delta \leq 8\epsilon, \tag{9.108}$$

as required. □

Finally, observe that, if so desired, we may eliminate the square in the conclusion of the above proposition using the convexity of $x \mapsto x^2$ to conclude that

$$\sum_{i=1}^{r} p_i \| u_i u_i^* - v_i v_i^* \|_1 = \sqrt{2} \sum_{i=1}^{r} p_i \| u_i u_i^* - v_i v_i^* \|_2 \leq 4\sqrt{\epsilon}. \tag{9.109}$$

Chapter 10

Conclusion

In this thesis we have considered questions regarding the usefulness of entanglement in single-shot quantum channel discrimination. Specifically, we have examined how much entanglement is in general necessary to achieve an optimal discrimination strategy, as well as how large the advantage provided by entanglement can be in principle. Due to the connection of this problem with the induced trace-norm and completely bounded trace-norm, these problems translate into questions about properties of these norms on various classes of linear maps.

Our results extend the literature on the well-known examples of matrix transposition and the Werner-Holevo channels. For example, matrix transposition has long been known to saturate the inequality

$$\|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_1 \leq k\|\Psi\|_1, \tag{10.1}$$

and we have proven that, at least for linear maps between matrix algebras, transposition is essentially unique in this property. By leveraging this we have proven uniqueness results for the Werner-Holevo channel game regarding the gap between entangled and unentangled performance. We have also defined a family of channel discrimination games that extend the Werner-Holevo channel game, where the output dimension can be arbitrarily small compared to the input dimension, but it is still necessary to use as much entanglement as is possible to perfectly discriminate the channels. These results depend on an understanding of the structure of maximally entangled states, as well as particular characterizations of complete trace-norm isometries and reversible quantum channels.

There are many natural questions that remain open. For example, establishing and characterizing the maximal gap between entangled and unentangled performance in single-shot quantum channel discrimination, and the corresponding mathematical problem of bounding $\|\Psi \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^k)}\|_{1,H}$ in terms of $\|\Psi\|_{1,H}$. As discussed in the previous chapter, this problem requires additional insights beyond simply adapting the proof of the bound in Equation (10.1). Furthermore, these questions motivate conjectures regarding completely bounded norm-type questions for real linear maps

$$\Psi : \mathrm{Herm}(\mathcal{X}) \to \mathrm{Herm}(\mathcal{Y}). \tag{10.2}$$

Our observations throughout this thesis also motivate the following conjecture:

Conjecture 10.1. *For any completely positive map $\Phi \in \mathrm{CP}(\mathcal{X}, \mathcal{Y})$ and integer $k \geq 1$, it holds that*

$$\|\Phi \otimes T_k\|_{1,H} = \|\Phi \otimes T_k\|_1. \tag{10.3}$$

We pose this conjecture for the following reasons:

- In Appendix B, numerical computations of both $\|\Psi_{n,k} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\|_1$ and $\|\Psi_{n,k} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\|_{1,H}$ appear to yield the same value. The map $\Psi_{n,k}$ is a completely positive map composed with transposition, and so these norms can be rewritten as a completely positive map tensored with the transpose.
- When $\mathcal{Y} = \mathcal{X}$ and $\Phi = \mathbb{1}_{\mathrm{L}(\mathcal{X})}$, the above conjecture holds by Proposition 4.2.
- For Φ a quantum channel, it holds that $\|\Phi \otimes T_k\|_{1,H} = k$ if and only if $\|\Phi \otimes T_k\|_1 = k$ (Theorem 6.12), and so in this extreme case the two norms are necessarily equal.

It is unclear how to prove such a result for arbitrary completely positive maps, or even for the restricted class of quantum channels.

Another clear direction for future work is to make some of the results in this thesis robust. That is, the characterizations in this thesis are exact, and so it is natural to seek approximate versions of these theorems. The simplest thing to try first is to characterize the states that almost have maximal entanglement negativity, and this is discussed in the previous chapter. This question could be considered relative to other entanglement measures as well. We could also try to characterize the structure of maps whose completely bounded trace-norm is almost maximal; i.e. characterize the maps $\Phi \in \mathrm{T}(\mathbb{C}^n, \mathcal{Y})$ satisfying $\|\Phi\|_1 = 1$ and $|||\Phi|||_1 \geq n(1-\epsilon)$ for $\epsilon \in [0,1]$. For small ϵ, are these maps necessarily close (in some distance measure) to the transpose followed by a complete trace-norm isometry?

Overall, the operational questions in single-shot quantum channel discrimination seem to motivate interesting questions regarding norms. In this thesis we have worked in finite dimensions and used the trace-norm and its variants, but any general results about these norms in this context could potentially have their proofs adapted to the operator norm and applied in the setting of general C*-algebras. In finite dimensions it is convenient that either norm can be used when approaching these questions. The trace-norm seems to be well-suited to establishing equality conditions for many of the inequalities considered, as it seems to contain information about the entire matrix, as opposed to only its largest singular value.

We hope that this thesis stimulates further research into the role of entanglement in single-shot quantum channel discrimination, and motivates new and natural questions on the well-studied norms arising in this context.

