Separable State Discrimination Using Local Quantum Operations and Classical Communication

by

Laura Mančinska

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Combinatorics and Optimization – Quantum Information

Waterloo, Ontario, Canada, 2013

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

In this thesis we study the subset of quantum operations that can be implemented using only local quantum operations and classical communication (LOCC). This restricted paradigm serves as a tool to study not only quantum correlations and other nonlocal quantum effects, but also resource transformations such as channel capacities.

The mathematical structure of LOCC is complex and difficult to characterize. In the first part of this thesis we provide a precise description of LOCC and related operational classes in terms of quantum instruments. Our formalism captures both finite round protocols as well as those that utilize an unbounded number of communication rounds. This perspective allows us to measure the distance between two LOCC instruments and hence discuss the closure of LOCC in a rigorous way. While the set LOCC is not topologically closed, we show that the operations that can be implemented using some fixed number rounds of communication constitute a compact subset of all quantum operations. We also exhibit a subset of LOCC measurements that is closed. Additionally we establish the existence of an open ball around the completely depolarizing map consisting entirely of LOCC implementable maps.

In the second part of this thesis we focus on the task of discriminating states from some known set S by LOCC. Building on the work in the paper *Quantum nonlocality without entanglement*, we provide a framework for lower bounding the error probability of any LOCC protocol aiming at discriminating the states from S. We apply our framework to an orthonormal product basis known as the domino states. This gives an alternative and simplified bound quantifying how well these states can be discriminated using LOCC. We generalize this result for similar bases in larger dimensions, as well as the "rotated" domino states, resolving a long-standing open question. These results give new examples of quantitative gaps between the classes of separable and LOCC operations.

In the last part of this thesis, we ask what differentiates separable from LOCC operations. Both of these classes play a key role in the study of entanglement. Separable operations are known to be strictly more powerful than LOCC ones, but no simple explanation of this phenomenon is known. We show that, in the case of bipartite von Neumann measurements, the ability to interpolate is an operational principle that separates LOCC from all separable operations.

Acknowledgements

First and foremost, I would like to thank my supervisors Debbie Leung and Andrew Childs. Without their guidance and insight this project would not have been possible. I want to thank Debbie for suggesting the investigation of the domino states, which was the starting point of this thesis. I am very fortunate to have been part of both IQC and Combinatorics and Optimization department of University of Waterloo. Both of these institutions have given me an example of academic excellence to strive for.

I want to thank my co-authors Eric Chitambar, Maris Ozols, and Andreas Winter for giving me the opportunity to work with them. In particular, I want to thank Maris with whom we have spent endless hours trying to find an alternative for the notorious appendix B. I wish to acknowledge Dagmar Bruß, Sarah Croke, Graeme Smith, Hermann Kampermann, Matthias Kleinmann, Marco Piani, David Reeb and John Watrous for their helpful discussions regarding LOCC. I also want to thank John Watrous for his course notes to which I turned in doubtful moments. I am grateful in advance to my committee members Dagmar Bruß, Andrew Childs, Debbie Leung, Michele Mosca, Ashwin Nayak, and John Watrous for reading my thesis and their feedback.

I have been blessed with great friends who have made the years spent in Waterloo full of fun and adventures. I am most grateful to my dear Mom and Dad who have always been there to support and encourage me. Finally, I am indebted to David, who has patiently listened to my monologues on LOCC and always been there with me in the good times and bad.

To my parents
(Mammai un Tuksim)

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

The "distant lab" paradigm plays a crucial role in both theoretical and experimental aspects of quantum information. Here, a multipartite quantum system is distributed to various parties, and they are restricted to act locally on their respective subsystems by performing quantum operations. However, in order to enhance their measurement strategies, the parties are free to communicate any classical data, which includes the sharing of randomness and previous measurement results. Quantum operations implemented in such a manner are known as LOCC (local operations with classical communication). We can think of LOCC as a special subset of all physically realizable operations on the global system. This restricted paradigm serves as a tool to study not only quantum correlations and other nonlocal quantum effects, but also resource transformations such as channel capacities.

Using LOCC operations to study resource transformation is illustrated in quantum teleportation [BBC+93]. Two parties, called Alice and Bob, are separated in distant labs. Equipped with some pre-shared quantum states that characterize their entanglement resource, they are able to transmit quantum states from one location to another using LOCC; specifically, the exchange rate is one quantum bit (qubit) transmitted for one entangled bit (ebit) plus two classical bits (cbits) consumed. Via teleportation then, LOCC operations become universal in the sense that Alice and Bob can implement *any* physical evolution of their joint system given a sufficient supply of pre-shared entanglement. Thus, entanglement represents a fundamental resource in quantum information theory, with LOCC being the class of operations that manipulates and consumes this resource [BBPS96, Nie99, BPR+00]. Indeed, the class of non-entangled or separable quantum states are precisely those that can be generated exclusively by the action of LOCC on pure product states [Wer89], and any sensible measure of entanglement must satisfy the crucial property that it is non-increasing under LOCC [BDSW96, VPRK97, HHH00, PV07]. In fact, the present day entanglement theory is shaped by the viewing entanglement as a resource that is manipulated by LOCC [HHHH09].

The intricate structure of LOCC was perhaps first realized over 20 years ago by Peres and Wootters, who observed that when some classical random variable is encoded into an ensemble of bipartite product states, the accessible information may be appreciably reduced if the decoders—played by Alice and Bob—are restricted to LOCC operations [PW91]. (In fact, results in [PW91] led to the discovery of teleportation.) Several years later, Massar and Popescu analytically confirmed the spirit of this conjecture by considering pairs of particles that are polarized in the same randomly chosen direction [MP95]. It was shown that when Alice and Bob are limited to a finite number of classical communication exchanges, their LOCC ability to identify the polarization angle is strictly less than if they were allowed to make joint measurements on their shared states.

The examples concerning accessible information mentioned above demonstrate that a gap between the LOCC and globally accessible information exists even in the absence of entanglement. This finding suggests that nonlocality and entanglement are two distinct concepts, with the former being more general than the latter. Bennett *et al.* in [BDF+99] were able to sharpen this intuition by constructing a set of orthogonal bipartite pure product states that demonstrate "nonlocality without entanglement" in the sense that elements of the set could be perfectly distinguished by so-called separable operations (SEP) but not by LOCC [BDF+99]. This was also the first demonstration of a quantitative gap between the classes of LOCC and separable operations. Separable operations are precisely the class of maps whose Choi matrices are separable, and consequently SEP inherits the relatively well-understood mathematical structure possessed by separable states [HHH96, BCH+02]. The fact that SEP and LOCC are distinct classes means that LOCC lacks this nice mathematical characterization, and its structure is therefore much more subtle than SEP.

If quantum information is stored in a physical medium that allows easy restriction of incoming and outgoing quantum information[1], then the interaction of several such systems could be modeled using LOCC framework. In such cases, the fact that not all global information is accessible via LOCC can be harnessed for cryptographic purposes.

[1]Most of the current implementations except the photonic qubit meet this requirement.

Perhaps the best known example is quantum data hiding [TDL01, DLT02, EW02]. Here, some classical data is encoded into a bipartite state such that Alice and Bob have arbitrarily small accessible information when restricted to LOCC, while the data can be perfectly retrieved by global measurements. The LOCC model has also enabled construction of one-time memories (OTMs) [GKR08] that are information-theoretically secure against certain LOCC adversaries [Liu13]. An OTM is a tamper resistant cryptographic hardware that does not exist in fully quantum or fully classical setting (see *e.g.*, [Liu13] for more details).

The LOCC model has also been used in quantum complexity theory. A quantum multiprover analogue of NP is the class $\mathrm{QMA}(k)$ which consists of all languages that Arthur can decide in polynomial time using k unentangled polynomially-sized quantum proofs provided by Merlins [KMY03]. To understand whether many Merlins can be more helpful to Arthur, the class $\mathrm{QMA}_{\mathrm{LOCC}}(k)$ was introduced [ABD+08]. Here, Arthur is only allowed to perform k-party LOCC measurements on the proof provided by the k Merlins. In [BaCY11] the authors show that the class QMA (see *e.g.*, [Wat12]) coincides with the class of languages that Arthur can decide by measuring the k unentangled proofs with some one-way LOCC measurements.

In this thesis we study the class of LOCC operations and its relation to the class of separable operations. We provide precise definitions of different classes of LOCC operations and establish some of their basic topological properties (Chapter 2). Building on the work of [BDF+99], we investigate separable state discrimination with LOCC and provide new quantitative gaps between the classes of separable and LOCC operations (Chapter 4). Finally, we propose an operational principle that sets apart LOCC from general separable operations (Chapter 5).

1.2 Overview

In this section we explain the organization of the rest of this thesis and state the main results. See Figure 1.1 for an overview of this thesis.

In Chapter 2 we provide precise definitions of different classes of LOCC in terms of quantum instruments. Our formalism captures both finite round protocols as well as those that utilize an unbounded number of communication rounds. Although the class LOCC is known not to be closed [CCL12a, CCL12b, CLM+12], we show that the class of instruments that can be implemented using any fixed number of communication rounds is compact (Theorem 2.21). Additionally, we establish the existence of a ball of

1. Introduction

2. LOCC model [CLM+12]

3. State discrimination

4. Framework for bounding nonlocality [CLMO12]

5. Interpolatability distinguishes LOCC from SEP [CLMO13]

Figure 1.1: Overview of the thesis.

LOCC instruments around the maximally depolarizing channel, as well as provide a lower bound on its radius (Theorem 2.17). The material in Chapter 2 is based on results from [CLM+12] which is joint work with Eric Chitambar, Debbie Leung, Maris Ozols, and Andreas Winter.

In Chapter 3 we provide the common background for the results concerning state discrimination and implementation of von Neumann measurements developed in the later chapters. Most importantly, we define the state discrimination problem and explain its relation to that of implementing von Neumann measurements. We also introduce the concept of non-disturbing measurements and describe its relevance for state discrimination with LOCC. Finally, we prove that asymptotic and finite LOCC are equally powerful for implementing projective measurements with tensor product operators (Theorem 3.17). This generalizes a similar result from [KKB11] regarding von Neumann measurements.

In Chapter 4 we revisit the problem of separable state discrimination with LOCC from [BDF+99]. We show that any LOCC measurement for discriminating states from a set S errs with probability

$$p_{\text{error}} \geq \frac{2}{27} \frac{\eta^2}{|S|^5}, \tag{1.1}$$

where η is a constant that depends on S (Theorem 4.7). Intuitively, η measures the nonlocality of S (see Definition 4.4). We systematically bound η for a large class of product bases. This lets us quantify the hardness of LOCC discrimination for the domino states from [BDF+99] (Corollary 4.15), domino-type states (Corollary 4.17), and rotated domino states (Corollary 4.19). The material in Chapter 4 is based on results from [CLMO12], which is joint work with Andrew Childs, Debbie Leung, and Maris Ozols.

We conclude with Chapter 5 where we ask why some separable operations can be implemented with LOCC while others cannot. We answer this question for von Neumann measurements. More precisely, we show that a von Neumann measurement can be interpolated if and only if it can be decomposed into two parts, the first of which can be implemented by LOCC (Theorem 5.10). This suggests the ability to interpolate as an operational principle that distinguishes LOCC from all separable operations. The material in Chapter 5 is based on results from [CLMO13], which is joint work with Andrew Childs, Debbie Leung, and Maris Ozols.

1.3 Preliminaries

1.3.1 Notation

The following notation is used in this thesis. Let $\mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$ be the set of all linear operators from $\mathbb{C}^n$ to $\mathbb{C}^m$ and let $\mathrm{L}(\mathbb{C}^n) := \mathrm{L}(\mathbb{C}^n, \mathbb{C}^n)$. Next, let $\mathrm{Herm}(\mathbb{C}^n) \subseteq \mathrm{L}(\mathbb{C}^n)$ be the set of all Hermitian operators on $\mathbb{C}^n$ and let $\mathrm{Pos}(\mathbb{C}^n) \subseteq \mathrm{Herm}(\mathbb{C}^n)$ be the set of all positive semidefinite operators on $\mathbb{C}^n$.

Let $\|M\|_{\max} := \max_{ij}|M_{ij}|$ denote the largest entry of $M \in \mathrm{L}(\mathbb{C}^n)$ in absolute value and $\mathrm{spec}(M)$ be the spectrum of M. Finally, for any natural number n, let $[n] := \{1, \ldots, n\}$ and let I_n be the $n \times n$ identity matrix.

1.3.2 Quantum operations

We are interested in studying quantum operations, *i.e.*, maps that send quantum states to quantum states. Any quantum state can be described using its density matrix $\rho \in \mathrm{L}(\mathbb{C}^{d_1})$ where $d_1 \in \mathbb{N}$. Thus, quantum operations are maps that send elements of $\mathrm{L}(\mathbb{C}^{d_1})$ to elements of $\mathrm{L}(\mathbb{C}^{d_2})$ for some $d_1, d_2 \in \mathbb{N}$. Let $\mathrm{T}(\mathbb{C}^{d_1}, \mathbb{C}^{d_2})$ be the complex vector space consisting of the linear operators of the form $\mathcal{E} : \mathrm{L}(\mathbb{C}^{d_1}) \to \mathrm{L}(\mathbb{C}^{d_2})$. We say that a map $\mathcal{E} \in \mathrm{T}(\mathbb{C}^{d_1}, \mathbb{C}^{d_2})$ is

- *trace-preserving*, if $\mathrm{Tr}(\mathcal{E}(\rho)) = \mathrm{Tr}(\rho)$ for all $\rho \in \mathrm{L}(\mathbb{C}^{d_1})$;
- *trace-nonicreasing*, if $\mathrm{Tr}(\mathcal{E}(\rho)) \leq \mathrm{Tr}(\rho)$ for all $\rho \in \mathrm{L}(\mathbb{C}^{d_1})$;
- *completely positive* (CP), if for all $d_3 \in \mathbb{N}$ and all $\rho \in \mathrm{Pos}(\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_3})$ we have $(\mathcal{E} \otimes \mathcal{I})\rho \in \mathrm{Pos}(\mathbb{C}^{d_2} \otimes \mathbb{C}^{d_3})$, where $\mathcal{I}$ is the identity map on $\mathrm{L}(\mathbb{C}^{d_3})$.

A *quantum operation* (or channel) is any map $\mathcal{E} \in \mathrm{T}(\mathbb{C}^{d_1}, \mathbb{C}^{d_2})$ that is trace-preserving and completely positive (TCP).

1.3.3 Quantum measurements

Quantum measurements can be viewed as quantum operations having both classical and quantum output registers.

A k-outcome *(quantum) measurement* $\mathcal{M}$ on an n-dimensional vector space can be specified by a set of operators $\{M_1, \ldots, M_k\} \subseteq \mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$ where $\sum_{i=1}^{k} M_i^\dagger M_i = I_n$ and m is finite. We refer to the operators M_i as the *measurement operators of* $\mathcal{M}$. The probability of obtaining outcome i upon measuring state ρ is $\mathrm{Tr}(M_i^\dagger M_i \rho)$; the corresponding post-measurement state is $M_i \rho M_i^\dagger \,/\, \mathrm{Tr}(M_i^\dagger M_i \rho)$.

If we are only interested in the inner products between the post-measurement states, we can specify the measurement $\mathcal{M}$ using its *POVM elements* $\{E_1, \ldots, E_k\}$, where $E_i := M_i^\dagger M_i$ for all $i \in [k]$. We refer to the set $\{E_1, \ldots, E_k\}$ as the *POVM of measurement* $\mathcal{M}$. Such a POVM only specifies the post-measurement states up to an isometry. Since for state discrimination we are only interested in producing the correct classical outcome, most of the time we will specify a measurement using its POVM elements rather than its measurement operators. Sometimes, we use the term *POVM measurement* to denote the measurement with only a classical outcome and no quantum post-measurement state.

A *projective* measurement is a measurement whose measurement operators are orthogonal projectors. A *von Neumann* measurement is a projective measurement whose measurement operators have rank one. Such measurements correspond naturally to orthonormal bases of the appropriate space. Therefore, we can specify a von Neumann measurement by indicating in which orthonormal basis it measures.

We say that a matrix $E \in \mathrm{Pos}(\mathbb{C}^d)$ is *trivial* if it is proportional to the identity matrix. We call a measurement $\mathcal{M}$ *trivial* or an *identity measurement*, if each of its POVM elements $E \in \mathcal{M}$ is trivial. The choice of these terms is motivated by the fact that identity measurements do not reveal any information about the state being measured. Given two measurements $\mathcal{M}_A = \{E_i\}_i \subseteq \mathrm{Pos}(\mathbb{C}^{d_A})$ and $\mathcal{M}_B = \{F_j\}_j \subseteq \mathrm{Pos}(\mathbb{C}^{d_B})$, we use $\mathcal{M}_A \otimes \mathcal{M}_B$ to denote the measurement with POVM elements $E_i \otimes F_j$. We use $\mathcal{I}_A$ and $\mathcal{I}_B$ to denote the trivial measurement with exactly one POVM element on Alice and Bob respectively. For example, $\mathcal{I}_A \otimes \mathcal{M}_B$ is the measurement with POVM elements $I_{d_A} \otimes F_j$.

1.3.4 Contraction diagrams

Contraction diagrams (also known as tensor networks) are a useful graphical tool for manipulating algebraic expressions. We now give a brief overview of the elements used

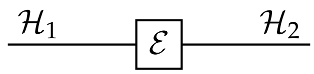


Figure 1.2: Application of $\mathcal{E} \in \mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)$ to vector space $\mathcal{H}_1$.

(a) Column vector $|\Phi\rangle$.

(b) Row vector $\langle\Phi|$.

Figure 1.3: Graphical representation of row and column vectors corresponding to an unnormalized maximally entangled state.

in later sections. For a more thorough discussion see, for example, [WBC11]. We use the contraction diagrams mostly for illustrative purposes and none of the presented proofs depend crucially on them.

Each wire represents some complex vector space; parallel wires are used to represent tensor product of spaces. Sometimes we use labels to indicate the complex vector space corresponding to a wire. We depict application of a linear map $\mathcal{E} \in \mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)$ to some complex vector space $\mathcal{H}$ by drawing a labeled block on the corresponding wire (see Figure 1.2). An unnormalized maximally entangled state on complex vector space $\mathcal{H}$

$$|\Phi\rangle := \sum_{i \in [\dim(\mathcal{H})]} |i, i\rangle \tag{1.2}$$

and its row version $\langle\Phi|$ are represented using right and left angles respectively (see Figure 1.3). Since $\big(\langle\Phi| \otimes I_{\mathcal{H}}\big)\big(I_{\mathcal{H}} \otimes |\Phi\rangle\big) = I_{\mathcal{H}}$, we can also compose the corresponding graphical representations to form the identity operator. This allows to straighten the wire with two bends by "pulling at both ends". We find this operation particularly convenient (see Figure 1.4 for an example).

Figure 1.4: An example of straightening a wire with two bends.

Chapter 2

Local operations and classical communication model

2.1 Introduction

While LOCC emerges as the natural class of operations in many important quantum information tasks, its mathematical structure is complex and difficult to characterize. The goal of the first part of this chapter is to provide a precise description of LOCC and related operational classes in terms of quantum instruments. Our formalism captures both finite round protocols as well as those that utilize an unbounded number of communication rounds.

Using the description from the first part, we continue with the discussion of some topological properties of the set of LOCC operations. It is known that the set of LOCC operations is not topologically closed [CCL12a, CCL12b, CLM+12]. In contrast to this, we show that the LOCC operations that can be implemented using some fixed number of communication rounds constitute a compact subset of all quantum operations. Additionally we show that the interior of of LOCC implementable maps is nonempty. We achieve this by exhibiting a ball of LOCC implementable maps around around the completely depolarizing map.

This chapter is based on the results obtained in collaboration with Eric Chitambar, Debbie Leung, Maris Ozols and Andreas Winter [CLM+12]. We start this chapter by explaining our motivation in Section 2.1.1 and stating our contributions as well as outlining the rest of this chapter in Section 2.1.2.