References

[1] A. Acín. Statistical distinguishability between unitary operations. *Physical Review Letters*, 87(17):177901, 2001.

[2] D. Aharonov, A. Kitaev, and N. Nisan. Quantum circuits with mixed states. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 20–30, 1998.

[3] I. Bjelakovic and R. Siegmund-Schultze. Quantum Stein's lemma revisited, inequalities for quantum entropies, and a concavity theorem of Lieb. Available as arXiv.org e-Print quant-ph/0307170, 2003.

[4] F. Brandao, M. Christandl, and J. Yard. Faithful squashed entanglement. *Communications in Mathematical Physics*, 306(3):805–830, 2011.

[5] P. Busch. Stochastic Isometries in Quantum Mechanics. *Mathematical Physics, Analysis and Geometry*, 2(1):83–106, 1999.

[6] E. Carlen. Trace inequalities and quantum entropy: an introductory course. In *Entropy and the quantum*, volume 529 of *Contemp. Math.*, pages 73–140. Amer. Math. Soc., Providence, RI, 2010.

[7] J.-T. Chan, C.-K. Li, and N.-S. Sze. Isometries for unitarily invariant norms. *Linear Algebra and its Applications*, 399:53–70, 2005.

[8] W.-S. Cheung, C.-K. Li, and Y.-T. Poon. Isometries between matrix algebras. *Journal of the Australian Mathematical Society*, 77(1):1–16, 2004.

[9] A. Childs, J. Preskill, and J. Renes. Quantum information and precision measurement. *Journal of Modern Optics*, 47(2–3):155–176, 2000.

[10] G. Chiribella, G. D'Ariano, and P. Perinotti. Memory effects in quantum channel discrimination. *Physical Review Letters*, 101(18):180501, 2008.

[11] M.-D. Choi. A Schwarz inequality for positive linear maps on C*-algebras. *Illinois Journal of Mathematics*, 18(4):565–574, 1974.

[12] M.-D. Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, 1975.

[13] M. Christandl and A. Winter. Squashed entanglement: An additive entanglement measure. *Journal of Mathematical Physics*, 45(3):829–840, 2004.

[14] I. L. Chuang and M. A. Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *Journal of Modern Optics*, 44(11-12):2455–2467, 1997.

[15] M. Donald, M. Horodecki, and O. Rudolph. The uniqueness theorem for entanglement measures. *Journal of Mathematical Physics*, 43(9):4252–4272, 2002.

[16] R. Duan, Y. Feng, and M. Ying. Local distinguishability of multipartite unitary operations. *Physical Review Letters*, 100(2):020503, 2008.

[17] R. Duan, Y. Feng, and M. Ying. Perfect distinguishability of quantum operations. *Physical Review Letters*, 103(21):210501, 2009.

[18] A. Gilchrist, N. Langford, and M. Nielsen. Distance measures to compare real and ideal quantum processes. *Physical Review A*, 71(6):062310, 2005.

[19] C. Granade, J. Combes, and D. G. Cory. Practical Bayesian tomography. *New Journal of Physics*, 18(3):033024, 2016.

[20] U. Haagerup. Injectivity and decomposition of completely bounded maps. In H. Araki, C. Moore, Ş.-V. Stratila, and D.-V. Voiculescu, editors, *Operator Algebras and their Connections with Topology and Ergodic Theory*, number 1132 in Lecture Notes in Mathematics, pages 170–222. Springer Berlin Heidelberg, 1985.

[21] A. Harrow, A. Hassidim, D. Leung, and J. Watrous. Adaptive versus non-adaptive strategies for quantum channel discrimination. *Physical Review A*, 81(3):032339, 2010.

[22] C. Helstrom. Detection theory and quantum mechanics. *Information and Control*, 10(3):254–291, 1967.

[23] F. Hiai and D. Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Communications in Mathematical Physics*, 143(1):99–114, 1991.

[24] A. Holevo. An analog of the theory of statistical decisions in noncommutative probability theory. *Transactions of the Moscow Mathematical Society*, 26:133–149, 1972.

[25] M. Horodecki, P. Horodecki, and R. Horodecki. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A*, 223(1):1–8, 1996.

[26] M. Horodecki, P. Horodecki, and R. Horodecki. Mixed-State Entanglement and Distillation: Is there a "Bound" Entanglement in Nature? *Physical Review Letters*, 80(24):5239–5242, 1998.

[27] A. Jenčová. Reversibility conditions for quantum operations. *Reviews in Mathematical Physics*, 24(07):1250016, 2012.

[28] A. Jenčová and M. Plávala. Conditions for optimal input states for discrimination of quantum channels. *Journal of Mathematical Physics*, 57(12):122203, 2016.

[29] L. Jian, K. He, Q. Yuan, and F. Wang. On Partially Trace Distance Preserving Maps and Reversible Quantum Channels, On Partially Trace Distance Preserving Maps and Reversible Quantum Channels. *Journal of Applied Mathematics, Journal of Applied Mathematics*, 2013, 2013:e474291, 2013.

[30] N. Johnston. How to compute hard-to-compute matrix norms [weblog post]. http://www.njohnston.ca/2016/01/, 2016.

[31] N. Johnston. QETLAB: A MATLAB toolbox for quantum entanglement, version 0.9. http://qetlab.com, 2016.

[32] N. Johnston and E. Patterson. The inverse eigenvalue problem for entanglement witnesses. *Linear Algebra and its Applications*, 550:1–27, 2018.

[33] R. Kadison. Isometries of Operator Algebras. *Annals of Mathematics*, 54(2):325–338, 1951.

[34] A. Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191–1249, 1997.

[35] A. Kitaev, A. Shen, and M. Vyalyi. *Classical and Quantum Computation*, volume 47 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002.

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

[37] C. Lancien, S. Di Martino, M. Huber, M. Piani, G. Adesso, and A. Winter. Should Entanglement Measures be Monogamous or Faithful? *Physical Review Letters*, 117(6):060501, 2016.

[38] R. H. Levene and R. M. Timoney. Completely bounded norms of right module maps. *Linear Algebra and its Applications*, 436(5):1406–1424, 2012.

[39] C.-K. Li and S. Pierce. Linear Preserver Problems. *The American Mathematical Monthly*, 108(7):591–605, 2001.

[40] C.-K. Li, Y.-T. Poon, and N.-S. Sze. Isometries for Ky Fan Norms between Matrix Spaces. *Proceedings of the American Mathematical Society*, 133(2):369–377, 2005.

[41] W. Matthews, M. Piani, and J. Watrous. Entanglement in channel discrimination with restricted measurements. *Physical Review A*, 82(3):032302, 2010.

[42] L. Molnár. Fidelity preserving maps on density operators. *Reports on Mathematical Physics*, 48(3):299–303, 2001.

[43] A. Müller-Hermes, D. Reeb, and M. M. Wolf. Positivity of linear maps under tensor powers. *Journal of Mathematical Physics*, 57(1):015202, 2015.

[44] A. Natarajan and T. Vidick. A Quantum Linearity Test for Robustly Verifying Entanglement. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, pages 1003–1015, New York, NY, USA, 2017. ACM.

[45] A. Nayak and P. Sen. Invertible quantum operations and perfect encryption of quantum states. *Quantum Information and Computation*, 7(1):103–110, 2007.

[46] V. Paulsen. *Completely Bounded Maps and Operator Algebras*. Cambridge University Press, Cambridge, 2003.

[47] M. Piani and J. Watrous. All Entangled States are Useful for Channel Discrimination. *Physical Review Letters*, 102(25):250501, 2009.

[48] D. Puzzuoli. `ancilla_dimension`. https://github.com/DanPuzzuoli/ancilla_dimension, 2016.

[49] D. Puzzuoli. Characterization of linear maps on M_n whose multiplicity maps have maximal norm, with an application in quantum information. *Quantum*, 2:51, 2018.

[50] D. Puzzuoli and J. Watrous. Ancilla Dimension in Quantum Channel Discrimination. *Annales Henri Poincaré*, 18(4):1153–1184, 2017.

[51] B. W. Reichardt, F. Unger, and U. Vazirani. Classical command of quantum systems via rigidity of CHSH games. *arXiv:1209.0449 [quant-ph]*, 2012. arXiv: 1209.0449.

[52] B. W. Reichardt, F. Unger, and U. Vazirani. Classical command of quantum systems. *Nature*, 496(7446):456–460, 2013.

[53] B. Rosgen. Additivity and distinguishability of random unitary channels. *Journal of Mathematical Physics*, 49(10):102107, 2008.

[54] B. Rosgen. Distinguishing short quantum computations. In *Proceedings of the 25th International Symposium on Theoretical Aspects of Computer Science*, pages 597–608, 2008.

[55] B. Rosgen and J. Watrous. On the hardness of distinguishing mixed-state quantum computations. In *Proceedings of the 20th Annual Conference on Computational Complexity*, pages 344–354, 2005.

[56] M. Sacchi. Entanglement can enhance the distinguishability of entanglement-breaking channels. *Physical Review A*, 72(1):014305, 2005.

[57] M. Sacchi. Optimal discrimination of quantum operations. *Physical Review A*, 71(6):062340, 2005.

[58] B. Schumacher and M. Nielsen. Quantum data processing and error correction. *Physical Review A*, 54(4):2629–2635, 1996.

[59] P. Shor, J. Smolin, and B. Terhal. Nonadditivity of bipartite distillable entanglement follows from a conjecture on bound entangled Werner states. *Physical Review Letters*, 86(12):2681–2684, 2001.

[60] R. Smith. Completely Bounded Maps between C*-Algebras. *Journal of the London Mathematical Society*, s2-27(1):157–166, 1983.

[61] D. Sutter, M. Berta, and M. Tomamichel. Multivariate Trace Inequalities. *Communications in Mathematical Physics*, 352(1):37–58, 2017.