2.1.1 Motivation

Like all quantum operations, an LOCC operation is a trace-preserving completely positive map acting on the space of density operators (or by a "quantum instrument" as we describe below), and the difficulty is in describing the precise structure of these maps. Part of the challenge stems from the way in which LOCC operations combine the globally shared classical information at one time with the particular choice of local measurements at a later time. The potentially unrestricted number of rounds of communication further complicates the analysis. A thorough definition of finite-round LOCC has been presented in [DHR02], thus formalizing the description given in [BDF+99].

Recently, there has been a renewed wave of interest in LOCC alongside new discoveries concerning asymptotic resources in LOCC processing [Rin04, Chi11, KKB11, CLM+12]. It has now been shown that when an unbounded number of communication rounds are allowed, or when a particular task needs only to be accomplished with an arbitrarily small but nonzero failure rate, more can be accomplished than in the setting of finite rounds and perfect success rates. Consequently, we can ask whether a task can be performed by LOCC, and failing that, whether or not it can be approximated by LOCC, and if so, whether a simple recursive procedure suffices. To make these notions precise, a definition of LOCC and its topological closure is needed. Here, we aim to extend the formalisms developed in [BDF+99, DHR02, KKB11] to facilitate an analysis of asymptotic resources and a characterization of the most general LOCC protocols. Indeed, we hope that this work will provide a type of "LOCC glossary" for the research community.

2.1.2 Our contributions

Our first contribution is to provide a precise description of LOCC in terms of quantum instruments.[1] This enables us to measure the distance between two LOCC instruments using a metric induced by the diamond norm (see Equation (2.3)). The introduced metric further allows us to rigorously define the class of instruments that can be

[1] An alternative characterization of LOCC in terms of physical tasks was described to us privately by the authors of [KKB11]. Instead of considering how well an LOCC map approximates a target map (by a distance measure on maps), one can define a success measure for a particular task and study the achievable values via LOCC. This is particularly useful when the task does not uniquely define a target map. Here, we define LOCC in terms of quantum instruments, which admits a more precise mathematical description, and a further optimization over possible target maps can be added if one wishes to focus on success rates.

implemented with infinite-round LOCC protocols as well as the more general class of LOCC-closure (Definition 2.5).

Our second contribution is to establish some topological properties of the classes of LOCC instruments. In particular, we show that the set of LOCC instruments that require no more than some fixed number of communication rounds is compact (Theorem 2.21). We also show that there exists a ball of LOCC instruments around the completely depolarizing instrument (Theorem 2.17).

The rest of this chapter is organized as follows. We start with the definition of general quantum instruments in Section 2.2.1. We proceed with the definitions of different classes of LOCC instruments in Section 2.2.2, as well as the related classes of separable and positive partial transpose-preserving instruments in Section 2.3. In Section 2.4 we discuss the relationships among these classes that follow from previously known results. In section Section 2.6 we prove Theorem 2.21 and Theorem 2.17. We conclude in Section 2.7 with a brief summary of results and a discussion of open problems.

2.2 How to define LOCC

2.2.1 Quantum instruments

Throughout this chapter we consider finite dimensional quantum systems. We use $\mathcal{H}$ to denote the associated complex vector space $\mathbb{C}^d$ and refer to it as the underlying state space of the system. Recall from Section 1.3.2 that $\mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)$ is the complex vector space consisting of the linear operators of the form $\mathcal{E} : \mathrm{L}(\mathcal{H}_1) \to \mathrm{L}(\mathcal{H}_2)$. We now define quantum instruments as discussed in [DL70].

Definition 2.1. An n-outcome *quantum instrument* $\mathfrak{J}$ is a tuple $(\mathcal{E}_1, \ldots, \mathcal{E}_n)$ of completely positive (CP) maps where $\mathcal{E}_i \in \mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)$ and the CP map $\sum_{j=1}^{n} \mathcal{E}_j$ is trace-preserving.

If we apply the instrument $\mathfrak{J}$ to a state ρ we obtain outcome j with probability $\mathrm{Tr}(\mathcal{E}_j(\rho))$ and the corresponding post-measurement state is given by $\mathcal{E}_j(\rho)$ up to normalization. To develop intuition for the not so commonly used notion of quantum instruments, we now explain how to think of measurements and completely positive trace-preserving (TCP) maps in terms of quantum instruments.

Example. Consider a measurement $\mathcal{M}$ specified by Kraus operators $M_1, \ldots, M_n$, where $\sum_{j\in[n]} M_j^\dagger M_j = I$. Upon measuring ρ with $\mathcal{M}$ and obtaining outcome j the resulting post-measurement state is $M_j \rho M_j^\dagger / \operatorname{Tr}(M_j^\dagger M_j \rho)$. The quantum instrument corresponding to this measurement is $(\mathcal{E}_1, \ldots, \mathcal{E}_n)$ where each CP map $\mathcal{E}_j(\rho) = M_j \rho M_j^\dagger$ has only one Kraus operator.

Example. Consider a TCP map $\mathcal{N}$ with Kraus operators $M_1, \ldots, M_n$, *i.e.*, the action of $\mathcal{N}$ is given by $\mathcal{N}(\rho) = \sum_{i \in [n]} M_i \rho M_i^\dagger$ where $[n] := \{1, \ldots, n\}$. The instrument corresponding to this TCP map is $(\mathcal{N})$ with only one CP map (which is necessarily trace-preserving).

We use $\mathrm{CP}[n, \mathcal{H}_1, \mathcal{H}_2] \subseteq \mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)^n$ to denote the set of instruments with index set $[n]$ that send quantum states from $\mathrm{L}(\mathcal{H}_1)$ to quantum states in $\mathrm{L}(\mathcal{H}_2)$. If the underlying state spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ have been fixed before, we simply write $\mathrm{CP}[n]$. Furthermore, if the index set is unimportant or implicitly clear from the context we often omit that as well and simply write CP.

Note that the set of quantum instruments $\mathrm{CP}[n, \mathcal{H}_1, \mathcal{H}_2]$ is convex, where addition of two instruments and multiplication by a real scalar is defined componentwise. If convenient, an instrument $\mathfrak{J} \in \mathrm{CP}[n, \mathcal{H}_1, \mathcal{H}_2]$ may naturally be viewed as an element of $\mathrm{CP}[m, \mathcal{H}_1, \mathcal{H}_2]$ with $m \geq n$ by padding $\mathfrak{J}$ with zero maps, *i.e.*, $\mathcal{E}_j = 0$ for $j > n$.

Let us fix the underlying input and output spaces to be $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. Given an index set $[n]$, we define a *quantum-classical* (QC) *map* over $[n]$ as the TCP map which sends $\mathrm{L}(\mathcal{H}_1)$ to $\mathrm{L}(\mathcal{H}_2) \otimes \mathrm{L}(\mathbb{C}^n)$ and is of the form

$$\rho \mapsto \sum_{j \in [n]} \mathcal{E}_j(\rho) \otimes |j\rangle\langle j|, \tag{2.1}$$

where the vectors $|j\rangle$ form the standard basis of $\mathbb{C}^n$. In this way, we see that the set $\mathrm{CP}[n, \mathcal{H}_1, \mathcal{H}_2]$ is in a one-to-one correspondence with the set of QC maps over $[n]$. For instrument $\mathfrak{J} = (\mathcal{E}_j : j \in [n])$, we denote its corresponding QC map by $\mathcal{E}[\mathfrak{J}](\cdot) = \sum_j \mathcal{E}_j(\cdot) \otimes |j\rangle\langle j|$.

We now introduce the concept of coarse graining that allows us to post-process the classical outcomes of one instrument, to obtain another instrument with a smaller number of outcomes.

Definition 2.2. (Coarse graining) Let $\mathfrak{J} = (\mathcal{E}_j : j \in [n])$ and $\mathfrak{L} = (\mathcal{F}_k : k \in [m])$ be quantum instruments with the same underlying input and output spaces and $n \geq m$. We say that $\mathfrak{L}$ is a *coarse graining* of $\mathfrak{J}$ if there exists a partition $\Lambda = (\Lambda_1, \ldots, \Lambda_m)$ of the index set $[n]$ such that

$$\mathcal{F}_k = \sum_{j \in \Lambda_k} \mathcal{E}_j \tag{2.2}$$

for all $k \in [m]$.

Equivalently, and perhaps more intuitively, one can describe coarse graining by the action of a *coarse graining map* $f : [n] \to [m]$, which is simply the function $f(k) = j$ for $k \in \Lambda_j$. In this picture we are using the coarse graining map f to post-process the classical information from the instrument $\mathfrak{J}$. Physically, this action corresponds to the discarding of classical information if the coarse graining is nontrivial. The fully coarse-grained instrument of $\mathfrak{J}$ corresponds to the TCP map $\sum_j \mathcal{E}_j$, obtained by tracing out the classical register of $\mathcal{E}[\mathfrak{J}]$. We say that an instrument $(\mathcal{E}_j : j \in [n])$ is *fine-grained* if each of the $\mathcal{E}_j$ has action of the form $\rho \mapsto M_j \rho M_j^\dagger$ for some operator M_j. In this way we see that the most general instrument can be implemented by performing a fine-grained instrument followed by coarse graining.

We now explain how to measure the distance between two instruments with the same underlying input and output spaces $\mathcal{H}_1, \mathcal{H}_2$ and the same index set $[n]$. The space $\mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)$ carries a metric, induced by the diamond norm $\|\cdot\|_\diamond$ [KSV02, Wat05b]. We use this metric and the L_1 norm to define a product metric $D_\diamond(\cdot, \cdot)$ on the n-fold product space $\mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)^n$. The distance between two instruments $\mathfrak{J} = (\mathcal{E}_j : j \in [n])$ and $\mathfrak{L} = (\mathcal{F}_j : j \in [n])$ is given by:

$$D_\diamond(\mathfrak{J}, \mathfrak{L}) := \sum_{j \in [n]} \|\mathcal{E}_j - \mathcal{F}_j\|_\diamond. \tag{2.3}$$

Since $\mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)^n$ is a finite dimensional complex vector space, all metrics defined on it are equivalent and give the same resulting topology (see *e.g.*, Theorem 8.7 in [BN00]). Therefore, any later discussed topological properties of the LOCC instruments do not depend on the specific metric we have chosen here.

We say that a sequence $\mathfrak{J}_1, \mathfrak{J}_2, \ldots$ of instruments from $\mathrm{CP}[n, \mathcal{H}_1, \mathcal{H}_2]$ converges to an instrument $\mathfrak{J} = (\mathcal{E}_j : j \in [n]) \in \mathrm{CP}[n, \mathcal{H}_1, \mathcal{H}_2]$, if

$$\lim_{\mu \to \infty} D_\diamond(\mathfrak{J}_\mu, \mathfrak{J}) = 0. \tag{2.4}$$

Since n is finite and fixed, the above is equivalent to requiring convergence in each coordinate. That is, for all $j \in [n]$ we require that $\lim_{\nu \to \infty} \|\mathcal{E}_{\nu,j} - \mathcal{E}_j\|_\diamond = 0$, where $\mathfrak{J}_\mu = (\mathcal{E}_{\mu,j} : j \in [n])$ for all μ.

We end this section by arguing that the set $\mathrm{CP}[n, \mathcal{H}_1, \mathcal{H}_2]$ forms a compact subset of the finite dimensional complex vector space $\mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)^n$. First, the set of all trace-nonincreasing maps $\mathcal{N} \in \mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)$ is compact. Next, any n-fold Cartesian product of compact sets from $(\mathrm{T}(\mathcal{H}_1, \mathcal{H}_2), \|\cdot\|_\diamond)$ gives a compact set in the product metric space

$(\mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)^n, D_\diamond)$. Finally, any compact subset of $\mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)^n$ remains compact if we restrict it to tuples $\big(\mathcal{E}_j : j \in [n]\big)$ for which the sum $\sum_{j\in[n]} \mathcal{E}_j$ is trace-preserving.

2.2.2 LOCC instruments

To discuss LOCC one has to fix the partition of the space. We do this by specifying the tensor product structure of the underlying input and output spaces beforehand. When we later refer to LOCC instruments it is with respect to this previously fixed partition.

For an N-partite quantum system, the underlying state space takes the form $\mathcal{H} := \mathcal{H}^{(1)} \otimes \cdots \otimes \mathcal{H}^{(N)}$, where party $K \in [N]$ holds the system corresponding to $\mathcal{H}^{(K)}$. For the remainder of this section, we fix the underlying input and output spaces to be $\mathcal{H}_1 := \mathcal{H}_1^{(1)} \otimes \cdots \otimes \mathcal{H}_1^{(N)}$ and $\mathcal{H}_2 := \mathcal{H}_2^{(1)} \otimes \cdots \otimes \mathcal{H}_2^{(N)}$, respectively. When convenient, we indicate which part of the space a map is acting on by specifying the number of the corresponding party in the superscript. In this way we can, for example, write $\mathcal{E}^{(2)} \otimes \mathcal{F}^{(1)}$ instead of $\mathcal{F} \otimes \mathcal{E}$.

Definition 2.3. We say that an instrument $\mathfrak{L} = (\mathcal{F}_1, \ldots, \mathcal{F}_n) \in \mathrm{CP}[n, \mathcal{H}_1, \mathcal{H}_2]$ is *one-way local* (with respect to party K) if each of its CP maps takes the form

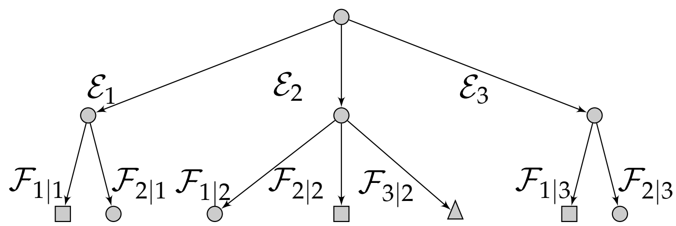
$$\mathcal{F}_j = \Big(\bigotimes_{J\neq K} \mathcal{T}_j^{(J)}\Big) \otimes \mathcal{E}_j^{(K)}, \tag{2.5}$$

where $\mathfrak{J} = (\mathcal{E}_1 \ldots, \mathcal{E}_n) \in \mathrm{CP}[n, \mathcal{H}_1^{(K)}, \mathcal{H}_2^{(K)}]$ and $\mathcal{T}_j^{(J)} \in \mathrm{T}(\mathcal{H}_1^{(J)}, \mathcal{H}_2^{(J)})$ is some TCP map (*i.e.*, an instrument with one outcome) for each $J \neq K$.

Operationally, this one-way local operation consists of party K applying an instrument $\mathfrak{J} = (\mathcal{E}_1, \ldots, \mathcal{E}_n)$, broadcasting the classical outcome j to all the other parties, and each party $J \neq K$ applying a TCP map $\mathcal{T}_j^{(J)}$ after receiving this information.

Now we formally describe which instruments can be obtained from some given instrument $\mathfrak{J}$ by one round of LOCC.

Definition 2.4. We say that an instrument $\mathfrak{L}$ is *LOCC linked* to $\mathfrak{J} = (\mathcal{E}_1, \ldots, \mathcal{E}_n)$ if there exists a collection of one-way local instruments $\{\mathfrak{J}_j = (\mathcal{F}_{1|j}, \ldots, \mathcal{F}_{m|j}) : j \in [n]\}$ such that $\mathfrak{L}$ is a coarse-graining of the instrument with CP maps $\mathcal{F}_{j'|j} \circ \mathcal{E}_j$.



Coarse-grained instrument:

$$\mathfrak{L} = (\mathcal{F}_{1|1} \circ \mathcal{E}_1 + \mathcal{F}_{2|2} \circ \mathcal{E}_2 + \mathcal{F}_{1|3} \circ \mathcal{E}_3, \\ \mathcal{F}_{2|1} \circ \mathcal{E}_1 + \mathcal{F}_{2|2} \circ \mathcal{E}_2,\ \mathcal{F}_{3|2} \circ \mathcal{E}_2)$$

(coarse grain over matching shapes)

Figure 2.1: The instrument $\mathfrak{L}$ is LOCC linked to the instrument $\mathfrak{J} = (\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3)$. The conditional instruments $(\mathcal{F}_{1|1}, \mathcal{F}_{2|1})$, $(\mathcal{F}_{1|2}, \mathcal{F}_{2|2}, \mathcal{F}_{3|2})$ and $(\mathcal{F}_{1|3}, \mathcal{F}_{2|3})$ composed with the three CP maps of $\mathfrak{J}$ yield the instrument $\mathfrak{L}$ after coarse graining.

Operationally, we first apply instrument $\mathfrak{J}$ and then, conditioned on the measurement outcome j, we apply instrument $\mathfrak{J}_j$ possibly followed by a coarse graining (see Fig. 2.1). Note that we allow the acting party in the conditional instrument $\mathfrak{J}_j$ to vary according to the outcome j of $\mathfrak{J}$. Also note that we can assume that the instruments $\mathfrak{J}_j$ all have the same number of outcomes, since otherwise we can pad the instruments with fewer outcomes with 0 maps.

Having introduced the above definitions, we now define different classes of LOCC instruments.

Definition 2.5. Given a quantum instrument $\mathfrak{J}$, we say that:

- $\mathfrak{J} \in \mathrm{LOCC}_1$ if $\mathfrak{J}$ is a coarse-graining of some one way local instrument $\mathfrak{L}$;
- $\mathfrak{J} \in \mathrm{LOCC}_r$ (with $r \geq 2$) if it is LOCC linked to some $\mathfrak{L} \in \mathrm{LOCC}_{r-1}$;
- $\mathfrak{J} \in \mathrm{LOCC}_{\mathbb{N}}$ if $\mathfrak{J} \in \mathrm{LOCC}_r$ for some $r \in \mathbb{N} := \{1, 2, \ldots\}$;
- $\mathfrak{J} \in \mathrm{LOCC}$ if there exists a sequence of $\mathrm{LOCC}_{\mathbb{N}}$ instruments $\mathfrak{J}_1, \mathfrak{J}_2, \ldots$ such that
  - (i) $\mathfrak{J}_\nu$ is LOCC linked to $\mathfrak{J}_{\nu-1}$,
  - (ii) each $\mathfrak{J}_\nu$ has a coarse-graining $\mathfrak{J}'_\nu$ such that the sequence $\mathfrak{J}'_1, \mathfrak{J}'_2, \ldots$ converges to $\mathfrak{J}$;
- $\mathfrak{J} \in \overline{\mathrm{LOCC}_{\mathbb{N}}}$ if there exists a sequence of $\mathrm{LOCC}_{\mathbb{N}}$ instruments $\mathfrak{J}_1, \mathfrak{J}_2, \ldots$ that converges to $\mathfrak{J}$.

Operationally, LOCC_r is the set of all instruments that can be implemented by some r-round LOCC protocol. Here, one round of communication involves one party communicating to all the others, and the sequence of communicating parties can depend on the intermediate measurement outcomes. The set of instruments that can be implemented by some finite round protocol is then $\text{LOCC}_{\mathbb{N}}$. We use the term *finite LOCC* to refer to this class of instruments. On the other hand, the so-called infinite round protocols, or those having an unbounded number of nontrivial communication rounds, correspond to instruments in $\text{LOCC} \setminus \text{LOCC}_{\mathbb{N}}$. The full set of LOCC then consists of both bounded-round protocols as well as unbounded ones [Chi11]. As the notation suggests, the class $\overline{\text{LOCC}_{\mathbb{N}}}$ is the topological closure of $\text{LOCC}_{\mathbb{N}}$. If we let $\overline{\text{LOCC}}$ be the closure of LOCC, then from the chain of inclusions $\text{LOCC}_{\mathbb{N}} \subseteq \text{LOCC} \subseteq \overline{\text{LOCC}_{\mathbb{N}}}$, we obtain $\overline{\text{LOCC}_{\mathbb{N}}} = \overline{\text{LOCC}}$. Sometimes the class $\overline{\text{LOCC}}$ is referred to as *asymptotic LOCC* [KKB11, CLMO12] and we follow this practice in the coming sections.