[62] R. M. Timoney. Computing the norms of elementary operators. *Illinois Journal of Mathematics*, 47(4):1207–1226, 2003.

[63] J. Tomiyama. On the transpose map of matrix algebras. *Proceedings of the American Mathematical Society*, 88(4):635–638, 1983.

[64] J. Tomiyama. Recent Development of the Theory of Completely Bounded Maps between C*-Algebras. *Publications of the Research Institute for Mathematical Sciences*, 19(3):1283–1303, 1983.

[65] R. Tucci. Quantum entanglement and conditional information transmission. Available as arXiv.org e-Print quant-ph/9909041, 1999.

[66] G. Vidal and R. F. Werner. Computable measure of entanglement. *Physical Review A*, 65(3):032314, 2002.

[67] J. Watrous. Notes on Super-operator Norms Induced by Schatten Norms. *Quantum Info. Comput.*, 5(1):58–68, 2005.

[68] J. Watrous. Distinguishing quantum operations having few Kraus operators. *Quantum Information and Computation*, 8(9):819–833, 2008.

[69] J. Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.

[70] R. F. Werner and A. S. Holevo. Counterexample to an additivity conjecture for output purity of quantum channels. *Journal of Mathematical Physics*, 43(9):4353–4357, 2002.

[71] K. Wiesner. The careless use of language in quantum information. *arXiv:1705.06768*, 2017.

APPENDICES

Appendix A

Optimal value for independent strategies in the $k = 2$ case

To be precise, what we mean by an *independent strategy* for optimizing

$$\begin{aligned}&\left\|\left(\Psi_{n,2} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\right)(X)\right\|_1 \\ &\quad = \left\|\left(T \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\right)\left(\mathrm{Tr}_{\mathcal{X}_2}(X)\right)\right\|_1 + \left\|\left(T \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\right)\left(\mathrm{Tr}_{\mathcal{X}_1}(X)\right)\right\|_1\end{aligned} \tag{A.1}$$

for $X \in \mathrm{L}(\mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{Y})$, is an attempt at optimizing the above expression with an matrix of the following form. For $a, b \in \{1, \ldots, \dim(\mathcal{Y})\}$ with $ab \leq \dim(\mathcal{Y})$ and some $U \in \mathrm{U}(\mathbb{C}^a \otimes \mathbb{C}^b, \mathcal{Y})$, X takes the form

$$X = (\mathbb{1}_{\mathcal{X}_1 \otimes \mathcal{X}_2} \otimes U) \underbrace{(Y_1 \otimes Y_2)}_{\in \mathrm{L}(\mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathbb{C}^a \otimes \mathbb{C}^b)} (\mathbb{1}_{\mathcal{X}_1 \otimes \mathcal{X}_2} \otimes U^*) \tag{A.2}$$

for some $Y_1 \in \mathrm{L}(\mathcal{X}_1 \otimes \mathbb{C}^a)$ and $Y_2 \in \mathrm{L}(\mathcal{X}_2 \otimes \mathbb{C}^b)$ with $\|Y_1\|_1 = \|Y_2\|_1 = 1$, and we are again using the implicit permutation notation introduced in Section 5.1.1, before the proof of Theorem 5.4. For a matrix of this form we have

$$\begin{aligned}\left\|\left(\Psi_{n,2} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\right)(X)\right\|_1 = &\left\|\left(T \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^a)}\right)(Y_1)\right\|_1 \left\|\mathrm{Tr}_{\mathcal{X}_2}(Y_2)\right\|_1 + \\ &\left\|\left(T \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^b)}\right)(Y_2)\right\|_1 \left\|\mathrm{Tr}_{\mathcal{X}_1}(Y_1)\right\|_1.\end{aligned} \tag{A.3}$$

Corollary 7.2 says that when $\dim(\mathcal{Y}) \geq n^2$, optimal matrices are *necessarily* of this form. We now give the optimal value for these matrices when

$$n \leq \dim(\mathcal{Y}) < n^2. \tag{A.4}$$

Proposition A.1. *Let $\mathcal{X}_1$ and $\mathcal{X}_2$ denote copies of $\mathbb{C}^n$ and let $\mathcal{Y} = \mathbb{C}^m$ with*

$n \leq m < n^2$. *If* $X \in \mathrm{L}(\mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{Y})$ *is of the form given in Equation (A.2), then*

$$\left\| \left(\Psi_{n,2} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})} \right)(X) \right\|_1 \leq n + \lfloor m/n \rfloor, \tag{A.5}$$

and furthermore equality is achieved for some matrix of this form.