Both LOCC and $\overline{\text{LOCC}_{\mathbb{N}}}$ consist of instruments that can be approximated better and better with more LOCC rounds. The two sets are distinguished by noting that for any instrument in LOCC, its approximation in finite rounds can be made tighter by just continuing for more rounds within a *fixed* LOCC protocol; whereas for instruments in $\overline{\text{LOCC}_{\mathbb{N}}} \setminus \text{LOCC}$, different protocols will be needed for different degrees of approximation. Both LOCC and $\overline{\text{LOCC}}$ are operationally motivated. If it is very expensive to design the apparatus for implementing an LOCC protocol or it is necessary to control the precision of approximation in real time, then membership to LOCC is the relevant question. On the other hand, if one is willing to commit to a certain precision beforehand or redesign the physical apparatus to achieve a higher precision, then $\overline{\text{LOCC}}$ becomes the relevant class of instruments to consider.

Note that according to our definitions, every LOCC instrument is defined with respect to some fixed index set $[n]$. However, the instruments implemented during intermediate rounds of a protocol might range over different index sets. The requirement is that the intermediate instruments can each be coarse grained into an n-outcome instrument to form a convergent sequence of instruments. This coarse-graining need not correspond to an actual discarding of information. Indeed, discarding the measurement record midway through the protocol will typically prohibit the parties from completing the final LOCC instrument since the choice of measurement in each round depends on the full measurement history. On the other hand, often there will be an accumulation of classical data superfluous to the task at hand, and the parties will physically perform some sort of coarse graining (discarding of information) at the very end of the protocol.

Every instrument belonging to LOCC_r, LOCC or $\overline{\text{LOCC}}$ has an associated index set $[n]$. Thus, for fixed input and output spaces these operational classes naturally become subsets of $\text{CP}[n]$, and we denote them as $\text{LOCC}_r[n]$, $\text{LOCC}[n]$ and $\overline{\text{LOCC}}[n]$, respectively. If we wish to explicitly specify the partition of the input and output spaces we write, for example, $\text{LOCC}_r[n, \mathcal{H}_1^{(1)} : \cdots : \mathcal{H}_1^{(N)}, \mathcal{H}_2^{(1)} : \cdots : \mathcal{H}_2^{(N)}]$.

2.3 Separable and PPT instruments

To complete the picture from the previous section, we now provide definitions for the related classes of separable (SEP) and positive partial transpose-preserving (PPT) instruments. The PPT states and operations have been well-studied due to their relation to entanglement. In particular, bipartite entangled states with positive partial transpose are called bound entangled, because no LOCC protocol can be used to distill entanglement from them [HHH98].

Similar to the previous section we fix the underlying N-partite input and output spaces to be $\mathcal{H}_1 := \mathcal{H}_1^{(1)} \otimes \cdots \otimes \mathcal{H}_1^{(N)}$ and $\mathcal{H}_2 := \mathcal{H}_2^{(1)} \otimes \cdots \otimes \mathcal{H}_2^{(N)}$, respectively. As in the case of LOCC, to discuss SEP and PPT instruments we need to fix some partition of the space. Let us fix the partition to be $\mathcal{H}_1^{(1)} : \cdots : \mathcal{H}_1^{(N)}$ and $\mathcal{H}_2^{(1)} : \cdots : \mathcal{H}_2^{(N)}$ for the spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively.

An N-partite operator $\rho \in \mathrm{Pos}(\mathcal{H}_1)$ is said to be separable if it can be expressed as

$$\rho = \sum_i \rho_i^{(1)} \otimes \cdots \otimes \rho_i^{(N)}, \tag{2.6}$$

where each $\rho_i^{(K)} \in \mathrm{Pos}(\mathcal{H}_1^{(K)})$. Likewise, $\rho \in \mathrm{Pos}(\mathcal{H}_1)$ is said to possess the PPT property if the operator obtained by taking the partial transpose in any fixed basis with respect to any subset of the parties is positive semidefinite. Since conjugation by a unitary does not change the eigenvalues of a matrix, the PPT property is not affected by the choice of the basis used to define the transpose map. Note that any separable $\rho \in \mathrm{Pos}(\mathcal{H}_1)$ is necessarily PPT.

Intuitively, we want to define separable and PPT quantum operations to be the class of operations $\mathcal{E}$ that preserve the separability or the PPT property, respectively. However, as in the case of complete positivity, it is useful to introduce an auxiliary space and require that the map $\mathcal{E} \otimes \mathcal{I}$ also preserves the property in question. In this way, a tensor product of separable (PPT) maps is always a separable (PPT) map.

Let $\mathcal{H}'_1 := \mathcal{H}'^{(1)}_1 \otimes \cdots \otimes \mathcal{H}'^{(N)}_1$ be a copy of the space $\mathcal{H}_1$ and let $\mathcal{I}$ be the identity map on $\mathcal{H}'_1$. Further let $\mathcal{X}_t^{(K)} := \mathcal{H}_t^{(K)} \otimes \mathcal{H}'^{(K)}_1$ for all $t \in \{1,2\}$ and $K \in [N]$. If we consider quantum maps of the form $\mathcal{E} \otimes \mathcal{I}$ for $\mathcal{E} \in \mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)$, then $\mathcal{X}_1^{(K)}$ and $\mathcal{X}_2^{(K)}$ are the Kth party's input and output spaces, respectively.

Definition 2.6. Let $\mathfrak{J} = (\mathcal{E}_j : j \in [n]) \in \mathrm{CP}[n, \mathcal{H}_1, \mathcal{H}_2]$ be an instrument. We say that

- $\mathfrak{J}$ is *separable* and write $\mathfrak{J} \in \mathrm{SEP}[n]$ if each $\mathcal{E}_j$ is a separable map [Rai97], meaning that $(\mathcal{E}_j \otimes \mathcal{I})\rho$ is separable with respect to $\mathcal{X}_2^{(1)} : \cdots : \mathcal{X}_2^{(N)}$ whenever ρ is separable with respect to $\mathcal{X}_1^{(1)} : \cdots : \mathcal{X}_1^{(N)}$;
- $\mathfrak{J}$ is *positive partial transpose-preserving* and write $\mathfrak{J} \in \mathrm{PPT}[n]$ if each $\mathcal{E}_j$ is a PPT map [Rai99, DLT02], meaning that $(\mathcal{E}_j \otimes \mathcal{I})\rho$ is PPT with respect to $\mathcal{X}_2^{(1)} : \cdots : \mathcal{X}_2^{(N)}$ whenever ρ is PPT with respect to $\mathcal{X}_1^{(1)} : \cdots : \mathcal{X}_1^{(N)}$.

If we want to explicitly specify the underlying partition of the space we write

$$\mathfrak{J} \in \mathrm{SEP}[n, \mathcal{H}_1^{(1)} : \cdots : \mathcal{H}_1^{(N)}, \mathcal{H}_2^{(1)} : \cdots : \mathcal{H}_2^{(N)}] \text{ and} \tag{2.7}$$
$$\mathfrak{J} \in \mathrm{PPT}[n, \mathcal{H}_1^{(1)} : \cdots : \mathcal{H}_1^{(N)}, \mathcal{H}_2^{(1)} : \cdots : \mathcal{H}_2^{(N)}] \tag{2.8}$$

respectively.

For any two complex vector spaces $\mathcal{H}_1, \mathcal{H}_1'$ of the same dimension, let

$$|\Phi_{\mathcal{H}_1:\mathcal{H}_1'}\rangle := \sum_{t \in [\dim(\mathcal{H}_1)]} |t, t\rangle \quad \text{and} \quad \Phi_{\mathcal{H}_1:\mathcal{H}_1'} := |\Phi_{\mathcal{H}_1:\mathcal{H}_1'}\rangle\langle\Phi_{\mathcal{H}_1:\mathcal{H}_1'}|. \tag{2.9}$$

When the spaces $\mathcal{H}_1$ and $\mathcal{H}_1'$ possess the above specified tensor product structure we have

$$\Phi_{\mathcal{H}_1:\mathcal{H}_1'} = \bigotimes_{K \in [N]} \Phi_{\mathcal{H}_1^{(K)}:\mathcal{H}_1'^{(K)}}, \tag{2.10}$$

where $\Phi_{\mathcal{H}_1^{(K)}:\mathcal{H}_1'^{(K)}} \in \mathrm{Pos}(\mathcal{X}^{(K)})$ and $\mathcal{X}^{(K)} = \mathcal{H}_1^{(K)} \otimes \mathcal{H}_1'^{(K)}$ Thus, $\Phi_{\mathcal{H}_1:\mathcal{H}_1'}$ is both separable and PPT with respect to the partition $\mathcal{X}_1^{(1)} : \cdots : \mathcal{X}_1^{(N)}$. Now from the definition we see that the Choi matrix [Jam72, Cho75]

$$J_{\mathcal{E}} = (\mathcal{E} \otimes \mathcal{I})\Phi_{\mathcal{H}_1:\mathcal{H}_1'} \in \mathrm{Pos}(\mathcal{H}_2 \otimes \mathcal{H}_1') \tag{2.11}$$

of a separable (PPT) map $\mathcal{E} \in \mathrm{CP}(\mathcal{H}_1, \mathcal{H}_2)$ is always separable (PPT). In fact, the converse is also true. That is, if the Choi matrix $J_{\mathcal{E}}$ of a CP map $\mathcal{E}$ is separable (PPT) then the map $\mathcal{E}$ is also separable (PPT) [CDKL01]. Since any PPT operator ρ is separable, the characterization of separable and PPT maps in terms of their Choi matrices shows that $\mathrm{SEP}[n] \subseteq \mathrm{PPT}[n]$.

Both the set of separable operators and that of PPT ones are known to be closed. The characterization of separable and PPT maps in terms of their Choi matrices allows us to reach the same conclusion about the sets of trace nonincreasing separable and PPT maps. Any n-fold Cartesian product of closed sets from $(\mathrm{T}(\mathcal{H}_1, \mathcal{H}_2), \|\cdot\|_\diamond)$ gives a closed set in the product metric space $(\mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)^n, D_\diamond)$. Finally, any closed subset of $\mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)^n$ remains closed if we restrict to tuples $\big(\mathcal{E}_j : j \in [n]\big)$ for which the sum $\sum_{j\in[n]} \mathcal{E}_j$ is trace-preserving. Hence, the sets $\mathrm{SEP}[n]$ and $\mathrm{PPT}[n]$ are closed.

We now characterize separable maps in terms of their Kraus decompositions. If $\mathcal{E} \in \mathrm{CP}(\mathcal{H}_1, \mathcal{H}_2)$ is separable, then so is its Choi matrix $J_\mathcal{E} \in \mathrm{Pos}(\mathcal{H}_2 \otimes \mathcal{H}_1')$. Thus, for an appropriate choice of vectors $|v_t^{(K)}\rangle \in \mathcal{H}_2^{(K)} \otimes \mathcal{H}_1'^{(K)}$ we have

$$J_\mathcal{E} = \sum_t |v_t^{(1)}\rangle\langle v_t^{(1)}| \otimes \cdots \otimes |v_t^{(N)}\rangle\langle v_t^{(N)}| \tag{2.12}$$

$$= \sum_t \bigotimes_{K\in[N]} \big(V_t^{(K)} \otimes \mathcal{I}_{\mathcal{H}_1'^{(K)}}\big) \Phi_{\mathcal{H}_1^{(K)}:\mathcal{H}_1'^{(K)}} \big(V_t^{(K)} \otimes \mathcal{I}_{\mathcal{H}_1'^{(K)}}\big)^\dagger \tag{2.13}$$

$$= \sum_t \big(V_t^{(1)} \otimes \cdots \otimes V_t^{(N)} \otimes \mathcal{I}_{\mathcal{H}_1'}\big) \Phi_{\mathcal{H}_1:\mathcal{H}_1'} \big(V_t^{(1)} \otimes \cdots \otimes V_t^{(N)} \otimes \mathcal{I}_{\mathcal{H}_1'}\big)^\dagger, \tag{2.14}$$

where $|v_t^{(K)}\rangle = \mathrm{vec}\,\big(V_t^{(K)}\big)$ and $\mathrm{vec} : \mathrm{L}(\mathcal{H}_1^{(K)}, \mathcal{H}_2^{(K)}) \to \mathcal{H}_2^{(K)} \otimes \mathcal{H}_1'^{(K)}$ is the row vectorization defined via $\mathrm{vec}(|i\rangle\langle j|) := |i\rangle|j\rangle$. So we see that any separable map $\mathcal{E}$ admits a Kraus decomposition with tensor product Kraus operators. It also easily follows from the definition that whenever a decomposition of this form exists, the map $\mathcal{E}$ is necessarily separable. This completes the characterization of separable maps in terms of their Kraus decompositions. We are not aware of any such characterization for PPT maps.

Note that any instrument $\mathfrak{J} \in \mathrm{LOCC}_\mathbb{N}[n]$ can be viewed as a coarse-graining of some fine-grained instrument $\mathfrak{L} \in \mathrm{LOCC}_\mathbb{N}[m]$, where $m \geq n$. Recall that each of the CP maps $\mathcal{F}_j$ of $\mathfrak{L}$ has only one Kraus operator M_j. By the construction of $\mathrm{LOCC}_\mathbb{N}$ in Definition 2.5 we see that M_j is an N-fold tensor product. So each of the CP maps of $\mathfrak{L}$ and hence also of $\mathfrak{J}$ has a Kraus representation with tensor-product operators. Thus, $\mathfrak{J} \in \mathrm{SEP}[n]$ in particular and $\mathrm{LOCC}_\mathbb{N}[n] \subseteq \mathrm{SEP}[n]$ in general. Since the set of separable instruments is closed, we also have $\overline{\mathrm{LOCC}}[n] \subseteq \mathrm{SEP}[n]$. Combining this with the previously discussed inclusion $\mathrm{SEP}[n] \subseteq \mathrm{PPT}[n]$, we get $\overline{\mathrm{LOCC}}[n] \subseteq \mathrm{SEP}[n] \subseteq \mathrm{PPT}[n]$.

Although the classes of SEP and PPT instruments are more powerful than LOCC ones, they are still more restrictive than the most general quantum instruments. At the same time, they admit a simpler mathematical characterization than LOCC, and this can be used to derive many limitations on LOCC, such as entanglement distillation [Rai97, Rai01, HHH98] and state discrimination [TDL01, Che04].

We end this section by showing that any separable map $\mathcal{E} \in \mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)$ admits a Kraus representation with no more than $(\dim(\mathcal{H}_1)\dim(\mathcal{H}_2))^2$ elements each of which is a tensor product. To prove this, we use Carathéodory's theorem (see *e.g.* [Roc96]):

Theorem 2.7 (Carathéodory)**.** *Let S be a subset of $\mathbb{R}^d$ for some $d \in \mathbb{Z}_+$. Then any element $v \in \mathrm{conv}(S)$ can be expressed as a convex combination of $m \leq d+1$ vectors $v_1, \ldots, v_m \in S$.*

Let us first give a simple lemma bounding the number of terms in a decomposition certifying the separability of an operator from $\mathrm{Pos}(\mathcal{H})$.

Lemma 2.8. *Let $M \in \mathrm{Pos}(\mathcal{H})$, where $\mathcal{H} = \mathcal{H}^{(1)} \otimes \ldots \otimes \mathcal{H}^{(N)}$ and $\dim(\mathcal{H}) = d$. If M is separable with respect to partition $\mathcal{H}^{(1)} : \cdots : \mathcal{H}^{(N)}$, then it can be expressed as*

$$M = \sum_{i \in [m]} M_i^{(1)} \otimes \ldots \otimes M_i^{(N)}, \tag{2.15}$$

for some rank one matrices $M_i^{(K)} \in \mathrm{Pos}(\mathcal{H}^{(K)})$ and $m \leq \mathrm{rank}(M)^2$.

Proof. Let V be the set of Hermitian operators on $\mathcal{H}$ having the same trace as M and whose column space is contained in that of M. Note that V can be viewed as a real vector space of dimension $\mathrm{rank}(M)^2 - 1$. Next, since M is separable, we can write it as

$$M = \sum_{i \in [t]} \lambda_i N_i^{(1)} \otimes \ldots \otimes N_i^{(N)}, \tag{2.16}$$

for some $t \in \mathbb{Z}_+$, $\lambda_i > 0$, and some rank one matrices $N_i^{(K)} \in \mathrm{Pos}(\mathcal{H}^{(K)})$. Let $N_i := N_i^{(1)} \otimes \ldots \otimes N_i^{(N)}$. Without loss of generality, we can assume that $\mathrm{Tr}(N_i) = \mathrm{Tr}(M)$. In such a case $\sum_{i \in [T]} \lambda_i = 1$. Moreover, we can argue that the column space of each N_i must be contained in that of M. If it was not the case, we could find a vector $v \in \mathrm{Null}(M)$ that does not belong to $\mathrm{Null}(N_i)$. Yet this would contradict Equation (2.16), where we have expressed M as a positive linear combination of the positive semidefinite matrices N_i. So $N_i \in V$ for all $i \in [t]$ and thus M is a convex combination of t elements from V. Hence, if $t > \mathrm{rank}(M)^2$ we can apply Carathéodory's theorem to reduce the number of summands in Equation (2.16) and write M as

$$M = \sum_{i \in R} \mu_i N_i^{(1)} \otimes \ldots \otimes N_i^{(N)}, \tag{2.17}$$

where $R \subseteq [t]$ and $|R| \leq \mathrm{rank}(M)^2$. Finally, absorbing the coefficients μ_i into the matrices $N_i^{(1)} \otimes \ldots \otimes N_i^{(N)}$ gives an expression of the desired form (2.15). $\square$

Having established the above lemma, we are ready to bound the number of operators in a minimal tensor product Kraus representation of a separable map.

Corollary 2.9. *Consider a CP map $\mathcal{E} \in \mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)$ with Choi matrix $J_{\mathcal{E}}$. Let $d_1 := \dim(\mathcal{H}_1)$ and $d_2 := \dim(\mathcal{H}_2)$. If $\mathcal{E}$ is separable then it admits a Kraus representation of the form*

$$\{M_i^{(1)} \otimes \ldots \otimes M_i^{(N)}\}_{i \in [m]} \tag{2.18}$$

with $m \leq \mathrm{rank}(J_{\mathcal{E}})^2 \leq (d_1 d_2)^2$.

Proof. Consider the Choi matrix $J_{\mathcal{E}} \in \mathrm{Pos}(\mathcal{H}_2 \otimes \mathcal{H}'_1)$. Since $\mathcal{E}$ is separable, so must be $J_{\mathcal{E}}$. Hence, according to Lemma 2.8, we can express it as

$$J_{\mathcal{E}} = \sum_{i \in [m]} J_i^{(1)} \otimes \ldots \otimes J_i^{(N)} \tag{2.19}$$

for some rank one matrices $J_i^{(N)} \in \mathrm{Pos}(\mathcal{H}_2^{(K)} \otimes \mathcal{H'}_1^{(K)})$ and $m \leq \mathrm{rank}(J_{\mathcal{E}})^2$. From here, via a calculation similar to the one in Equations (2.12)–(2.14), we obtain the desired tensor product Kraus representation with no more than $\mathrm{rank}(J_{\mathcal{E}})^2$ operators. □

2.4 Relationships between different classes of LOCC

In this section we discuss the known relationships between the different classes of instruments introduced in the previous two sections. For a fixed number of parties $N \geq 2$ with local dimension at least two, and over a sufficiently large index set $[n]$, the known relationships can be summarized as follows:

$$\mathrm{LOCC}_1 \subsetneq \mathrm{LOCC}_r \subsetneq \mathrm{LOCC}_{r+1} \subsetneq \mathrm{LOCC}_{\mathbb{N}} \subsetneq \mathrm{LOCC} \subsetneq \overline{\mathrm{LOCC}} \subsetneq \mathrm{SEP} \subsetneq \mathrm{PPT} \tag{2.20}$$

for any $r \geq 2$. Before we explain why the above inclusions are proper, let us prove the following lemma.

Lemma 2.10. *If $\mathrm{LOCC}_r \subsetneq \mathrm{LOCC}_{r+k}$ for some $k \in \mathbb{Z}_+$ then $\mathrm{LOCC}_r \subsetneq \mathrm{LOCC}_{r+1}$.*

Proof. Suppose that $\mathrm{LOCC}_r = \mathrm{LOCC}_{r+1}$, and consider $\mathfrak{J} \in \mathrm{LOCC}_{r+k} \setminus \mathrm{LOCC}_r$. Then there exists an implementation of $\mathfrak{J}$ consuming $r + k$ rounds. Let $\mathfrak{J}_{r+1}$ being the instrument performed during the first $r + 1$ rounds of this particular implementation. Since $\mathrm{LOCC}_r = \mathrm{LOCC}_{r+1}$, we have $\mathfrak{J}_{r+1} \in \mathrm{LOCC}_r$ and so $\mathfrak{J} \in \mathrm{LOCC}_{r+k-1}$. Here, we have considered n sufficiently large such that both $\mathfrak{J}$ and $\mathfrak{J}_{r+1}$ are instruments over the same index set (this can always be done by Theorem 2.20). Repeating this argument k times gives that $\mathfrak{J} \in \mathrm{LOCC}_r$, which is a contradiction. □

We now explain why the inclusions in Equation (2.20) are proper. The operational advantage of LOCC_2 over LOCC_1 is well-known, having been observed in entanglement distillation [BDSW96], quantum cryptography [GL03], and state discrimination [Coh07, OH08]. On the other hand, only a few examples have been proven to demonstrate the separation between LOCC_r and LOCC_{r+1}. For $N = 2$, Xin and Duan have constructed sets containing $O(n^2)$ mutually orthogonal pure states in two n-dimensional systems that require $O(n)$ rounds of LOCC to distinguish perfectly [XD08]. For $N \geq 3$, a stronger separation is shown for random distillation of bipartite entanglement from a three-qubit state. Here the dimension is fixed, and two extra LOCC rounds can always increase the probability of success by a quantifiable amount; moreover, certain distillations only become possible by infinite-round LOCC, thus demonstrating $\text{LOCC}_{\mathbb{N}} \neq \text{LOCC}$ [Chi11]. By studying the same random distillation problem, one can show that $\text{LOCC} \neq \overline{\text{LOCC}}$ [CCL12a, CCL12b]. For $N = 2$ and local dimension two, it has been shown that there exists an instrument $\mathfrak{J} \notin \text{LOCC}$ that can nevertheless be approximated to arbitrary precision with instruments from $\text{LOCC}_{\mathbb{N}}$ [CLM+12]. So there are instruments that require different protocols to achieve better and better approximations when more and more LOCC rounds are available, and neither $\text{LOCC}_{\mathbb{N}}$ nor LOCC is closed. The difference between $\text{LOCC}_{\mathbb{N}}$ and SEP has emerged in various problems such as state discrimination [BDM+99, Coh07, DFXY09] and entanglement transformations [CD09]. Proving that $\overline{\text{LOCC}} \neq \text{SEP}$ is more difficult, but it indeed has been demonstrated in [BDF+99, KTYI07, CLMO12] for the task of state discrimination, as well as [CCL11, CCL12a, CCL12b] for random distillation. Finally, the strict inclusion between PPT and SEP follows from the existence of non-separable states that possess a positive partial transpose [Hor97].

2.5 Stochastic LOCC

Although $\text{LOCC} \subsetneq \text{SEP}$, it turns out that it is always possible to implement a separable instrument with some nonzero probability of success using LOCC. We formalize the notion of probabilistic implementation in the following definition. We use $\mathcal{D} \in \mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)$ to denote the completely depolarizing channel, *i.e.*, the TCP map which acts by $\mathcal{D}(\rho) = \frac{\mathrm{Tr}(\rho)}{\dim(\mathcal{H}_2)} I_{\mathcal{H}_2}$.

Definition 2.11. Let $\mathfrak{J} = (\mathcal{E}_1, \ldots, \mathcal{E}_n) \in \text{SEP}[n, \mathcal{H}_1\mathcal{H}_2]$ be an instrument. We say that $\mathfrak{J}$ can be performed by *Stochastic LOCC* (SLOCC) if for some $p > 0$, the instrument

$$\mathfrak{L} = (p\mathcal{E}_1, \ldots p\mathcal{E}_n, (1-p)\mathcal{D}) \in \text{CP}[n+1, \mathcal{H}_1, \mathcal{H}_2] \tag{2.21}$$

belongs to the class of LOCC instruments.

The classical outcome "$n+1$" indicates a failure in implementing $\mathfrak{J}$ and it occurs with probability $1-p$. Hence, we are always aware of whether or not we have succeeded at implementing $\mathfrak{J}$. Also, since $\mathcal{D} \circ \mathcal{E} = \mathcal{D}$ for every TCP map $\mathcal{E}$, there is no loss of generality in assuming that in the case of failure, we can implement the maximally depolarizing channel.

It has been shown in [DVC00] that every separable CP map has an SLOCC implementation. Here we describe a finite LOCC protocol for implementing any separable instrument, thus providing a lower bound on the success probability p. Our construction follows from quantum teleportation [BBC$^+$93]. As before let $\mathcal{H}_1^{(1)} : \cdots : \mathcal{H}_1^{(N)}$ and $\mathcal{H}_2^{(1)} : \cdots : \mathcal{H}_2^{(N)}$ be the fixed partition of the input and output spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively.

Lemma 2.12. *Let $\mathfrak{J} = (\mathcal{E}_1, \ldots, \mathcal{E}_n) \in CP[n, \mathcal{H}_1, \mathcal{H}_2]$ be a separable N-partite instrument and let $d_t := \dim \mathcal{H}_t$ for $t = 1, 2$. Then $\mathfrak{J}$ can be implemented by* SLOCC *with success probability of at least $\frac{1}{d_1^2}$ using a protocol with at most N communication rounds.*

Proof. Suppose that N parties wish to measure some shared state $\rho \in \mathcal{H}_1$ using the separable instrument $\mathfrak{J}$. First, using one round of LOCC, they choose $j \in [n]$ with probability $\mathrm{Tr}(\mathcal{E}_j(I))/d_1$ and prepare the separable state $\rho_j := \frac{1}{\mathrm{Tr}(\mathcal{E}_j(I))} J_{\mathcal{E}_j}$. Here, $J_{\mathcal{E}_j}$ is the Choi matrix of the CP map $\mathcal{E}_j$. Next, each of the parties measure their part of the last two registers of $\rho_j \otimes \rho \in \mathrm{L}(\mathcal{H}_2 \otimes \mathcal{H}_1' \otimes \mathcal{H}_1)$ in a basis containing the canonical maximally entangled state $\frac{1}{\sqrt{d_1}} |\Phi_{\mathcal{H}'^{(K)}_1 : \mathcal{H}^{(K)}_1}\rangle$ (see Equation (2.9)).

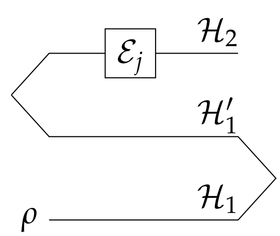


Figure 2.2: Contraction diagram of $\mathcal{E}_j(\rho)$.

From the diagram in Figure 2.2, we see that

$$\mathcal{E}_j(\rho) = (I \otimes \langle \Phi_{\mathcal{H}_1' : \mathcal{H}_1}|)(J_{\mathcal{E}_j} \otimes \rho)(I \otimes |\Phi_{\mathcal{H}_1' : \mathcal{H}_1}\rangle) \tag{2.22}$$

$$= \mathrm{Tr}\left(\mathcal{E}_j(I)\right) d_1 \Big(I \otimes \tfrac{1}{\sqrt{d_1}} \langle \Phi_{\mathcal{H}_1' : \mathcal{H}_1}|)(\rho_j \otimes \rho)(I \otimes \tfrac{1}{\sqrt{d_1}} |\Phi_{\mathcal{H}_1' : \mathcal{H}_1}\rangle \Big). \tag{2.23}$$

Hence, the probability that all the parties obtain the outcome corresponding to the canonical maximally entangled state is given by $\frac{1}{\mathrm{Tr}(\mathcal{E}_j(I))d_1}\,\mathrm{Tr}\left(\mathcal{E}_j(\rho)\right)$. In this case they have succeeded at applying CP map $\mathcal{E}_j$ to ρ. Overall, the described protocol succeeds at implementing $\mathfrak{J}$ with probability $\frac{1}{d_1^2}$. It is easy to see that this protocol takes N rounds of classical communication to broadcast the measurement outcomes of each of the N parties. $\square$

We now give an alternative protocol for implementing a separable instrument $\mathfrak{J} = (\mathcal{E}_1, \ldots, \mathcal{E}_n)$. This protocol achieves a success probability of $\frac{1}{nr^2}$, where

$$r := \max_{i\in[n]} \mathrm{rank}(J_{\mathcal{E}_i}). \tag{2.24}$$

Therefore, if all of the Choi matrices $J_{\mathcal{E}_i}$ have low rank, it can outperform the protocol from Lemma 2.12. In contrast, if at least one of the Choi matrices has full rank, it gives a success probability of only $\frac{1}{n(d_1d_2)^2}$. We now describe the protocol.

For each $j \in [n]$ let $\{M_{ij}^1 \otimes \ldots \otimes M_{ij}^N\}_{i\in[r^2]}$ be the set of operators in some Kraus representation of $\mathcal{E}_j$ with r^2 elements. Since we can always increase the number of operators in a Kraus representation by adding zero operators, the existence of such a representation follows from Corollary 2.9. For each Kraus operator $M_{ij}^1 \otimes \ldots \otimes M_{ij}^N$, we choose the individual matrices M_{ij}^K so that $\|M_{ij}^1\|_\infty = \ldots = \|M_{ij}^N\|_\infty$. This ensures that $I \geq (M_{ij}^K)^\dagger M_{ij}^K$ and thus

$$\mathcal{M}_{ij}^K := \left\{M_{ij}^K, \sqrt{I - (M_{ij}^K)^\dagger M_{ij}^K}\right\} \tag{2.25}$$

is a valid local measurement for each party K. The protocol consists of the parties first collectively choosing a pair $(i,j) \in [r^2] \times [n]$ uniformly at random. They then take turns to perform their respective local measurements $\mathcal{M}_{ij}^K$ and broadcast their result. If all parties obtain the outcome corresponding to M_{ij}^K, their implementation is a success and they coarse grain the classical data (i,j) over the index i. This coarse graining recovers the CP maps $\mathcal{E}_j$. If at least one party obtains the second outcome, all the parties apply local maximally depolarizing channels and this is a failure outcome. Coarse graining over all failure outcomes generates the LOCC instrument

$$\left(\frac{1}{nr^2}\mathcal{E}_1, \ldots, \frac{1}{nr^2}\mathcal{E}_n, \frac{nr^2-1}{nr^2}\mathcal{D}\right). \tag{2.26}$$

Note that in the initial round of the protocol, some fixed party K can first choose the random pair $(i, j) \in [r^2] \times [n]$, then measure $\mathcal{M}_{ij}^K$ and broadcast both the chosen pair (i, j) and the outcome of the measurement. Hence, the described protocol requires N rounds of communication.

2.6 Some topological properties

In this section we discuss some topological properties of different classes of LOCC instruments. As before, we let $\mathcal{H}_1$ be the input and $\mathcal{H}_2$ be the the output space. We also fix $\mathcal{H}_1^{(1)} : \cdots : \mathcal{H}_1^{(N)}$ and $\mathcal{H}_2^{(1)} : \cdots : \mathcal{H}_2^{(N)}$ to be their respective partitions.

In the coming subsections we consider the instruments that can be implemented using at most r rounds of communication, *i.e.*, the class LOCC_r for some finite r. However, we first give the following simple observation concerning the other previously introduced classes of LOCC instruments.

Theorem 2.13. *For any $n \in \mathbb{Z}_+$ the sets $\mathrm{LOCC}_{\mathbb{N}}[n], \mathrm{LOCC}[n]$, and $\overline{\mathrm{LOCC}}[n]$ are convex.*

Proof. Fix any $n \in \mathbb{Z}_+$. Given two instruments $\mathfrak{J}_1, \mathfrak{J}_2 \in \mathrm{LOCC}[n]$ and $\lambda \in [0, 1]$, consider the instrument $\mathfrak{J} := \lambda\mathfrak{J}_1 + (1 - \lambda)\mathfrak{J}_2$. To implement $\mathfrak{J}$, one of the parties starts the LOCC protocol by choosing $v = 1$ with probability λ and $v = 2$ with probability $1 - \lambda$ and broadcasting this choice to the other parties. Upon receiving the value v, the parties proceed according to some LOCC protocol for implementing $\mathfrak{J}_v$. This shows that $\mathrm{LOCC}[n]$ is convex. Since finite LOCC protocols for implementing $\mathfrak{J}_1$ and $\mathfrak{J}_2$ yield a finite LOCC protocol for implementing $\mathfrak{J}$, we conclude that $\mathrm{LOCC}_{\mathbb{N}}[n]$ is also convex. Finally, since the closure of a convex set is convex, we get that $\overline{\mathrm{LOCC}}[n]$ is convex. □

In contrast to the above, $\mathrm{LOCC}_1[n]$ for $n \geq 2$ is not convex. To see this, suppose that $\mathfrak{J}_v$ for $v \in \{1, 2\}$ corresponds to the v-th party performing a standard basis measurement on their state space $\mathcal{H}_1^{(K)}$. Then for any $\lambda \in (0, 1)$ the instrument $\lambda\mathfrak{J}_1 + (1 - \lambda)\mathfrak{J}_2 \notin \mathrm{LOCC}_1[n]$, since LOCC_1 essentially consists of instruments that are one-way local with respect to some fixed acting party (see Definition 2.5 and Definition 2.3).

2.6.1 A ball of LOCC_r instruments

In this section we establish that for any $r, n \in \mathbb{Z}_+$ the set $\mathrm{LOCC}_r[n]$ has nonempty interior within the set of all quantum instruments $\mathrm{CP}[n]$. To prove this, we show that there exists a non-empty ball consisting entirely of LOCC instruments around the completely depolarizing instrument[2]

$$\mathfrak{D} := \left(\frac{1}{n}\mathcal{D}, \ldots, \frac{1}{n}\mathcal{D}\right). \tag{2.27}$$

Here, $\mathcal{D} \in \mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)$ is the completely depolarizing channel as defined in Section 2.5. Operationally this means that any sufficiently noisy instrument can be implemented with LOCC. Also, it provides a (probably suboptimal) upper bound for the distance between any fixed instrument $\mathfrak{J}$ and the set $\overline{\mathrm{LOCC}}$.

Our argument will rely on the following result.

Theorem 2.14 (Gurvits and Barnum [GB03]). *Let $\mathcal{H} = \mathcal{H}^{(1)} \otimes \cdots \otimes \mathcal{H}^{(N)}$. If $A \in \mathrm{L}(\mathcal{H})$ and $\|A\|_2 \leq R_N := 2^{1-N/2}$, then $I + A$ is a separable operator with respect to the partition $\mathcal{H}_1^{(1)} : \cdots : \mathcal{H}_1^{(N)}$.*

As discussed in Section 2.3, a CP map $\mathcal{E} \in \mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)$ is separable if and only if its Choi matrix $J_{\mathcal{E}} \in \mathrm{L}(\mathcal{H}_2 \otimes \mathcal{H}_1')$ is separable. Using this fact, we can translate the above theorem into the following statement about separable maps.

Corollary 2.15. *Consider an instrument $\mathfrak{J} = (\mathcal{E}_1, \ldots, \mathcal{E}_n) \in CP[n]$. Then*

$$(1-\eta)\mathfrak{D} + \eta\mathfrak{J} \in \mathrm{SEP}[n] \tag{2.28}$$

for all $\eta \leq R_{\mathrm{SEP}} := R_N/(nd_1d_2 + R_N) = (nd_1d_2 2^{\frac{N}{2}-1} + 1)^{-1}$.

Proof. By the above discussion, it suffices to show that the Choi matrices of the CP maps of the resulting instrument in Equation (2.28) are all separable. For the completely depolarizing map $\mathcal{D}$, we have $J_{\mathcal{D}} = \frac{1}{d_2} I_{d_1 d_2}$, where $d_1 := \dim(\mathcal{H}_1)$ and $d_2 := \dim(\mathcal{H}_2)$. Hence, the Choi matrix of the jth CP map of the instrument in Equation (2.28) is

$$(1-\eta)\frac{I_{d_1 d_2}}{nd_2} + \eta J_{\mathcal{E}_j}. \tag{2.29}$$

According to Theorem 2.14 we have that

$$I_{d_1 d_2} + \frac{\eta n d_2}{1-\eta} J_{\mathcal{E}_i} \tag{2.30}$$

[2] Via private communication, we have learned that Marco Piani has independently obtained a similar result for TCP maps.

is separable whenever $\frac{\eta n d_2}{1-\eta}\|J_{\mathcal{E}_i}\|_2 \leq R_N$. Note that $\|J_{\mathcal{E}_i}\|_2 \leq \|J_{\mathcal{E}_i}\|_1 = d_1$. Therefore, choosing

$$\eta \leq \frac{R_N}{nd_1d_2 + R_N} \tag{2.31}$$

yields the desired statement. □

Combining this corollary with the observation that any separable instrument can be implemented with stochastic LOCC (see Lemma 2.12) gives the following theorem.

Theorem 2.16. *Consider an instrument $\mathfrak{J} = (\mathcal{E}_1, \ldots, \mathcal{E}_n) \in CP[n]$. Then*

$$(1-\gamma)\mathfrak{D} + \gamma\mathfrak{J} \in \mathrm{LOCC}_N[n] \tag{2.32}$$

for all $\gamma \leq \frac{R_{\mathrm{SEP}}}{d_1^2} = \left(nd_1^3 d_2 2^{\frac{N}{2}-1} + d_1^2\right)^{-1}$.

Proof. Pick any $\gamma \leq R_{\mathrm{SEP}}/d_1^2$ and choose $\eta \leq R_{\mathrm{SEP}}$ and $p \leq \frac{1}{d_1^2}$ such that $\gamma = p\eta$. First, according to Corollary 2.15 the instrument

$$\mathfrak{L} := (1-\eta)\mathfrak{D} + \eta\mathfrak{J} \in \mathrm{SEP}[n] \tag{2.33}$$

since $\eta \leq R_{\mathrm{SEP}}$. Next, by Lemma 2.12 it follows that there exists an N-round LOCC protocol that successfully implements $\mathfrak{L}$ with probability $p \leq \frac{1}{d_1^2}$ and depolarizes (fails) with probability $1-p$. By coarse graining the failure outcome into each of the n success outcomes by equal amounts, we have

$$(1-p)\mathfrak{D} + p\mathfrak{L} = (1-p\eta)\mathfrak{D} + p\eta\mathfrak{J} = (1-\gamma)\mathfrak{D} + \gamma\mathfrak{J} \in \mathrm{LOCC}_N[n], \tag{2.34}$$

which gives the desired statement. □

Having established the above theorem, we are ready to show the existence of an LOCC ball around the completely depolarizing channel. This is the main result of this subsection.

Theorem 2.17. *Any instrument $\mathfrak{J} = (\mathcal{E}_1, \ldots, \mathcal{E}_n) \in CP[n]$ such that*

$$D_\diamond(\mathfrak{D}, \mathfrak{J}) \leq R_{\mathrm{LOCC}} := \frac{R_{\mathrm{SEP}}}{nd_1^3 d_2} = (n^2 d_1^4 d_2^2 2^{\frac{N}{2}-1} + nd_1^3 d_2)^{-1} \tag{2.35}$$

can be implemented by an N-round LOCC protocol.

Proof. We begin by decomposing each of the CP maps $\mathcal{E}_i$ as

$$\mathcal{E}_i = (1 - nd_2\delta)\frac{1}{n}\mathcal{D} + nd_2\delta\mathcal{F}_i, \tag{2.36}$$

where $\mathcal{F}_i = \frac{1}{n}\mathcal{D} - \frac{1}{nd_2\delta}\left(\frac{1}{n}\mathcal{D} - \mathcal{E}_i\right) \in \mathrm{T}(\mathcal{H}_1, \mathcal{H}_2)$ and δ is to be chosen later. Note that the map

$$\sum_{i\in[n]} \mathcal{F}_i = \mathcal{D} - \frac{1}{nd_2\delta}\Big(\mathcal{D} - \sum_{i\in[n]} \mathcal{E}_i\Big) \tag{2.37}$$

is trace-preserving for any $\delta > 0$. Our goal now is to ensure that $\mathfrak{L} := (\mathcal{F}_1, \ldots, \mathcal{F}_n)$ is a valid instrument by choosing δ so that all the maps $\mathcal{F}_i$ are completely positive. This will allow us to apply Theorem 2.16, since according to Equation (2.36) instrument $\mathfrak{J}$ is a convex combination of $\mathfrak{D}$ and $\mathfrak{L}$.