Proof. First, for such an X the value achieved in Equation (A.3) can be upper bounded by

$$\begin{aligned} \left\| \left(T \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^a)} \right)(Y_1) \right\|_1 \left\| \mathrm{Tr}_{\mathcal{X}_2}(Y_2) \right\|_1 + \left\| \left(T \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^b)} \right)(Y_2) \right\|_1 \left\| \mathrm{Tr}_{\mathcal{X}_1}(Y_1) \right\|_1 \\ \leq \left\| \left(T \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^a)} \right)(Y_1) \right\|_1 + \left\| \left(T \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^b)} \right)(Y_2) \right\|_1 \\ \leq \min(n, a) + \min(n, b), \end{aligned} \tag{A.6}$$

where the first inequality is monotonicity of the 1-norm under partial trace, and the second is two applications of Proposition 4.2. Next, observe that for fixed a and b, this value is attained by some choice of Y_1 and Y_2 (again, by Proposition 4.2), and finally, observe that by virtue of the min functions, there is no reason to consider either $a > n$ or $b > n$. In summary, the optimal value for matrices of this form is the same as the optimal value of the following simpler optimization problem

$$\max\{a + b : a, b \in \{1, \ldots, n\}, ab \leq m\} = \alpha. \tag{A.7}$$

Note that $a = n$ and $b = \lfloor m/n \rfloor$ satisfy the constraints, so $\alpha \geq n + \lfloor m/n \rfloor$.

To see that $\alpha \leq n + \lfloor m/n \rfloor$, consider the relaxed optimization problem

$$\max\{a + b : a, b \in [1, n], ab \leq m\} = \beta \geq \alpha. \tag{A.8}$$

For a given a the optimal value of b is $\min(n, m/a)$, so

$$\beta = \max\{a + \min(n, m/a) : a \in [1, n]\}. \tag{A.9}$$

The function $f(a) = a + \min(n, m/a)$ is strictly increasing over the interval $[1, m/n]$, so the optimum is achieved at some point in the interval $[m/n, n]$, on which $f(a) = a + m/a$. f is convex on $[m/n, n]$ as $f''(a) = 2m/a^3 > 0$, so the optimum is achieved at an endpoint, and in this case

$$f(m/n) = f(n) = n + m/n. \tag{A.10}$$

Hence

$$\alpha \leq \beta = n + m/n, \tag{A.11}$$

and since α is a natural number this implies $\alpha \leq n + \lfloor m/n \rfloor$. $\square$

Appendix B

Numerical tests

For $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, computing $\|\Phi\|_1$ is hard in general. However, as detailed in [30], there are nice algorithms for computing lower bounds to $\|\Phi\|_1$. For $2 \leq n \leq 6$ and $n \leq m \leq n^2$, Table B.1 contains computed lower bounds for $\|\Psi_{n,2} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\|_1$, as well as computed lower bounds for $\|\Psi_{n,2} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\|_{1,H}$, where

$$\|\Phi\|_{1,H} = \max\{\|\Phi(H)\|_1 : H \in \mathrm{Herm}(\mathcal{X}), \|H\|_1 = 1\}. \tag{B.1}$$

The computations were done in MATLAB using modified versions of the function `InducedSchattenNorm` in the QETLAB [31] package (which uses the algorithm in [30]). For $n = 5$ and $n = 6$, plots ranging over $n \leq m \leq n^2$ are given in Figure B.1. The code and data used in this appendix can be found in the GitHub repository at [48].

One feature of the data is that the lower bounds for $\|\Psi_{n,2} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\|_1$ and $\|\Psi_{n,2} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\|_{1,H}$ almost always agree (up to stopping precision), and in cases of disagreement the value computed for Hermitian inputs is always the larger of the two. This lends evidence to the conjecture that

$$\left\|\Psi_{n,2} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\right\|_1 = \left\|\Psi_{n,2} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\right\|_{1,H}, \tag{B.2}$$

and the stronger conjecture that

$$\left\|\Psi_{n,k} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\right\|_1 = \left\|\Psi_{n,k} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\right\|_{1,H} \tag{B.3}$$

for all k.

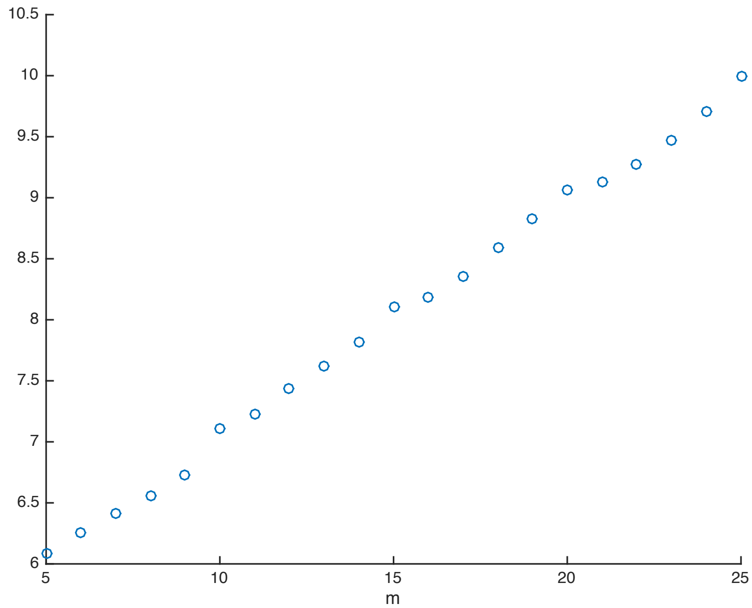
Another curious feature, displayed in Figure B.1, is that while seeming to increase roughly linearly in m, there is a bump when m is a multiple of n, with dips between these points. It is unclear whether this is an actual feature of $\left\| \Psi_{n,2} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)} \right\|_1$ or is a peculiarity of the lower bounds found by the algorithm.