Considering the Choi matrices

$$J_{\mathcal{F}_i} := \frac{1}{n}J_{\mathcal{D}} + \frac{1}{nd_2\delta}\Big(J_{\mathcal{E}_i} - \frac{1}{n}J_{\mathcal{D}}\Big) \tag{2.38}$$

$$= \frac{1}{nd_2}\bigg(I_{d_1d_2} + \frac{1}{\delta}\Big(J_{\mathcal{E}_i} - \frac{1}{n}J_{\mathcal{D}}\Big)\bigg), \tag{2.39}$$

we see that taking $\delta := \max_{i\in[n]}\|J_{\mathcal{E}_i} - \frac{1}{n}J_{\mathcal{D}}\|_\infty$ ensures that $\mathfrak{L}$ is a valid instrument.

To complete the proof, we only need to show that $nd_2\delta \le R_{\mathrm{SEP}}/d_1^2$ as then according to Theorem 2.16, we can conclude that

$$\mathfrak{J} = (1 - nd_2\delta)\mathfrak{D} + nd_2\delta\mathfrak{L} \in \mathrm{LOCC}_N[n] \tag{2.40}$$

as desired. To achieve this, we will make use of the fact that $D_\diamond(\mathfrak{D}, \mathfrak{J}) \le R_{\mathrm{SEP}}/(nd_1^3d_2)$.

By definition of the diamond norm and δ we have

$$
\begin{aligned}
D_\diamond(\mathfrak{D},\mathfrak{J}) &= \sum_{i\in[n]} \Big\| \mathcal{E}_i - \frac{1}{n}\mathcal{D} \Big\|_\diamond && (2.41)\\
&= \sum_{i\in[n]} \max\Big\{ \Big\| \Big(\Big(\mathcal{E}_i - \frac{1}{n}\mathcal{D} \Big) \otimes \mathcal{I}_{\mathcal{H}_1'} \Big) \rho \Big\|_1 : \rho \in \mathrm{L}(\mathcal{H}_1 \otimes \mathcal{H}_1'), \|\rho\|_1 \le 1 \Big\} && (2.42)\\
&\ge \sum_{i\in[n]} \frac{1}{d_1} \Big\| J_{\mathcal{E}_i} - \frac{1}{n} J_{\mathcal{D}} \Big\|_1 && (2.43)\\
&\ge \sum_{i\in[n]} \frac{1}{d_1} \Big\| J_{\mathcal{E}_i} - \frac{1}{n} J_{\mathcal{D}} \Big\|_\infty && (2.44)\\
&\ge \frac{\delta}{d_1}. && (2.45)
\end{aligned}
$$

Hence, $nd_2\delta \le nd_1d_2D_\diamond(\mathfrak{D},\mathfrak{J}) \le R_{\mathrm{SEP}}/d_1^2$ as desired and the proof is complete. □

2.6.2 Compactness of LOCC_r

Recall from Section 2.4 that the set of LOCC instruments is not closed and hence it clearly cannot be compact. In this section we show that when restricted to finite round protocols with a finite number of outcomes, compactness indeed holds. Let us point out that such a result might not be entirely obvious. It is conceivable that there exists an instrument for which better approximations using $\mathrm{LOCC}_r[n,\mathcal{H}_1,\mathcal{H}_2]$ instruments can only be achieved by allowing unbounded number of intermediate outcomes in the protocols for implementing these instruments. We rule out this possibility, by showing that it is always possible to find a protocol for implementing any instrument from $\mathrm{LOCC}_r[n,\mathcal{H}_1,\mathcal{H}_2]$ with a bounded number of intermediate outcomes.

Suppose that an instrument $\mathfrak{J} = (\mathcal{E}_1,\ldots,\mathcal{E}_n) \in \mathrm{CP}[n,\mathcal{H}_1,\mathcal{H}_2]$ can be implemented by first measuring instrument $\mathfrak{M} = (\mathcal{M}_1,\ldots,\mathcal{M}_t) \in \mathrm{CP}[t,\mathcal{H}_1,\mathcal{H}_3]$ and then, conditioned on outcome $j \in [t]$, measuring instrument $\mathfrak{L}_j = (\mathcal{F}_1,\ldots,\mathcal{F}_n) \in \mathrm{CP}[n,\mathcal{H}_3,\mathcal{H}_2]$. In other words, $\mathcal{E}_i = \sum_{j\in[t]} \mathcal{F}_i \circ \mathcal{M}_j$ for all $i \in [n]$. For brevity, let us use $\big(\bigoplus_{j\in[t]} \mathfrak{L}_j\big) \circ \mathfrak{M}$ to denote the above described implementation. Our goal is to show the measurement in the first step can always be chosen so that t is not too large. Moreover, this can be achieved by essentially discarding some of the CP maps of $\mathfrak{M}$.

Lemma 2.18. *Suppose that an instrument $\mathfrak{J} \in CP[n, \mathcal{H}_1, \mathcal{H}_2]$ admits an implementation $\left(\bigoplus_{j\in[t]} \mathfrak{L}_j\right) \circ \mathfrak{M}$, where $\mathfrak{M} = (\mathcal{M}_1, \ldots, \mathcal{M}_t) \in CP[t, \mathcal{H}_3, \mathcal{H}_2]$. Then $\mathfrak{J}$ also admits an implementation*

$$\left(\bigoplus_{j\in T} \mathfrak{L}_j\right) \circ \mathfrak{M}', \tag{2.46}$$

where $T \subseteq [t]$, $|T| \leq (nd_1d_2d_3^2)^2$, $\mathfrak{M}' = (q_j\mathcal{M}_j : j \in T)$ for some $q_j > 0$ and $d_s = \dim(\mathcal{H}_s)$ for $s = 1, 2, 3$.

Proof. Recall from Section 2.2.1 that the quantum-classical (QC) map $\mathcal{E}[\mathfrak{J}]$ of an instrument $\mathfrak{J} = (\mathcal{E}_1, \ldots, \mathcal{E}_n)$ is given by

$$\mathcal{E}[\mathfrak{J}](\cdot) = \sum_{j\in[n]} \mathcal{E}_j(\cdot) \otimes |j\rangle\langle j|. \tag{2.47}$$

From the contraction diagram in Figure 2.3, we see that the Choi matrix of the QC map $\mathcal{E}[\mathfrak{J}]$ is given by

$$J_{\mathcal{E}[\mathfrak{J}]} = \left(I_{\mathcal{H}_2\otimes\mathcal{H}_c} \otimes \langle\Phi_{\mathcal{H}_3':\mathcal{H}_3}| \otimes I_{\mathcal{H}_1'}\right)\left(\sum_{j\in[t]} J_{\mathcal{E}[\mathfrak{L}_j]} \otimes J_{\mathcal{M}_j}\right)\left(I_{\mathcal{H}_2\otimes\mathcal{H}_c} \otimes |\Phi_{\mathcal{H}_3':\mathcal{H}_3}\rangle \otimes I_{\mathcal{H}_1'}\right), \tag{2.48}$$

where $\mathcal{H}_c = \mathbb{C}^n$ is the classical register of $\mathcal{E}[\mathfrak{L}_j]$ and $|\Phi_{\mathcal{H}:\mathcal{H}'}\rangle := \sum_{s\in[\dim(H)]} |s, s\rangle$. Let

$$J := \sum_{j\in[t]} J_{\mathcal{E}[\mathfrak{L}_j]} \otimes J_{\mathcal{M}_j} = \sum_{j\in[t]} \lambda_j\left(\lambda_j^{-1} J_{\mathcal{E}[\mathfrak{L}_j]} \otimes J_{\mathcal{M}_j}\right) \tag{2.49}$$

where $\lambda_j = \mathrm{Tr}\left(J_{\mathcal{E}[\mathfrak{L}_j]} \otimes J_{\mathcal{M}_j}\right) / \mathrm{Tr}(J)$ and $\sum_{j\in[t]} \lambda_j = 1$. So the last expression shows that J is a convex combination of positive semidefinite matrices with trace $\mathrm{Tr}(J)$.

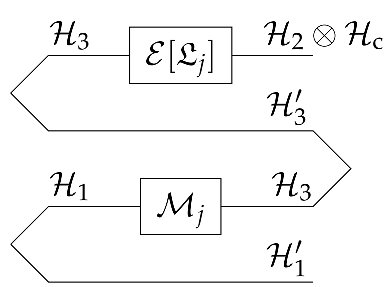


Figure 2.3: Contraction diagram for one of the summands in Equation (2.48).

Since fixed-trace Hermitian matrices over $\mathcal{H}$ can be viewed as a real vector space of dimension $\dim(H)^2 - 1$, we can apply Carathéodory's theorem to Equation (2.49). As $J \in \operatorname{Pos}\left(\mathcal{H}_2 \otimes \mathcal{H}_c \otimes \mathcal{H}_3' \otimes \mathcal{H}_3 \otimes \mathcal{H}_1'\right)$, this yields

$$J = \sum_{j \in T} q_j J_{\mathcal{E}[\mathfrak{L}_j]} \otimes J_{\mathcal{M}_j} = \sum_{j \in T} J_{\mathcal{E}[\mathfrak{L}_j]} \otimes J_{q_j \mathcal{M}_j} \tag{2.50}$$

for some $T \subseteq [t]$ with $|T| \leq (nd_1 d_2 d_3^2)^2$ and some $q_j > 0$. Since $q_j \mathcal{M}_j$ are all completely positive, to show that $\mathfrak{M}' := (q_j \mathcal{M}_j : j \in T)$ is a valid instrument it suffices to check that $\mathcal{M} := \sum_{j \in T} q_j \mathcal{M}_j$ is trace-preserving. To this end we proceed to verify that $\operatorname{Tr}_{\mathcal{H}_3}(J_{\mathcal{M}}) = I_{d_1}$. Using the expressions of J from Equation (2.49) and Equation (2.50) and the fact that $\sum_{j \in [t]} \mathcal{M}_j$ is trace-preserving, we get

$$\begin{aligned}
\operatorname{Tr}_{\mathcal{H}_3}(J_{\mathcal{M}}) =& \operatorname{Tr}_{\mathcal{H}_3}\Big(\sum_{j \in T} J_{q_j \mathcal{M}_j}\Big) && (2.51)\\
=& \frac{1}{d_3} \operatorname{Tr}_{\mathcal{H}_2 \otimes \mathcal{H}_c \otimes \mathcal{H}_3' \otimes \mathcal{H}_3}\Big(\sum_{j \in T} J_{\mathcal{E}[\mathfrak{L}_j]} \otimes J_{q_j \mathcal{M}_j}\Big) && (2.52)\\
=& \frac{1}{d_3} \operatorname{Tr}_{\mathcal{H}_2 \otimes \mathcal{H}_c \otimes \mathcal{H}_3' \otimes \mathcal{H}_3}\Big(\sum_{j \in [t]} J_{\mathcal{E}[\mathfrak{L}_j]} \otimes J_{\mathcal{M}_j}\Big) && (2.53)\\
=& I_{d_1}. && (2.54)
\end{aligned}$$

Hence, $\mathfrak{M}'$ is a valid instrument. Plugging the last expression of J from Equation (2.50) in Equation (2.48) shows that $\left(\bigoplus_{j \in T} \mathfrak{L}_j\right) \circ \mathfrak{M}'$, where $\mathfrak{M}' = (q_j \mathcal{M}_j : j \in T)$. $\square$

The above lemma describes a situation similar to the one encountered at any point in the execution of an LOCC protocol. However, in the latter case we will be able to assume that the instrument $\mathfrak{M}$ acts nontrivially only on the current acting party's space. In such a case the bound on the number of CP maps of $\mathfrak{M}'$ can be improved as follows:

Corollary 2.19. *Suppose that the CP maps of $\mathfrak{M}$ from Lemma 2.18 all take the form*

$$\mathcal{M}_j = \Big(\bigotimes_{L \neq K} \mathcal{I}_{\mathcal{H}_1^{(L)}}\Big) \otimes \mathcal{M}_j^{(K)} \tag{2.55}$$

for some fixed party $K \in [N]$. Then $|T|$, the number of CP maps of $\mathfrak{M}'$, can be upper bounded by $\left(nd_2 d_3 d_1^{(K)} d_3^{(K)}\right)^2$, where $d_i^{(K)} := \dim\left(\mathcal{H}_i^{(K)}\right)$ and $\mathcal{H}_i = \mathcal{H}_i^{(1)} \otimes \ldots \otimes \mathcal{H}_i^{(N)}$ for $i = 1, 2, 3$.

Proof. In this case we can write J from Equation (2.49) in the proof of Lemma 2.18 as

$$J = \Big(\sum_{j\in[t]} J_{\mathcal{E}[\mathfrak{L}_j]} \otimes J_{\mathcal{M}_j^{(K)}} \Big) \otimes \Big(\bigotimes_{L\neq K} \Phi_{\mathcal{H}_1^{(L)}:\mathcal{H}'^{(L)}_1} \Big) \tag{2.56}$$

Applying Carathéodory's theorem to

$$\sum_{j\in[t]} J_{\mathcal{E}[\mathfrak{L}_j]} \otimes J_{\mathcal{M}_j^{(K)}} \in \operatorname{Pos}\left(\mathcal{H}_2 \otimes \mathcal{H}_{\mathrm{c}} \otimes \mathcal{H}'_3 \otimes \mathcal{H}_3^{(K)} \otimes \mathcal{H}'^{(K)}_1 \right) \tag{2.57}$$

allows to reduce the number of summands to at most $\big(nd_2d_3d_1^{(K)}d_3^{(K)}\big)^2$. Next, we proceed similar to the proof of Lemma 2.18 to obtain an implementation $\big(\bigoplus_{j\in T}\mathfrak{L}_j\big)\circ\mathfrak{M}'$ of instrument $\mathfrak{J}$, where $\mathfrak{M}'$ has at most $\big(nd_2d_3d_1^{(K)}d_3^{(K)}\big)^2$ CP maps. □

Having established the last corollary, we are ready to show that any instrument from $\mathrm{LOCC}_r[n,\mathcal{H}_1,\mathcal{H}_2]$ admits an LOCC protocol with bounded number of intermediate measurement outcomes. Recall from Section 2.2.1 that an instrument $\mathfrak{J}\in\mathrm{CP}[n,\mathcal{H}_1,\mathcal{H}_2]$ is said to be fine-grained if each of its CP maps takes the form $M(\cdot)M^\dagger$ for some matrix $M\in\mathrm{L}(\mathcal{H}_1,\mathcal{H}_2)$.

Theorem 2.20. *Consider $\mathfrak{J}\in\mathrm{LOCC}_r[n,\mathcal{H}_1,\mathcal{H}_2]$ and let $D_1 := \max_{K\in[N]}\dim\big(\mathcal{H}_1^{(K)}\big)$. Then $\mathfrak{J}$ can be implemented by an r-round LOCC protocol in which all the instruments are fine-grained. Moreover, the instruments applied in any but the final round have at most $(nd_2d_1D_1^2)^2$ outcomes while those applied in the final round have at most nd_1d_2 outcomes.*

Proof. A general r-round LOCC instrument can be represented as a tree partitioned into r levels. Within each level are nodes that correspond to the different one-way local LOCC instruments performed in that round. The nodes in round $l+1$ are specified by their respective measurement histories $(i_1i_2\ldots i_l)$. At each node many parties may apply a nontrivial TCP map (*i.e.*, instrument with one outcome), but only the acting party can perform an instrument with more than one outcome, and this party may vary across different nodes at each level. Some paths of execution can terminate before round r; in this case "final round" refers to some round $l<r$.

We now explain how to convert a general protocol into one for which

(i) the instruments applied before the final round act nontrivially only on the acting party's space (*i.e.*, other parties apply the identity TCP map $\mathcal{I}$);

(ii) the instruments applied before the final round are fine-grained and have the same input and output dimension;

(iii) the instruments applied in the final round have at most n outcomes.

To obtain (i) each of the non-acting parties postpones the application of any nontrivial TCP map to the closest round in future in which that party is the acting party or to the final round (in the case when the party in question is never the acting party in the later rounds).

Consider a protocol satisfying (i). To obtain (ii) we first replace the nontrivial local instruments applied before the final round with their completely fine-grained versions. This modification can only increase the number of edges leaving a node, but does not change the total number of rounds. The original instrument can be recovered by suitable coarse graining at the end. Then, we apply the polar decomposition $M = UA$ where A is positive semidefinite and U is an isometry. Thus whenever $M(\cdot)M^\dagger$ is performed within the original protocol, it can be replaced with $A(\cdot)A^\dagger$ combined with a pre-application of U in the next level. In other words, if the same party applies Kraus operator N next in the original protocol, then NU is applied instead. Polar decomposing NU, we obtain $U'A'$ where A' again maps the initial system to itself, and U' is some other isometry to be moved to yet the next level. Doing this inductively for all levels yields condition (i) in which the state lives in the same input system throughout, except for the final round.

To obtain (iii), we recall that the instrument $\mathfrak{J}$ that we are trying to implement has only n outcomes. Therefore, if some instrument $\mathfrak{L}$ from the final round has more than n outcomes some of them correspond to the same outcome of $\mathfrak{J}$. We can coarse grain all CP maps of $\mathfrak{L}$ corresponding to the same outcome of $\mathfrak{J}$ to obtain an instrument $\mathfrak{L}'$ with at most n outcomes.

Let $\mathcal{T}$ be the tree corresponding to an r-round LOCC protocol that implements $\mathfrak{J}$ and satisfies (i)–(iii). Before the final round at any node m of $\mathcal{T}$ the parties apply some fine-grained instrument $\mathfrak{M} = (\mathcal{M}_1, ..., \mathcal{M}_t) \in \mathrm{LOCC}_1[t, \mathcal{H}_1, \mathcal{H}_1]$ whose outcomes lead to the children nodes $c_1, \ldots, c_t$ of m. Next, conditioned on outcome $j \in [t]$, the parties apply instrument $\mathfrak{L}_j \in \mathrm{LOCC}_{\mathbb{N}}[n, \mathcal{H}_1, \mathcal{H}_2]$ which corresponds to the subtree of $\mathcal{T}$ rooted at node c_j. Thus, we are in a situation described by Lemma 2.18. Moreover, property (iii) guarantees that $\mathfrak{M}$ is nontrivial only on the acting party's space and so we can apply Corollary 2.19. This allows us to prune tree $\mathcal{T}$ at node m by replacing instrument $\mathfrak{M}$ with $\mathfrak{M}' = (q_j\mathcal{M}_j : j \in T)$, where $T \subseteq [t]$, $|T| \leq (nd_1d_2D_1^2)^2$ and $q_j > 0$. We apply Corollary 2.19 to any instrument applied before the final round. This yields an r-round protocol $\mathcal{P}$ that implements $\mathfrak{J}$ and the instruments applied in any but the final round have no more than $(nd_1d_2D_1^2)^2$ outcomes.

Finally, we fine-grain the n-outcome instruments at the final round of $\mathcal{P}$. Consider any CP map $\mathcal{F}$ of some instrument $\mathfrak{L}$ from the final round. Note that $\mathcal{F}$ is an N-fold tensor product of local CP maps on each of the parties. From the Choi representation one sees that any CP map in $\mathrm{T}(\mathcal{H}_1^{(K)}, \mathcal{H}_2^{(K)})$ can be represented using at most $d_1^{(K)} d_2^{(K)}$ Kraus operators, where $d_s^{(K)} = \dim\left(\mathcal{H}_s^{(K)}\right)$ for $s = 1, 2$. Therefore, we can fine-grain each of the n completely positive maps of $\mathfrak{L}$ into at most $d_1 d_2$ maps of the form specified by the theorem, *i.e.*, $\rho \mapsto M\rho M^\dagger$. This yields a fine-graining $\mathfrak{L}'$ of $\mathfrak{L}$ with at most nd_1d_2 outcomes as desired.

□

We are now ready to establish the main result of this subsection.

Theorem 2.21. *The set of instruments* $\mathrm{LOCC}_r[n, \mathcal{H}_1, \mathcal{H}_2]$ *is compact for all positive integers* r *and* n*.*

Proof. Consider an instrument $\mathfrak{J} \in \mathrm{LOCC}_r[n, \mathcal{H}_1, \mathcal{H}_2]$ and let $\mathcal{P}$ be an r-round LOCC protocol implementing $\mathfrak{J}$ from Theorem 2.20. We can specify $\mathcal{P}$ using the Choi matrices of the N-fold tensor product CP maps of the instruments in each round. We can ensure that all paths of execution in $\mathcal{P}$ have length exactly r by inserting intermediate trivial instruments $(\mathcal{I}) \in \mathrm{CP}[1, \mathcal{H}_1, \mathcal{H}_1]$. We can also add zero maps to instruments to ensure that the instruments up to level $r-1$ have exactly $B := (d_1 d_2 D_1^2)^2$ outcomes, while those in the level r have exactly $C := nd_1d_2$ outcomes. This way, we can specify each instrument from $\mathrm{LOCC}_r[n, \mathcal{H}_1, \mathcal{H}_2]$ using exactly $D := (B^r - B)/(B-1) + B^{r-1}C$ matrices.

With the above in mind, let S be the set consisting of D-tuples

$$\left(J_{i_1}^{()}, J_{i_2}^{(i_1)}, \ldots, J_{i_r}^{(i_1 \ldots i_{r-1})} : i_1, \ldots, i_{r-1} \in [B] \text{ and } i_r \in [C]\right), \tag{2.58}$$

satisfying the following conditions:

1. the matrices $J_{i_r}^{(i_1 \ldots i_{r-1})} \in \mathrm{Pos}(\mathcal{H}_2 \otimes \mathcal{H}_1')$ and $\sum_{i_r=1}^{C} \mathrm{Tr}_{\mathcal{H}_2}\left(J_{i_r}^{(i_1 \ldots i_{r-1})}\right) = I_{\mathcal{H}_1'}$ for any measurement history $(i_1 \ldots i_{r-1})$;
2. for any $l \in [r-1]$ and for any measurement history $(i_1 \ldots i_{l-1})$, the matrices $J_{i_l}^{(i_1 \ldots i_{l-1})} \in \mathrm{Pos}(\mathcal{H}_1 \otimes \mathcal{H}_1')$ and $\sum_{i_l=1}^{B} \mathrm{Tr}_{\mathcal{H}_1}\left(J_{i_l}^{(i_1 \ldots i_{l-1})}\right) = I_{\mathcal{H}_1'}$.
3. for all $l \in [r]$ the matrices $J_{i_l}^{(i_1 \ldots i_{l-1})}$ are N-fold tensor products;

The set S consists of D-tuples whose each coordinate belongs to either set S_1 or S_2, where

$$S_t := \{J \in \mathrm{Pos}(\mathcal{H}_t \otimes \mathcal{H}_1') \text{ is an } N\text{-fold tensor product} : I_{\mathcal{H}_1'} - \mathrm{Tr}_{\mathcal{H}_t}(J) \in \mathrm{Pos}(\mathcal{H}_1')\} \tag{2.59}$$

for $t = 1, 2$. Both S_1 and S_2 are closed and bounded. Therefore, S is also bounded. Since the only other constraints for D-tuples in S are of equality type, the set S is closed. Finally, as we are working in finite dimensional space, we can conclude that S is compact.

Let $f : S \to \mathrm{LOCC}_r[n, \mathcal{H}_1, \mathcal{H}_2]$ be a function that assigns every tuple in S its corresponding instrument in $\mathrm{LOCC}_r[n, \mathcal{H}_1, \mathcal{H}_2]$. Function f is both continuous and surjective. Since continuous functions map compact sets to compact ones, we get that $\mathrm{LOCC}_r[n, \mathcal{H}_1, \mathcal{H}_2]$ is compact. □

2.7 Discussion and open problems

In this chapter we have studied the structure of LOCC operations. In light of recent findings concerning the nature of asymptotic LOCC processes, we have adopted the formalism of quantum instruments to precisely characterize the topological closure of LOCC.

There are a few interesting questions related to this work that deserve additional investigation. First, all known examples that separate LOCC from $\overline{\mathrm{LOCC}}$ or $\overline{\mathrm{LOCC}}$ from SEP make use of the classical information obtained from a quantum measurement. For quantum channels with no classical output register, it is unknown whether the same separation results hold (although, $\mathrm{LOCC}_{\mathbb{N}}$ can be separated from SEP by such channels [CD09]).

On the other hand, one can ask how the operational classes compare if one is only interested in the classical information extracted from a quantum measurement. While the state discrimination results take such an approach to separate $\overline{\mathrm{LOCC}}$ from SEP, the random distillation examples demonstrating $\mathrm{LOCC}_{\mathbb{N}} \neq \mathrm{LOCC} \neq \overline{\mathrm{LOCC}}$ depend crucially on the quantum outputs of the measurement. Thus, it may be possible that finite and asymptotic LOCC are equally powerful for implementing quantum maps with only classical output register.

In Section 2.6.1 we discussed an LOCC protocol for probabilistically implementing any separable instrument. It is not known whether this protocol achieves optimal worst-case success probability. Understanding which separable instruments are hardest to implement (in terms of success probability) with LOCC would further our understanding of the relationship between the two classes.

Finally, using the distance measure between instruments described in Section 2.2.1, one can meaningfully inquire about the size in separation between operational classes. When given some instrument in SEP, what is the closest LOCC instrument? Furthermore, is this distance related to the nonlocal resources needed to implement the separable instrument? We hope this work stimulates further research into such questions concerning the structure of LOCC.

Chapter 3

Bipartite state discrimination with LOCC

3.1 Introduction

In the previous chapter we were concerned with general LOCC instruments. The rest of this thesis focuses on the problem of discriminating separable states from some known set $S = \{|\psi_i\rangle\}_i$ using LOCC. To accomplish this task we only need to produce a classical answer (*i.e.*, the index "i"), hence we are interested in implementing a measurement rather than a general quantum instrument. In this chapter we provide the common background for the results concerning state discrimination and implementation of von Neumann measurements developed in the later chapters. We also discuss the problem of discriminating states from an orthonormal basis with finite LOCC. Finally, we ask whether asymptotic LOCC offers any advantage over finite LOCC for implementing projective measurements.

This chapter is organized as follows. We start by discussing separable and LOCC measurements and describing the tree structure of an LOCC protocol in Section 3.2. Then in Section 3.3 we formally define the state discrimination problem and clarify its relation to that of implementing von Neumann measurements. We continue with a review of previous results on state discrimination with LOCC in Section 3.4. Next, in Section 3.5 we introduce the concept of non-disturbing measurements and explain their relevance for state discrimination with LOCC. In Section 3.6 we show that there exists a polynomial time algorithm for deciding whether states from a basis can be discriminated with finite LOCC. In Section 3.7 we give a more detailed account of the results on state discrimination with asymptotic LOCC presented in [KKB11]. In the later chapters we use these results, and their interpolation technique in particular, in an essential way. Finally in Section 3.8 we prove that asymptotic and finite LOCC are equally powerful for implementing a projective measurement with tensor product operators. This result generalizes a theorem of [KKB11] reviewed in Section 3.7.

3.2 Separable and LOCC measurements

3.2.1 Separable measurements

In this section we specialize the definition of separable instruments from Chapter 2 for 2-party POVMs.

A POVM $\{E_i\}_{i\in[n]} \subseteq \mathrm{Pos}(\mathbb{C}^{d_A}\otimes\mathbb{C}^{d_B})$ corresponds to a 2-party instrument $\mathfrak{J} = (\mathcal{E}_1,\ldots,\mathcal{E}_n)$, where the CP maps $\mathcal{E}_i$ are defined via $\mathcal{E}_i(\rho) = \mathrm{Tr}(E_i\rho)$ for all $i\in[n]$. The Choi matrix of $\mathcal{E}_i$ is given by

$$J_{\mathcal{E}_i} = \left(\mathcal{E}_i\otimes\mathcal{I}\right)\Phi_{AB:A'B'} = \sum_{\substack{a,c\in[d_A]\\ b,d\in[d_B]}} \mathrm{Tr}\left(E_i|a,b\rangle\langle c,d|\right)|a,b\rangle\langle c,d| = E_i, \tag{3.1}$$

where $\Phi_{AB:A'B'}$ is an unnormalized maximally entangled state (see Equation (2.9)). Recall from Section 2.3 that an instrument $\mathfrak{J}$ is separable if and only if the Choi matrices of all its CP maps $\mathcal{E}_i$ are separable. Since $J_{\mathcal{E}_i} = E_i$, the POVM $\{E_i\}_{i\in[n]}$ is separable if and only if all the E_i are separable, *i.e.*,

$$E_i = \sum_j E_j^A \otimes E_j^B \tag{3.2}$$

for some $E_j^A \in \mathrm{Pos}(\mathbb{C}^{d_A})$ and $E_j^B \in \mathrm{Pos}(\mathbb{C}^{d_B})$.

3.2.2 LOCC measurements

In this section we specialize the definition of finite LOCC instruments from Chapter 2 for 2-party POVMs. We also describe a general LOCC protocol for implementing such a POVM.

Let $\mathcal{M}$ be a 2-party measurement with POVM $\{E_i\}_{i\in[n]} \subseteq \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_A} d_b)$ and let $\mathfrak{J} = (\mathcal{E}_i : i \in [n])$ be the instrument corresponding to this POVM. We say that $\mathcal{M}$ is a 2-party *LOCC measurement* if the instrument $\mathfrak{J}$ belongs to 2-party $\mathrm{LOCC}_{\mathbb{N}}$. In other words, the measurement statistics of $\mathcal{M}$ can be reproduced exactly by some finite 2-party LOCC protocol $\mathcal{P}$. We will say that such a $\mathcal{P}$ *implements* the measurement $\mathcal{M}$. It would be more accurate to use the term "finite LOCC measurement" to refer to $\mathcal{M}$, however for shortness we have chosen to omit the description "finite".

As described in Section 2.2.2, if $\mathfrak{J} \in \mathrm{LOCC}_{\mathbb{N}}$ then it can be implemented by Alice and Bob taking finitely many turns in applying one-way local instruments and communicating the outcome to the other party. The instrument applied at round t depends on the *measurement record* $m = (m_1, \ldots, m_{t-1})$ accumulated during the previous rounds.

We now describe a general LOCC protocol for implementing an n-outcome POVM $\{E_i\}_{i\in[n]}$. Since coarse graining can always be postponed to the end of the protocol, we can assume that each of the one-way local instruments is completely fine-grained (see Section 2.2.1 for terminology). Note that the non-acting party can postpone the application of any nontrivial TCP map to the next round. Therefore, we can assume that the applied instruments act non-trivially only on the acting party's space (this might require extending the protocol by one additional round). Finally, in the case of only two parties, we can assume that Alice is the acting party in all odd numbered rounds and Bob is the acting party in all even numbered rounds or vice versa.

Let $m = (m_1, \ldots, m_t)$ be the measurement record after the execution of the first t rounds of the protocol and let Λ be the empty string, corresponding to no messages having been sent yet. Also, let $A_{m_i}(m_1, \ldots, m_{i-1})$ (and $B_{m_i}(m_1, \ldots, m_{i-1})$) be the Kraus operator of the m_ith CP map of the instrument Alice (Bob) measures in round i. Then the Kraus operator that Alice and Bob have implemented at the end of round t is a product operator $A_m \otimes B_m$, where[1]

$$A_m := A_{m_{t-1}}(m_1, \ldots, m_{t-2}) \ldots A_{m_3}(m_1, m_2) A_{m_1}(\Lambda), \tag{3.3}$$

$$B_m := B_{m_t}(m_1, \ldots, m_{t-1}) \ldots B_{m_4}(m_1, m_2, m_3) B_{m_2}(m_1). \tag{3.4}$$

Certain measurement records will cause Alice and Bob to terminate the protocol. Prior to the final coarse-graining, the quantum operation implemented by the LOCC protocol acts on any state ρ as

$$\bigoplus_m (A_m \otimes B_m)\rho(A_m \otimes B_m)^\dagger \tag{3.5}$$

[1]Here we assume for simplicity that t is even and Alice starts the protocol; in the other cases the operators A_m and B_m can be defined similarly.

where m ranges over all the terminating measurement records. When the output dimensions of A_m and B_m are equal for all m, the above expression can be written as

$$\sum_m |m\rangle\langle m| \otimes (A_m \otimes B_m)\rho(A_m \otimes B_m)^\dagger. \tag{3.6}$$

Here the first register stores the classical measurement record and is shared between the two parties; the last two registers belong to Alice and Bob, respectively. At the end of the protocol Alice and Bob must output one of the n classical outcomes of $\mathcal{M}$. Let $L(k)$ be the set of all terminating measurement records corresponding to outcome $k \in [n]$. Then coarse graining according to the partition $\big(L(1), \dots, L(n)\big)$ corresponds to measuring the classical register in Equation (3.5) according to this partition. The kth POVM element of the resulting measurement is

$$\sum_{m \in L(k)} A_m^\dagger A_m \otimes B_m^\dagger B_m, \tag{3.7}$$

which must equal E_k if the LOCC protocol indeed implements measurement $\mathcal{M}$.

3.2.3 Finite and asymptotic LOCC

We consider two scenarios: when a measurement can be performed in a finite number of rounds or asymptotically.

Definition 3.1. We say that a measurement $\mathcal{M}$ can be *implemented by finite LOCC*, *i.e.*, $\mathrm{LOCC}_{\mathbb{N}}$, if there exists a finite-round LOCC protocol that, for any input state, produces the same distribution of measurement outcomes as $\mathcal{M}$.

Definition 3.2. We say that a measurement $\mathcal{M}$ can be *implemented by asymptotic LOCC*, *i.e.*, $\overline{\mathrm{LOCC}}$, if there exists a sequence $\mathcal{P}_1, \mathcal{P}_2, \dots$ of finite-round LOCC protocols whose output distributions converge to that of $\mathcal{M}$.

The exact implementation scenario is not practical since any real-world device is susceptible to errors due to imperfections in implementation. However, proving that a certain task cannot be performed asymptotically is considerably harder than showing that it cannot be done (exactly) by any finite LOCC protocol.

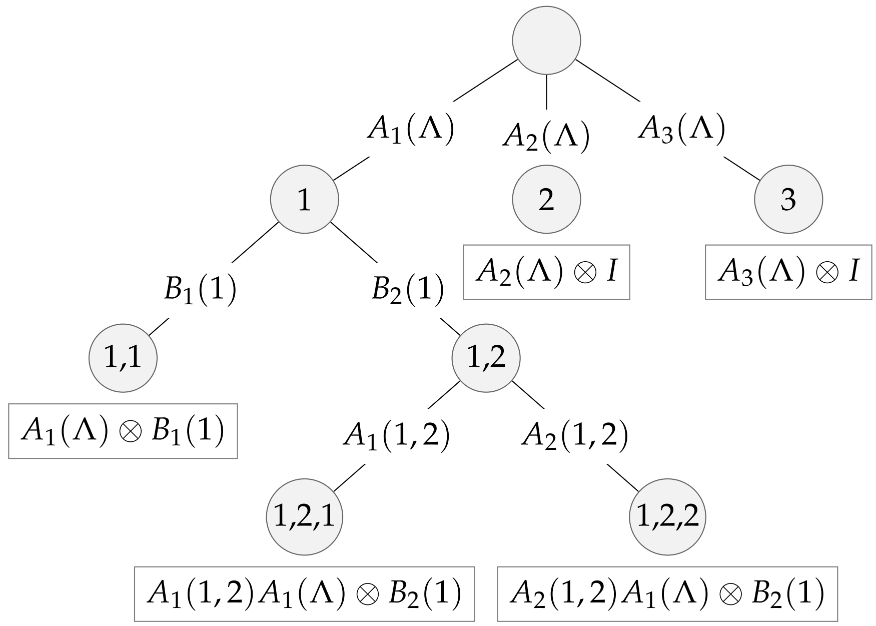


Figure 3.1: An example showing the tree structure of a specific LOCC measurement. In round one Alice performs a three-outcome measurement $\mathcal{A}(\Lambda)$; in round two, upon receiving message "1", Bob performs a two-outcome measurement $\mathcal{B}(1)$ and upon receiving message "2" or "3" he terminates the protocol; in round three, upon receiving message "1", Alice terminates the protocol and, upon receiving message "2", she performs a two-outcome measurement $\mathcal{A}(1,2)$. All nodes are labeled by the accumulated measurement record. The corresponding measurement operator is given below each leaf.

3.2.4 LOCC protocol as a tree

We represent an LOCC measurement protocol as a rooted tree (see Figure 3.1). The protocol begins at the root and proceeds downward along the edges. Each edge represents a certain measurement outcome obtained at its parent node, and *leaves* are the nodes where the protocol terminates. The set of all leaves is partitioned into subsets, each corresponding to an outcome of the LOCC measurement being implemented.

A path from the root to a leaf is called a *branch*. There is a one-to-one correspondence between the branches and the possible courses of execution of the LOCC protocol. Likewise, there is a one-to-one correspondence between the nodes of the tree and the accumulated measurement records.

The *measurement at node u* is the measurement performed by the acting party once the protocol has reached node u. In contrast, the *measurement operator corresponding to node u* is the measurement operator that has been implemented upon reaching node u. For example, consider the node $(1,2)$. The measurement at node $(1,2)$ is given by the measurement operators $\{A_1(1,2), A_2(1,2)\}$, whereas the measurement operator corresponding to the node $(1,2)$ is given by $A_1(\Lambda) \otimes B_2(1)$ (recall that Λ denotes the empty string). As another example, the measurement operators corresponding to the leaves are exactly the measurement operators of the LOCC protocol prior to coarse graining.

3.3 Bipartite state discrimination problem

The task of bipartite state discrimination is central to this thesis. Formally, the problem is as follows:

Let $S = \{\rho_1, \ldots, \rho_n\} \subseteq \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ be a known set of quantum states. Suppose that $k \in [n]$ is selected uniformly at random and Alice and Bob are given the corresponding parts of state $\rho_k \in S$. Their task is to determine the index k with certainty by performing a measurement on this state.

Sometimes we will be redundant and say that Alice and Bob can discriminate the states from S *with certainty*. The case when they must succeed with at least some fixed probability p is also important, but will not be considered in this thesis.

We will mostly be concerned with situations where S is a set of pure states, *i.e.*, each $\rho_i = |\psi_i\rangle\langle\psi_i|$ for some $|\psi_i\rangle \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. A case of special interest is when S is an orthonormal product basis, *i.e.*, each $|\psi_i\rangle = |\alpha_i\rangle|\beta_i\rangle$ for some unit vectors $|\alpha_i\rangle \in \mathbb{C}^{d_A}$ and $|\beta_i\rangle \in \mathbb{C}^{d_B}$. Such states can be perfectly discriminated by a separable measurement $\mathcal{M}$ with POVM elements

$$E_i := |\alpha_i\rangle\langle\alpha_i| \otimes |\beta_i\rangle\langle\beta_i|. \tag{3.8}$$

However, this measurement cannot always be implemented by finite [WH02, GV01] or even asymptotic LOCC [BDF+99]. In such cases we say that S possesses nonlocality (without entanglement).

We now show that in certain situations there is a close connection between the task of discriminating states from a set S with LOCC and implementing a certain projective measurement with LOCC. In these cases the following lemma will allow us to consider the two tasks interchangeably.

Lemma 3.3. *Let $\mathcal{M} = \{P_1, \ldots, P_n\} \subseteq \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ be a projective measurement. Also let $S = \{\rho_1, \ldots, \rho_n\}$, where $\rho_i := \frac{1}{\mathrm{rank}(P_i)} P_i$ for all $i \in [n]$. An LOCC protocol $\mathcal{P}$ implements measurement $\mathcal{M}$ if and only if $\mathcal{P}$ can be used to discriminate the states from S with certainty.*

Proof. Clearly, given an LOCC protocol $\mathcal{P}$ that implements the measurement $\mathcal{M}$, we can use the measurement outcome to discriminate the states from S with certainty.