Table B.1: Lower bounds for $\left\|\Psi_{n,2} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\right\|_1$ and $\left\|\Psi_{n,2} \otimes \mathbb{1}_{\mathrm{L}(\mathbb{C}^m)}\right\|_{1,H}$ (the columns with '-H') for $2 \leq n \leq 6$ (columns) and $n \leq m \leq n^2$ (rows), computed using 1000 initial guesses and a stopping tolerance of 10^{-5}.

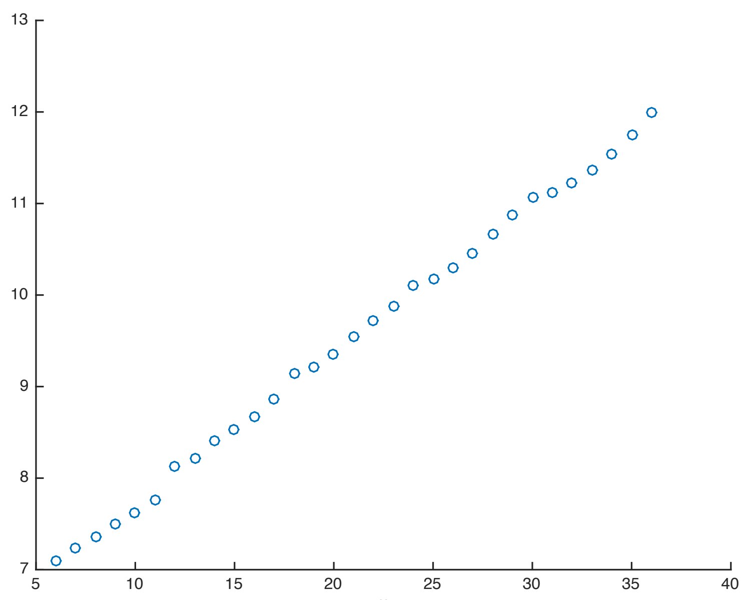
m\n	2	2-H	3	3-H	4	4-H	5	5-H	6	6-H
2	3.0448	3.0448								
3	3.4142	3.4142	4.0656	4.0656						
4	4.0000	4.0000	4.3307	4.3307	5.0777	5.0777				
5			4.6386	4.6386	5.2830	5.2830	6.0857	6.0857		
6			5.0551	5.0551	5.4711	5.4711	6.2527	6.2527	7.0914	7.0914
7			5.2361	5.2361	5.6949	5.6949	6.4100	6.4100	7.2319	7.2319
8			5.5615	5.5616	6.0896	6.0896	6.5593	6.5593	7.3666	7.3666
9			6.0000	6.0000	6.2240	6.2241	6.7331	6.7331	7.4961	7.4961
10					6.4873	6.4873	7.1136	7.1136	7.6209	7.6209
11					6.7635	6.7635	7.2207	7.2209	7.7611	7.7611
12					7.0596	7.0596	7.4396	7.4396	8.1312	8.1312
13					7.1622	7.1623	7.6222	7.6222	8.2202	8.2206
14					7.3722	7.3723	7.8151	7.8152	8.4068	8.4068
15					7.6457	7.6457	8.1023	8.1023	8.5342	8.5342
16					8.0000	8.0000	8.1873	8.1874	8.6700	8.6701
17							8.3605	8.3605	8.8563	8.8564
18							8.5850	8.5850	9.1344	9.1344
19							8.8297	8.8297	9.2058	9.2061
20							9.0623	9.0623	9.3479	9.3480
21							9.1295	9.1296	9.5437	9.5437
22							9.2749	9.2749	9.7192	9.7192
23							9.4641	9.4641	9.8829	9.8830
24							9.7016	9.7016	10.1101	10.1101
25							10.0000	10.0000	10.1708	10.1711
26									10.2970	10.2971
27									10.4621	10.4621
28									10.6717	10.6717
29									10.8717	10.8717
30									11.0639	11.0639
31									11.1145	11.1146
32									11.2170	11.2170
33									11.3589	11.3589
34									11.5311	11.5311
35									11.7416	11.7416
36									12.0000	12.0000

Figure B.1: Plots for the data in Table B.1 for $n = 5$ and $n = 6$.

(a) $n = 5$, $5 \leq m \leq 25$.



(b) $n = 6$, $6 \leq m \leq 36$.



Appendix C

Proof of Theorem 9.5

In this appendix we provide a proof of Theorem 9.5, as well as a proof of Conjecture 9.4 in the special case $\mathcal{Y} = \mathbb{C}$. We require some lemmas. The following is well-known.

Lemma C.1. *For $A, B \in \mathrm{L}(\mathcal{X})$, it holds that*

$$\|A + B\|_1 = \|A\|_1 + \|B\|_1, \text{ and } \|A - B\|_1 = \|A\|_1 + \|B\|_1 \tag{C.1}$$

if and only if $A^*B = AB^* = 0$. In particular, for $X \in \mathrm{L}(\mathcal{X})$, it holds that

$$\|X + X^*\|_1 = 2\|X\|_1, \text{ and } \|X - X^*\|_1 = 2\|X\|_1 \tag{C.2}$$

if and only if $X^2 = 0$, i.e. X is nilpotent.

Proof. First, observe that the matrices

$$Y = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}, \text{ and } Z = \begin{pmatrix} B & 0 \\ 0 & -B \end{pmatrix} \tag{C.3}$$

are Hilbert-Schmidt orthogonal. Furthermore,

$$\begin{aligned} \|A + B\|_1 + \|A - B\|_1 &= \left\| \begin{pmatrix} A + B & 0 \\ 0 & A - B \end{pmatrix} \right\|_1 && \text{(C.4)} \\ &= \|Y + Z\|_1 && \text{(C.5)} \\ &\leq \|Y\| + \|Z\|_1 && \text{(C.6)} \\ &= 2(\|A\|_1 + \|B\|_1), && \text{(C.7)} \end{aligned}$$

where the inequality is the triangle inequality. Thus, as individually

$$\|A \pm B\|_1 \leq \|A\|_1 + \|B\|_1, \tag{C.8}$$

it holds that both $\|A \pm B\|_1 = \|A\|_1 + \|B\|_1$ if and only if equality holds in the above triangle inequality. As Y and Z are Hilbert-Schmidt orthogonal, Proposition 5.1 gives that equality holds if and only if $Y^*Z = YZ^* = 0$, which is equivalent to $A^*B = AB^* = 0$. □

Lemma C.2. *For $A, B \in \mathrm{L}(\mathcal{X})$, $\|A\|_1 = \|A + \lambda B\|_1$ for all $\lambda \in \{\pm 1, \pm i\}$ if and only if $B = 0$.*

Proof. Let $U \in \mathrm{U}(\mathcal{X})$ be a unitary for which $\mathrm{Tr}(UA) = \|A\|_1$. For $\lambda \in \{1, i\}$, it holds that

$$\begin{aligned} 2\|A\|_1 = 2\mathrm{Tr}(UA) &= \mathrm{Tr}(U(A + \lambda B)) + \mathrm{Tr}(U(A - \lambda B)) && \text{(C.9)} \\ &\leq |\mathrm{Tr}(U(A + \lambda B))| + |\mathrm{Tr}(U(A - \lambda B))| && \text{(C.10)} \\ &\leq \|A + \lambda B\|_1 + \|A - \lambda B\|_1 && \text{(C.11)} \\ &= 2\|A\|_1. && \text{(C.12)} \end{aligned}$$

Hence, the above inequalities are equalities, implying that

$$\mathrm{Tr}(U(A \pm \lambda B)) = \|A \pm \lambda B\|_1, \tag{C.13}$$

and hence $U(A \pm \lambda B) \geq 0$. This implies that both UB and iUB are Hermitian, which is only possible when $B = 0$. □

Proof of Theorem 9.5. First, consider an arbitrary $X \in \mathrm{L}(\mathcal{X})$ with $\|X\|_1 = 1$ that satisfies $\|\Phi(X)\|_1 = \|\Phi\|_1 = 2\|\Phi\|_{1,H}$. For any $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ it holds that

$$(X + \lambda X^*)^* = X^* + \overline{\lambda} X = \overline{\lambda}(X + \lambda X^*). \tag{C.14}$$

In particular, this implies that $X \pm \lambda X^*$ is normal, and as we can alternatively write

$$\|\Phi\|_{1,H} = \max\{\|\Phi(N)\|_1 : N \in \mathrm{L}(\mathcal{X}), \|X\|_1 = 1, N^*N = NN^*\}, \tag{C.15}$$

it holds that

$$\|\Phi(X + \lambda X^*)\|_1 \leq \|\Phi\|_{1,H} \|X + \lambda X^*\|_1 \tag{C.16}$$

for any $\lambda \in \mathbb{C}$ with $|\lambda| = 1$. Thus,

$$
\begin{aligned}
2\|\Phi\|_1 = 2\|\Phi(X)\|_1 &= \|\Phi(X + \lambda X^*) + \Phi(X - \lambda X^*)\|_1 && \text{(C.17)} \\
&\leq \|\Phi(X + \lambda X^*)\|_1 + \|\Phi(X - \lambda X^*)\|_1 && \text{(C.18)} \\
&\leq \|\Phi\|_{1,H}(\|X + \lambda X^*\|_1 + \|X - \lambda X^*\|_1) && \text{(C.19)} \\
&\leq 4\|\Phi\|_{1,H}, && \text{(C.20)}
\end{aligned}
$$

where we have used that $\|X \pm \lambda X^*\|_1 \leq 2$ in the last line. As $\|\Phi\|_1 = 2\|\Phi\|_{1,H}$, all of the above inequalities are equalities.