To prove the other direction, suppose $\mathcal{P}$ can be used to discriminate the states from S with certainty. Consider any nonzero POVM element $E \otimes F$ that corresponds to some leaf of the protocol tree of $\mathcal{P}$. Since $\mathcal{P}$ discriminates the states from S with certainty, $\mathrm{Tr}\big((E \otimes F)\rho\big) = 0$ for all but one of the states $\rho \in S$. Hence, if $\mathrm{Tr}\big((E \otimes F)\rho_i\big) \neq 0$ then the column space of $E \otimes F$ is contained in that of ρ_i (or equivalently, in that of P_i). Let Λ_i be the set of POVM elements $E \otimes F$ of $\mathcal{P}$ for which $\mathrm{Tr}\big((E \otimes F)\rho_i\big) \neq 0$. Since POVM elements must sum to identity, we have

$$\sum_{E \otimes F \in \Lambda_i} E \otimes F = P_i \tag{3.9}$$

and the protocol $\mathcal{P}$ can be used to implement $\mathcal{M}$. □

Corollary 3.4. *Consider measurement $\mathcal{M}$ and set of states S from Lemma 3.3. Then $\mathcal{M} \in \overline{\mathrm{LOCC}}$ if and only if the states from S can be discriminated with asymptotic LOCC,* i.e., $\overline{\mathrm{LOCC}}$.

Proof. Consider a sequence of measurements $\mathcal{M}_1, \mathcal{M}_2 \ldots$ that converges to $\mathcal{M}$. Note that using the outcomes of $\mathcal{M}_i$ to discriminate the states from S will give a vanishing error probability as i tends to ∞. Thus, the states from S can be discriminated with asymptotic LOCC.

To obtain the other direction, suppose that $\mathcal{P}_1, \mathcal{P}_2 \ldots$ is a sequence of finite LOCC protocols that discriminate the states from S with vanishing error. Since the only measurement that discriminates the states from S with certainty is $\mathcal{M}$, the sequence of measurements $\mathcal{M}'_1, \mathcal{M}'_2 \ldots$ implemented by protocols $\mathcal{P}_1, \mathcal{P}_2 \ldots$, respectively, must converge to $\mathcal{M}$. So we conclude that $\mathcal{M} \in \overline{\mathrm{LOCC}}$. □

In general, there can be more than one measurement that discriminates the states from S. Hence, even if states from S can be discriminated with finite LOCC, it can happen that a measurement $\mathcal{M}$ that can be used to discriminate these states does not belong even to the closure of LOCC (*i.e.*, $\overline{\mathrm{LOCC}}$). For example, consider the set

$$S := \big\{|\psi\rangle, |00\rangle\big\}, \tag{3.10}$$

where $|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$. Clearly, the states from S can be discriminated with finite LOCC. Now consider the measurement

$$\mathcal{M} := \{|\psi\rangle\langle\psi|, I - |\psi\rangle\langle\psi|\}. \tag{3.11}$$

Although $\mathcal{M}$ can be used to discriminate the states from S, it does not belong to $\overline{\mathrm{LOCC}}$, since its POVM elements are not separable. Hence, we see that in general there does not exist such a convenient connection between the task of state discrimination and that of measurement implementation as in the special case of Lemma 3.3 and Corollary 3.4.

3.4 Previous results

In this section we review the known results on state discrimination with LOCC. This problem has been extensively studied in the past decade. We do not intend to give a comprehensive overview, but rather focus on the results that are relevant to product state discrimination or deal with asymptotic as opposed to finite LOCC.

The first example of an orthonormal product basis of bipartite quantum states that cannot be perfectly discriminated by even asymptotic LOCC was given in [BDF+99]. By considering the so-called domino states authors of [BDF+99] gave a striking illustration of the difference between the power of LOCC and separable operations. Furthermore, [BDF+99] quantifies the information deficit of any LOCC protocol for discriminating these states. This result has been a starting point for many other studies on state discrimination by LOCC, with the ultimate goal of understanding LOCC operations and how they differ from separable ones. We briefly describe some of the directions that have been explored. Unless otherwise stated, these results refer to the discrimination of *pure* states with *finite* LOCC.

First consider the problem of discriminating two states without any restrictions on their dimension. Surprisingly, any two orthogonal (possibly entangled) pure states can be perfectly discriminated by finite LOCC, even when they are held by more than two parties [WSHV00]. Furthermore, optimal discrimination of any two multipartite pure states can be achieved with finite LOCC both in the sense of minimum error probability [VSPM01] and unambiguous discrimination [CY01, CY02, JCY05]. Recently this has been generalized to implementing an arbitrary POVM by finite LOCC in any 2-dimensional subspace [Cro12].

Many authors have considered the problem of perfect state discrimination by finite LOCC. In particular, the case where one party holds a small-dimensional system is well understood. Reference [WH02] characterizes when a set of orthogonal (possibly entangled) states in $\mathbb{C}^2 \otimes \mathbb{C}^2$ can be perfectly discriminated by finite LOCC. A similar characterization for sets of orthogonal *product* states in $\mathbb{C}^3 \otimes \mathbb{C}^3$ has been given by [FS09]. In addition, [WH02] characterizes when a set of orthogonal states in $\mathbb{C}^2 \otimes \mathbb{C}^n$ can be perfectly discriminated by finite LOCC when Alice performs the first nontrivial measurement. It is also known that a generalization of domino states, the so-called θ-rotated domino states, cannot be perfectly discriminated by finite LOCC (unless $\theta = 0$) [GV01]. Furthermore, the original domino states have inspired a construction of n-partite d-dimensional product bases that cannot be perfectly discriminated with finite LOCC [NC06].

The role of entanglement in perfect state discrimination by LOCC has also been considered. It is not possible to perfectly discriminate more than two Bell states in $\mathbb{C}^2 \otimes \mathbb{C}^2$ by finite LOCC [GKR$^+$01]. In fact, the same is true for any set of more than n maximally entangled states in $\mathbb{C}^n \otimes \mathbb{C}^n$ [Nat05]. Multipartite states from an orthonormal basis can be perfectly discriminated by finite LOCC only if it is a product basis [HSSH03]. Also, no basis of the subspace orthogonal to a state with Schmidt rank 3 or greater can be perfectly discriminated even by asymptotic LOCC [DFXY09]. On the other hand, any three orthogonal maximally entangled states in $\mathbb{C}^3 \otimes \mathbb{C}^3$ can be perfectly discriminated by finite LOCC [Nat05]. In fact, if the number of dimensions is not restricted, one can find arbitrarily large sets of orthogonal maximally entangled states that can be perfectly discriminated by finite LOCC [Fan04]. Contrary to intuition, states with more entanglement can sometimes be discriminated perfectly with finite LOCC while their less entangled counterparts cannot [HSSH03]. Generally, however, a set of orthogonal multipartite states $S \subseteq \mathbb{C}^D$ can be perfectly discriminated with finite LOCC only if $|S| \leq \frac{D}{d(S)}$, where $d(S)$ measures the average entanglement of the states in S [HMM$^+$06].

It is known that local projective measurements are sufficient to discriminate states from an orthonormal product basis with finite LOCC [DR04, CL04]. Moreover, there is a polynomial-time algorithm for deciding if states from a given orthonormal product basis of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ can be perfectly discriminated with finite LOCC [DR04]. The state discrimination problem for *incomplete* orthonormal sets (*i.e.*, orthonormal sets of states that do not span the entire space) seems to be harder to analyze. However, unextendible product bases might be an exception (although commonly referred to as "bases" these are in fact incomplete orthonormal sets). It is known that states from an unextendible product basis cannot be perfectly discriminated by finite LOCC [BDM$^+$99]. In fact, the same holds for any basis of a subspace spanned by an unextendible product basis in $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ [DXY10]. Curiously, there are only two families of unextendible product bases in $\mathbb{C}^3 \otimes \mathbb{C}^3$, one of which is closely related to the domino states [DMS$^+$03].

The problem of state discrimination with asymptotic LOCC has been studied less. It is known that states from an unextendible orthonormal product set cannot be perfectly discriminated with LOCC even asymptotically [DR04]. Reference [KKB11] gives a necessary condition for perfect asymptotic LOCC discrimination, and also shows that for perfectly discriminating states from an orthonormal product basis, asymptotic LOCC gives no advantage over finite LOCC. The latter result implies that the algorithm from [DR04] also covers the asymptotic case. On the other hand, even in some very basic instances of state discrimination it remains unclear whether asymptotic LOCC is superior to finite LOCC (see [DFXY09, KKB11] for specific sets of states).

Another line of study originating from [BDF+99] aims at understanding the difference between the classes of separable and LOCC operations. To this end, [Coh11] constructs an r-round LOCC protocol implementing an arbitrary separable measurement whenever such a protocol exists. A different approach is to exhibit quantitative gaps between the two classes. To the best of our knowledge, only two quantitative gaps other than that of [BDF+99] are known. References [KTYI07, Koa09] demonstrate a gap between the success probabilities achievable by bipartite separable and LOCC operations for unambiguously discriminating $|00\rangle$ from a fixed rank-2 mixed state. The largest known difference between the two classes is a gap of 0.125 between the achievable success probabilities for *tripartite* EPR pair distillation [CCL12a, CCL12a]. Moreover, as the number of parties grows, the gap approaches 0.37 [CCL12a, CCL12b].

At a first glance one might think that the nonlocality without entanglement phenomenon is related to quantum discord. However, the quantum discord value cannot be used to determine whether states from a given ensemble can be discriminated with LOCC [BT10].

Finally, if a set of orthogonal (product or entangled) states cannot be perfectly discriminated by LOCC, one can measure their nonlocality by considering how much entanglement is needed to achieve perfect discrimination [Coh08, BBKW09].

3.5 Non-disturbing measurements

In this section we introduce the concept of non-disturbing operators, which is intrinsic to perfect state discrimination with both finite and asymptotic LOCC.

Definition 3.5. Let $S \subseteq \mathbb{C}^d$ be a set of orthogonal states. We say that $E \in \mathrm{Pos}(\mathbb{C}^d)$ is *non-disturbing for S*, if

$$\langle\psi|E|\phi\rangle = 0 \tag{3.12}$$

for all distinct $|\psi\rangle, |\phi\rangle \in S$. We say that a measurement $\mathcal{M}$ is *non-disturbing for S* if each of the POVM elements of $\mathcal{M}$ is non-disturbing for S.

The above definition suits the case when we want to discriminate pure states. If S consists of mixed states, in place of Condition (3.12) we require that

$$\mathrm{Tr}(E\rho E\sigma) = 0 \tag{3.13}$$

for all distinct $\rho, \sigma \in S$.

To better understand what the condition in Equation (3.12) means for state discrimination, note that two distinct states $|\psi\rangle, |\phi\rangle \in S$ and a POVM element E satisfy Equation (3.12) if and only if we are in one the following two situations.

- At least one of the states $|\psi\rangle, |\phi\rangle$ never leads to the outcome corresponding to E.
- Each of the states $|\psi\rangle$ and $|\phi\rangle$ can lead to the outcome corresponding to E but the obtained post-measurement states remain orthogonal.

Note that any measurement protocol that perfectly discriminates the states from S must consist solely of measurements that are non-disturbing for the current set of states. In particular, the protocol must start with a measurement that is non-disturbing for S. Since in finite LOCC protocols each measurement must be local and, without loss of generality, non-trivial, we obtain the following.

A necessary condition for perfect state discrimination with finite LOCC. *The states from a set S can be perfectly discriminated with finite LOCC only if S admits a non-disturbing product operator $a \otimes b$ where exactly one of the matrices a, b is the identity matrix.*

Showing that the above condition fails is a common approach for establishing that states from a given set cannot be perfectly discriminated with finite LOCC. It is not known whether this condition is also necessary for perfect state discrimination with asymptotic LOCC. However, in Section 3.7.2 we show that the concept of non-disturbing operators can also be used to formulate a necessary condition for state discrimination with asymptotic LOCC.

We conclude this section with a simple observation that the existence of a non-disturbing operator of the form $a \otimes I$ or $I \otimes b$ implies the existence of a non-disturbing measurement for the respective party. We first show that any nonnegative linear combination of non-disturbing operators is also non-disturbing.

Lemma 3.6. *Let $S \subseteq \mathrm{Pos}(\mathbb{C}^d)$ be a set of states and assume that $E_1, \dots, E_k \in \mathrm{Pos}(\mathbb{C}^d)$ are non-disturbing for S. Let $c_1, \dots, c_k$ be such that the operator*

$$E := \sum_{i \in [k]} c_i E_i \tag{3.14}$$

is positive semidefinite. Then E is non-disturbing for S.

Proof. For any $M, \rho, \sigma \in \mathrm{Pos}(\mathbb{C}^d)$

$$\mathrm{Tr}(M\rho M\sigma) = \mathrm{Tr}\left((\sqrt{\rho}M\sqrt{\sigma})^\dagger(\sqrt{\rho}M\sqrt{\sigma})\right) = \|\sqrt{\rho}M\sqrt{\sigma}\|_2^2. \tag{3.15}$$

Since $\|A\|_2 = 0$ implies that A is the zero matrix, $\sqrt{\rho}E_i\sqrt{\sigma}$ must be the zero matrix for all $i \in [k]$. By definition of E, also $\sqrt{\rho}E\sqrt{\sigma} = 0$ must be the zero matrix. This shows that $\mathrm{Tr}(E\rho E\sigma) = 0$ for all $\rho, \sigma \in S$ and hence E is non-disturbing for S. □

Since the identity operator is non-disturbing for any set of states S, we obtain the following corollary.

Corollary 3.7. *If $a \otimes I \in \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ is nontrivial and non-disturbing for S, then so is the measurement*

$$\{ta, I_{d_A} - ta\} \otimes \mathcal{I}_B, \tag{3.16}$$

where $t > 0$ is such that $I_{d_A} - ta \in \mathrm{Pos}(\mathbb{C}^{d_A})$.

3.6 Discriminating states from a basis with finite LOCC

The goal of this section is to show that if S is a basis, then it is easy to decide whether the states from S can be discriminated with finite LOCC. As shown in Lemma 3.3, discriminating the states from an orthonormal basis is equivalent to implementing a measurement in that basis. Since any measurement that can be implemented with LOCC must be separable, we can restrict our attention to orthonormal *product* bases. For the remainder of this section, unless specified otherwise, we assume that S is an orthonormal product basis of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$.

Let us start by making a few simple observations about the non-disturbing operators and measurements admitted by an orthonormal basis S of $\mathbb{C}^d$. First, note that Equation (3.12) asserts that the matrix E is diagonal in the basis S. Since the states in S are orthonormal, we arrive at the following.

Observation 3.8. *An operator $E \in \mathrm{Pos}(\mathbb{C}^d)$ is non-disturbing for S if and only if*

$$E = \sum_{|\psi\rangle \in S} \lambda_\psi |\psi\rangle\langle\psi| \tag{3.17}$$

for some values of $\lambda_\psi \in \mathbb{R}$.

Consider any operator $E \in \mathrm{Pos}(\mathbb{C}^d)$ that is non-disturbing for S and let E_λ be the projector onto its λ-eigenspace. From the above observation it is easy to see that E_λ is also non-disturbing for S. The fact that $\mathrm{spec}(E_\lambda) \subseteq \{0,1\}$ leads us to our next observation:

Observation 3.9. *If $E \in \mathrm{Pos}(\mathbb{C}^d)$ is non-disturbing for S then for all $\lambda \in \mathrm{spec}(E)$ the projector onto E_λ can be expressed as*

$$E_\lambda = \sum_{|\psi\rangle \in T} |\psi\rangle\langle\psi| \tag{3.18}$$

for some subset $T \subseteq S$.

Note that the subset $T = S$ if and only if E is trivial (*i.e.*, proportional to the identity matrix).

Using the above observations, we can relate the existence of nontrivial local measurements that are non-disturbing for S to the existence of a certain kind of subsets of S.

Lemma 3.10. *Let $S \subseteq \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ be an orthonormal basis and $d := d_A d_B$. Then S admits a nontrivial non-disturbing measurement of the form $\mathcal{M}_A \otimes \mathcal{I}_B$ if and only if there exists a subset $T \subsetneq S$ such that*

$$\sum_{|\psi\rangle \in T} |\psi\rangle\langle\psi| = a \otimes I \tag{3.19}$$

for some $a \in \mathrm{Pos}(\mathbb{C}^{d_A})$. Moreover, if $T \subsetneq S$ satisfies the above condition, then the measurement $\{a, I - a\} \otimes \mathcal{I}_B$ is non-disturbing for S.

Proof. Let $\mathcal{M}_A \otimes \mathcal{I}_B$ be a nontrivial measurement that is non-disturbing for S and let $a \otimes I_{d_B} \in \mathcal{M}_A \otimes \mathcal{I}_B$ be a nontrivial operator. Fix some $\lambda \in \mathrm{spec}(a \otimes I_{d_B})$ and consider the projector $E_\lambda = a_\lambda \otimes I_{d_B} \neq I_d$ onto the λ-eigenspace of $a \otimes I_{d_B}$. By Observation 3.9, we conclude that

$$a_\lambda \otimes I = \sum_{|\psi\rangle \in T} |\psi\rangle\langle\psi| \tag{3.20}$$

for some subset $T \subsetneq S$.

To prove the other direction, consider a subset $T \subsetneq S$ that satisfies Equation (3.19) for some $a \in \mathrm{Pos}(\mathbb{C}^{d_A})$. Since

$$(I_{d_A} - a) \otimes I_{d_B} = I_d - (a \otimes I_{d_B}) = \sum_{|\psi\rangle \in S \setminus T} |\psi\rangle\langle\psi|, \tag{3.21}$$

the operator $(I_{d_A} - a) \otimes \mathcal{I}_B$ is non-disturbing for S by Observation 3.8. Therefore, the projective measurement $\{a, I_{d_A} - a\} \otimes \mathcal{I}_B$ is non-disturbing for S and we are done. $\square$

The above lemma straightforwardly leads to an algorithm for checking whether S admits a nontrivial non-disturbing measurement of the form $\mathcal{M}_A \otimes \mathcal{I}_B$ that examines every subset of S and thus runs in time $2^{O(|S|)}$. In fact it is not necessary to exhaustively check all the subsets of S and there exists a polynomial time algorithm. We now present such an algorithm whose rough outline was first suggested in [DR04]. We later give the first formal analysis of this algorithm and the first rigorous proof of its correctness.

```
Nondisturbing Alice(S)
    Pick |α0, β0⟩ ∈ S
    ã := |α0⟩⟨α0|
    R := {|α0, β0⟩}
    continue := TRUE
    while (continue) do
        continue := FALSE
        T := S \ R;
        for each |α, β⟩ ∈ T do
            if (⟨α|ã|α⟩ ≠ 0) then
                ã := ã + |α⟩⟨α|
                R := R ∪ {|α, β⟩}
                continue := TRUE
    if (R = S) then return FALSE
    else return R
```

In the `while` cycle the set R is expanded by adding any vector $|\alpha, \beta\rangle \in T$ that is non-orthogonal on Alice's part to some vector from the current set R. To test non-orthogonality we compute $\langle\alpha|\tilde{a}|\alpha\rangle$ which takes time $O(|S|^2)$ as $|\alpha\rangle \in \mathbb{C}^{d_A}$ and $d_A \leq |S|$.

Since there are at most $|S|$ elements in $T \subseteq S$, one execution of the `while` cycle takes time $O(|S|^3)$. Finally, the `while` cycle gets repeated $O(|S|)$ times and thus **Nondisturbing Alice**(S) has time complexity $O(|S|^4)$.

The above algorithm determines whether S admits a non-disturbing measurement that Alice can apply to her state space. If the algorithm returns `FALSE` such a measurement does not exist. If the algorithm returns a subset T, then such measurement can be constructed easily form the set T. We now prove the correctness of the above algorithm.