Observe:

- Equality in Equation (C.20), for all $|\lambda| = 1$, implies by Lemma C.1 that $X^2 = 0$.
- $\|\Phi(X + \lambda X^*)\|_1 = 2\|\Phi\|_{1,H} = \|\Phi(X)\|_1$ for all $|\lambda| = 1$, and so Lemma C.2 gives that $\Psi(X^*) = 0$.

Restricting to the case $X = uv^*$ for unit vectors $u, v \in \mathcal{X}$, the first point above gives that u and v are orthogonal, and the second point gives $\Phi(vu^*) = 0$. Hence, to complete the proof, we need only show $\Phi(uu^* + vv^*) = 0$.

As u and v are orthogonal, it holds that $\|X \pm X^*\|_1 = 2$. Furthermore, as $\|\Psi(X \pm X^*)\|_1 = 2\|\Psi\|_{1,H}$, for each eigenvector z of either $X + X^*$ or $X - X^*$, it holds that $\|\Psi(zz^*)\|_1 = \|\Psi\|_{1,H}\|z\|^2$. The (un-normalized) eigenvectors of $X + X^*$ are $x_\pm = u \pm v$, and those of $X - X^*$ are $y_\pm = u \pm iv$. Hence,

$$
\|\Psi(uv^*)\| = 2\|\Psi\|_{1,H} = \|\Psi(x_\pm x_\pm^*)\|_1 = \|\Psi(uv^*) \pm \Psi(uu^* + vv^*)\|_1, \quad \text{(C.21)}
$$

and

$$
\|\Psi(uv^*)\| = 2\|\Psi\|_{1,H} = \|\Psi(y_\pm y_\pm^*)\|_1 = \|\Psi(uv^*) \pm i\Psi(uu^* + vv^*)\|_1. \quad \text{(C.22)}
$$

Thus, again by Lemma C.2, it holds that $\Psi(uu^* + vv^*) = 0$. □

It seems plausible that some kind of trick as above will work to show that $\Psi(uu^* - vv^*) = 0$ as well.

C.1 Special case: $\Phi \in \mathrm{T}(\mathcal{X}, \mathbb{C})$

Consider the special case that $\mathcal{Y} = \mathbb{C}$, i.e. $\Phi \in \mathrm{T}(\mathcal{X}, \mathbb{C})$ is a linear functional. As such a linear map is of the form $\Phi(X) = \langle A, X \rangle$ for some $A \in \mathrm{L}(\mathcal{X})$, the statement that $\|\Phi\|_1 = 2\|\Phi\|_{1,H}$ is the same as saying that $\|A\| = 2w(A)$. Hence, characterizing the elements of $\mathrm{T}(\mathcal{X}, \mathbb{C})$ with $\|\Phi\|_1 = 2\|\Phi\|_{1,H}$ is equivalent to characterizing matrices with maximal gap between their operator norm and numerical radius. This is certainly known but it is instructive to complete the proof in this case.

For unit vectors $u, v \in \mathcal{X}$ with $\|\Phi(uv^*)\|_1 = \|\Phi\|_1 = 2\|\Phi\|_{1,H} = 1$. By Theorem 9.5, u and v are orthogonal, so we may assume without loss of generality that $u = e_0$ and $v = e_1$, and we furthermore assume without loss of generality that $\Phi(E_{01}) = 1$. Theorem 9.5 also implies that $\Phi(E_{00} + E_{11}) = \Phi(E_{10}) = 0$. Hence, on matrices $X \in \operatorname{span}\{E_{00}, E_{01}, E_{10}, E_{11}\}$, it holds that

$$\Phi(X) = \langle \alpha(E_{00} - E_{11}) + E_{01}, X \rangle \tag{C.23}$$

for some $\alpha \in \mathbb{C}$, and hence $\|\Phi\|_1 = \|\alpha(E_{00} - E_{11}) + E_{01}\|$. However, it holds that

$$\begin{aligned}
\|\alpha(E_{00} - E_{11}) + E_{01}\|^2 &= \left\| \begin{pmatrix} \alpha & 1 \\ 0 & -\alpha \end{pmatrix} \right\|^2 && \text{(C.24)} \\
&= \left\| \begin{pmatrix} \alpha & 1 \\ 0 & -\alpha \end{pmatrix}^* \begin{pmatrix} \alpha & 1 \\ 0 & -\alpha \end{pmatrix} \right\| && \text{(C.25)} \\
&= \left\| \begin{pmatrix} |\alpha|^2 & \overline{\alpha} \\ \alpha & 1 + |\alpha|^2 \end{pmatrix} \right\| && \text{(C.26)} \\
&\geq 1 + |\alpha|^2. && \text{(C.27)}
\end{aligned}$$

Hence, $\|\Phi\|_1 \geq 1 + |\alpha|^2$, but as $\|\Phi\|_1 = 1$, necessarily $\alpha = 0$. Thus,

$$\Phi(E_{00}) = \Phi(E_{11}) = \Phi(E_{10}) = 0, \tag{C.28}$$

as required.