Lemma 3.11. *Suppose **Nondisturbing Alice**(S) = T for some $T \subseteq S$. Then set T satisfies Condition (3.19) in Lemma 3.10 for some $a \in \mathrm{Pos}(\mathbb{C}^{d_A})$.*

Proof. Let a be the projector onto $\mathrm{span}\{|\alpha\rangle : |\alpha, \beta\rangle \in T\}$. By the construction of the set T, we have that $\langle\alpha|\gamma\rangle = 0$ for all $|\alpha, \beta\rangle \in T$ and all $|\gamma, \delta\rangle \in (S \setminus T)$. Therefore, we get that

$$(a \otimes I)|\psi\rangle = \begin{cases} 0 \cdot |\psi\rangle & \text{if } |\psi\rangle \in (S \setminus T) \\ 1 \cdot |\psi\rangle & \text{if } |\psi\rangle \in T \end{cases} \tag{3.22}$$

which implies that

$$a \otimes I = \sum_{|\psi\rangle \in T} |\psi\rangle\langle\psi|. \tag{3.23}$$

This means that T satisfies Condition (3.19) in Lemma 3.10 with $a \in \mathrm{Pos}(\mathbb{C}^{d_A})$ defined above and we are done. $\square$

Lemma 3.12. *If there exists $T \subsetneq S$ satisfying Condition (3.19) in Lemma 3.10 then **Nondisturbing Alice**(S) $\neq$ `FALSE`.*

Proof. Suppose that a subset $T \subsetneq S$ satisfies the Condition (3.19) in Lemma 3.10 for some $a \in \mathrm{Pos}(\mathbb{C}^{d_A})$, *i.e.*

$$a \otimes I = \sum_{|\psi\rangle \in T} |\psi\rangle\langle\psi|. \tag{3.24}$$

Let $|\psi_0\rangle$ be the vector that is picked in the first step of the algorithm. Without loss of generality we can assume that $|\psi_0\rangle \in T$, as otherwise we could replace a with $I - a$ and T with $S \setminus T$.

From Equation (3.24) we see that

$$0 = \langle\gamma, \delta|(a \otimes I)|\gamma, \delta\rangle \tag{3.25}$$

for all $|\gamma, \delta\rangle \in (S \setminus T)$. Since

$$\sum_{|\alpha,\beta\rangle \in T} |\alpha\rangle\langle\alpha| = \mathrm{Tr}_B \Big(\sum_{|\psi\rangle \in T} |\psi\rangle\langle\psi| \Big) = \mathrm{Tr}_B\,(a \otimes I) = d_B \cdot a, \tag{3.26}$$

we can substitute $\frac{1}{d_B} \sum_{|\alpha,\beta\rangle \in T} |\alpha\rangle\langle\alpha|$ for a in Equation (3.25) to obtain

$$0 = \langle\gamma, \delta| \left(\sum_{|\alpha,\beta\rangle \in T} |\alpha\rangle\langle\alpha| \otimes I \right) |\gamma, \delta\rangle = \sum_{|\alpha,\beta\rangle \in T} |\langle\gamma|\alpha\rangle|^2 \tag{3.27}$$

for all $|\gamma, \delta\rangle \in (S \setminus T)$. Therefore, $\langle\alpha|\gamma\rangle = 0$ for all $|\alpha, \beta\rangle \in T$ and all $|\gamma, \delta\rangle \in (S \setminus T)$. Since the algorithm starts with a subset $R = \{|\psi_0\rangle\} \subseteq T$, the condition in Line 10 is never satisfied for $|\gamma, \delta\rangle \in (S \setminus T)$. So it must return some subset of T. □

Combining the above two lemmas with Lemma 3.10, we obtain the following theorem relating the result returned by the algorithm.

Theorem 3.13. *An orthonormal product basis S admits a nontrivial non-disturbing measurement of the form $\mathcal{M}_A \otimes \mathcal{I}_B$ if and only if **Nondisturbing Alice**(S) $\neq$ `FALSE`. Moreover, if **Nondisturbing Alice**(S) = T, then*

$$\sum_{|\psi\rangle \in T} |\psi\rangle\langle\psi| = a \otimes I \tag{3.28}$$

and the projective measurement

$$\mathcal{M}_T := \{a, I - a\} \otimes \mathcal{I}_B \tag{3.29}$$

is non-disturbing for S.

Note that by measuring $\mathcal{M}_T$ from Equation (3.29) we partition the set S into parts T and $S \setminus T$. If the states from S can be discriminated with finite LOCC then so can be the states from any subset $S' \subseteq S$. Therefore, measuring $\mathcal{M}_T$ as a first step in our finite LOCC protocol cannot render the set indistinguishable with finite LOCC, if it was not so to start with.

Similar to **Nondisturbing Alice**(S), we can define a function **Nondisturbing Bob**(S) that determines whether S admits a nontrivial non-disturbing measurement that Bob can perform on his part of the space. Bearing in mind the observations in the above paragraph, it is easy to see that the following algorithm determines whether the states from S can be discriminated with finite LOCC.

```
Discriminate(S)
    if (|S| ≤ 1) then return TRUE
    else
        A := Nondisturbing Alice(S)
        B := Nondisturbing Bob(S)
        if (A ≠ FALSE) then
            return (Discriminate(A) AND Discriminate(S \ A))
        else if (B ≠ FALSE) then
            return (Discriminate(B) AND Discriminate(S \ B))
        else return FALSE
```

Note that each of **Nondisturbing Alice**(S') and **Nondisturbing Bob**(S') either return FALSE or a proper subset of S'. Thus, **Discriminate**(S) will result in at most $O(|S|)$ recursive calls of **Discriminate**. Each such call invokes **Nondisturbing Alice**(S') and **Nondisturbing Bob**(S') that have time complexity $O(|S'|^4)$. Hence, **Discriminate**(S) runs in time $O(|S|^5)$. Therefore, one can decide in polynomial times whether states from an orthogonal basis S can be discriminated with finite LOCC.

3.7 Results of Kleinmann, Kampermann and Bruß

The questions concerning state discrimination with asymptotic LOCC are usually much harder to tackle than their finite LOCC analogues. In this chapter we review the results reported in the recent work [KKB11], which is a must-read for anybody interested in state discrimination with asymptotic LOCC. In Section 3.7.1 we discuss a particular technique for implementing any given measurement as a two-stage process that we use in later chapters. In Section 3.7.2 we give a necessary condition for perfect discrimination of states from a set S with asymptotic LOCC. As a corollary to this condition one can obtain that finite and asymptotic LOCC are equally powerful for perfect discrimination of the states from a full basis (see Corollary 3.16).

3.7.1 Pseudo-weak measurements

In this section we review a technique of splitting up a measurement into a two-stage process, where the measurement in the first stage can be made arbitrarily weak (*i.e.* close to an identity measurement). This technique will allow us to construct interpolated LOCC protocols in Chapter 4 and construct global ε-interpolations of any measurement in Chapter 5.

Definition 3.14. Let $\mathcal{M} = \{F_1, \ldots, F_k\} \subseteq \mathrm{Pos}(\mathbb{C}^n)$ be a measurement. The *pseudo-weak* implementation of $\mathcal{M}$ with interpolation parameters $c_1, \ldots, c_k \geq 0$ is the measurement $\mathcal{M}_1 = \{E_1, \ldots, E_k\}$ where

$$E_i := c\,(c_i I + F_i) \tag{3.30}$$

and $c := (1 + \sum_i c_i)^{-1}$.

The corresponding *recovery* measurements $\mathcal{M}_2^{(i)} = \{E_1^{(i)}, \ldots, E_k^{(i)}\}$ for $i \in [k]$ are then specified by

$$E_j^{(i)} := \begin{cases} \delta_{ij} I & \text{if } c_i = 0, \\ E_i^{-\frac{1}{2}} c \left(c_i F_j + \delta_{ij} F_j\right) E_i^{-\frac{1}{2}} & \text{otherwise.} \end{cases} \tag{3.31}$$

The idea is that we can implement $\mathcal{M}$ by first measuring the pseudo-weak measurement $\mathcal{M}_1$ and then, conditioned on outcome $i \in [k]$, performing the recovery measurement $\mathcal{M}_2^{(i)}$. Moreover, we can adjust the interpolation parameters c_i to make the POVM elements E_i arbitrarily weak (*i.e.*, close to some multiple of the identity operator).

Let us now verify that $\mathcal{M}_1$ and $\mathcal{M}_2^{(i)}$ as defined above are measurements that give us a two-stage implementation of $\mathcal{M}$. First, it is easy to see that $\mathcal{M}_1$ is a valid measurement, as is $\mathcal{M}_2^{(i)}$ when $c_i = 0$. To see this for $\mathcal{M}_2^{(i)}$ with $c_i > 0$, note that the matrix E_i has full rank and hence the inverse $E_i^{-\frac{1}{2}}$ is well-defined. Furthermore,

$$\sum_j E_j^{(i)} = E_i^{-\frac{1}{2}} c \sum_j \left(c_i F_j + \delta_{ij} F_j\right) E_i^{-\frac{1}{2}} = E_i^{-\frac{1}{2}} c\,(c_i I + F_i)\, E_i^{-\frac{1}{2}} = E_i^{-\frac{1}{2}} E_i E_i^{-\frac{1}{2}} = I. \tag{3.32}$$

To see that the pseudo-weak measurement followed by appropriate recovery measurement indeed implements $\mathcal{M}$, let i be the outcome of the pseudo-weak measurement and j be the outcome of the recovery measurement. Then the corresponding POVM element of the two-stage measurement process is given by

$$\sqrt{E_i} E_j^{(i)} \sqrt{E_i} = c\left(c_i + \delta_{ij}\right) F_j. \tag{3.33}$$

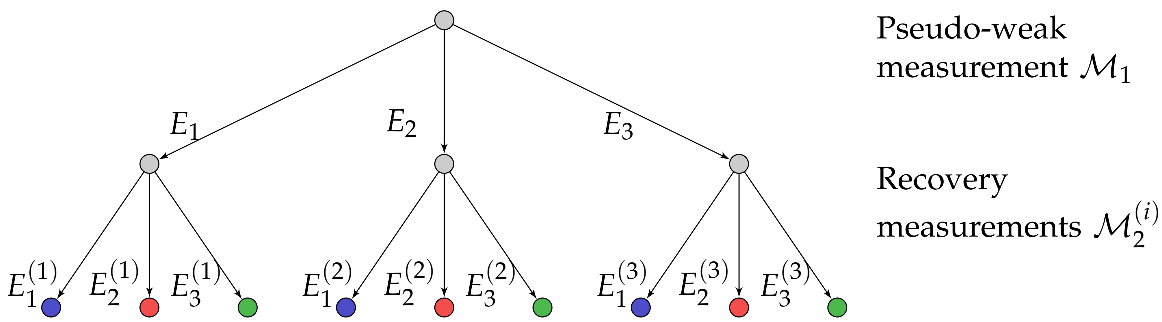


Figure 3.2: An example, where a three-outcome measurement $\mathcal{M}$ is implemented by first measuring a pseudo-weak measurement $\mathcal{M}_1 = \{E_1, E_2, E_3\}$ with *positive* interpolation parameters and then, conditioned on the outcome i, performing the recovery measurement $\mathcal{M}^{(i)} = \{E_1^{(i)}, E_2^{(i)}, E_3^{(i)}\}$. At the end, we perform a coarse graining according to the outcome (color) of the recovery measurement (*i.e.*, the same color outcomes correspond to the same outcome of the original measurement).

Since $\sum_i c\left(c_i + \delta_{ij}\right) F_j = F_j$, coarse graining according to the outcome j of the recovery measurements $\mathcal{M}^{(i)}$ implements measurement $\mathcal{M}$. See Figure 3.2 for an illustration.

The above shows that the pseudo-weak measurement followed by an appropriate recovery measurement reproduces the measurement statistics of $\mathcal{M}$. In fact we can also reproduce the post-measurement states. To discuss post-measurement states of $\mathcal{M}$, we have to specify $\mathcal{M}$ using the measurement operators. Let $\mathcal{M} = \{M_1, \ldots, M_k\}$ be such a specification, where $M_i^\dagger M_i = F_i$. Let us choose the following measurement operators for the pseudo-weak and recovery measurements respectively:

$$\mathcal{M}_1 = \{\sqrt{E_1}, \ldots, \sqrt{E_k}\} \tag{3.34}$$

$$\mathcal{M}_2^{(i)} = \left\{U_{1i}\sqrt{E_1^{(i)}}, \ldots U_{ki}\sqrt{E_k^{(i)}}\right\}. \tag{3.35}$$

where U_{ij} are some isometries to be chosen later. Note that the POVM elements of the above measurements are the same as in Definition 3.14. From Equation (3.33) we see that

$$\left(\sqrt{E_j^{(i)}}\sqrt{E_i}\right)^\dagger \left(\sqrt{E_j^{(i)}}\sqrt{E_i}\right) = c\left(c_i + \delta_{ij}\right) M_j^\dagger M_j. \tag{3.36}$$

According to polar decomposition, for any two matrices M, N, such that $M^\dagger M = N^\dagger N$ there exists an isometry U such that $M = UN$. Therefore, we can choose the isometries U_{ij} in Equation (3.35) so that

$$U_{ij}\sqrt{E_j^{(i)}}\sqrt{E_i} = \sqrt{c\left(c_i + \delta_{ij}\right)} M_j \tag{3.37}$$

Finally, since $\sum_i c\left(c_i + \delta_{ij}\right) = 1$, we conclude that pseudo-weak measurement followed by an appropriate recovery measurement and coarse graining according to the outcome of the recovery measurement, reproduces both the measurement statistics and post-measurement states of $\mathcal{M}$.

3.7.2 A necessary condition for perfect state discrimination by $\overline{\text{LOCC}}$

In this section we review a necessary condition for perfect discrimination of states from some fixed set S with asymptotic LOCC. Currently there are no other such conditions known except the one we develop in Chapter 4 (see Section 4.5.2 for a comparison of the two conditions).

Theorem 3.15 (Kleinmann, Kampermann, and Bruß [KKB11])**.** *Consider a set of states $S = \{\rho_1, \ldots, \rho_n\} \subseteq \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ such that $\bigcap_i \ker \rho_i$ does not contain any nonzero product vector. Then S can be discriminated with asymptotic LOCC only if for all χ with $1/n \leq \chi \leq 1$ there exists a positive semidefinite product operator $E = a \otimes b$ satisfying all of the following:*

1. $\sum_{\rho \in S} \mathrm{Tr}(E\rho) = 1$,
2. $\max_{\rho \in S} \mathrm{Tr}(E\rho) = \chi$,
3. $\mathrm{Tr}(E\rho E\sigma) = 0$ *for all distinct* $\rho, \sigma \in S$.

To prove that a certain set of states S cannot be discriminated using finite LOCC, a common approach is to show that S does not admit nontrivial non-disturbing (see Definition 3.5) operators of the form $a' \otimes I$ or $I \otimes b'$. Note that the last condition in the above theorem states that $a \otimes b$ is non-disturbing for S. Hence, the states from S cannot be discriminated with asymptotic LOCC if for some value of χ the set S does not admit non-disturbing $a \otimes b$.

The following corollary illustrates the usefulness of the above theorem:

Corollary 3.16 (Kleinmann, Kampermann, and Bruß [KKB11])**.** *If a (product) basis S of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ can be discriminated using asymptotic LOCC then it can already be discriminated with finite LOCC.*

In general, deciding whether a given set S of states can be discriminated using asymptotic LOCC can be a daunting task. However, the task becomes easy in the special case where S is a basis. This is because we can combine the above corollary with the polynomial time algorithm for deciding whether a given basis can be discriminated with finite LOCC from Section 3.6.

In the next section we will see how to generalize the above corollary to include sets of mixed product states.

3.8 $\mathrm{LOCC}_{\mathbb{N}}$ vs $\overline{\mathrm{LOCC}}$ for product state discrimination

As discussed in Chapter 2, the class of quantum operations that can be implemented with finite LOCC protocols is not closed. However, it could still contain natural subclasses of operations that are. Identification of such classes would simplify our understanding of the LOCC model, since it is usually much easier to test membership for closed sets.

By Lemma 3.3, we see that Corollary 3.16 from the previous section implies that the class of all von Neumann measurements that can be implemented with finite LOCC is closed. It is hence natural to ask whether a similar result holds for the class of all projective measurements, or for the class of all POVM measurements if we are more ambitious. In this section we take the first step towards answering these questions. We do this by showing that the class of all projective measurements with tensor product operators that can be implemented with finite LOCC is indeed closed.

Theorem 3.17. *Let $\mathcal{M} = \{P_i^A \otimes P_i^B\}_{i\in[n]} \subseteq \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ be a projective measurement. Then $\mathcal{M} \in \overline{\mathrm{LOCC}}$ implies that $\mathcal{M} \in \mathrm{LOCC}_{\mathbb{N}}$.*

We can use Lemma 3.3 and Corollary 3.4 to rephrase the above theorem in terms of state discrimination. When phrased in this way, it becomes clear that this theorem is a generalization of Corollary 3.16 to sets of tensor product mixed states.

Theorem 3.18. *Let $S = \{\rho_i := \tau_i \otimes \sigma_i\}_{i\in[n]} \subseteq \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ be a set of orthogonal states such that $\sum_i \rho_i$ has full rank. If the states from S can be discriminated with asymptotic LOCC then they can be discriminated with finite LOCC.*

An essential ingredient for proving this theorem is a construction of a local operator that is non-disturbing for S. Therefore, we start by better understanding the non-disturbing operators admitted by sets S from the above theorem.

Let S be a set of states satisfying the conditions of the above theorem. Also let $\mathcal{H}^{(i)}$ be the column space (*i.e.*, the image) of ρ_i. Recall that given a matrix M, a subspace $\mathcal{H}$

is said to be M-invariant if $M|\psi\rangle \in \mathcal{H}$ for all $|\psi\rangle \in \mathcal{H}$. Due to the special form of S, we can observe the following.

Observation 3.19. *An operator $M \in \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ is non-disturbing for S if and only if $\mathcal{H}^{(i)}$ is M-invariant for all $i \in [n]$.*

The following lemma provides us with alternative ways to think of M-invariance.

Lemma 3.20. *Let $M \in \mathrm{Herm}(\mathbb{C}^d)$. Also let $\mathcal{H} \subseteq \mathbb{C}^d$ be a subspace of dimension h and $Q = \sum_{i\in[h]} |v_i\rangle\langle i|$, where $\{|v_i\rangle\}_{i\in[h]}$ is some fixed orthonormal basis of $\mathcal{H}$ and $|i\rangle \in \mathbb{C}^h$. Then the following are equivalent:*

1. *$\mathcal{H}$ is M-invariant;*
2. *$MQ = QX$ for some square matrix X;*
3. *$\mathcal{H}$ has an orthonormal basis consisting of eigenvectors of M.*

Proof. Let us first prove that $(1) \Rightarrow (2)$. If $\mathcal{H}$ is M-invariant, then for all $i \in [h]$

$$M|v_i\rangle = \sum_{j\in[h]} x_{ji}|v_j\rangle \tag{3.38}$$

for some $x_{ji} \in \mathbb{C}$. If we let $X := (x_{ij})$, then

$$MQ = \sum_{i,j\in[h]} x_{ji}|v_j\rangle\langle i| = \sum_{j\in[h]} \left(|v_j\rangle \sum_{i\in[h]} \langle i|x_{ji} \right) = \sum_{j\in[h]} |v_j\rangle\langle j|X = QX. \tag{3.39}$$

Next, let us show that $(2) \Rightarrow (3)$. If $MQ = QX$, then $X = Q^\dagger MQ$ as $Q^\dagger Q = I_h$. Since M is Hermitian, so is X and we can consider its spectral decomposition $X = \sum_{i\in[h]} \lambda_i|w_i\rangle\langle w_i|$. Then for all $i \in [h]$

$$MQ|w_i\rangle = QX|w_i\rangle = \lambda_i Q|w_i\rangle. \tag{3.40}$$

Therefore, the vectors $Q|w_i\rangle \in \mathcal{H}$ are eigenvectors of M. Finally, for all $i, j \in [h]$ we have $\langle w_i|Q^\dagger Q|w_j\rangle = \langle w_i|w_j\rangle = \delta_{ij}$. So the set

$$\{Q|w_i\rangle\}_{i\in[h]} \tag{3.41}$$

is an orthonormal basis of $\mathcal{H}$ consisting of eigenvectors of M.

Last, we prove that $(3) \Rightarrow (1)$. Let $\{|u_i\rangle\}_{i\in[h]} \subseteq \mathcal{H}$ be a set of orthogonal eigenvectors of M with corresponding eigenvalues μ_i. Then any vector $u \in \mathcal{H}$ can be expressed as $|u\rangle = \sum_{i\in[h]} c_i|u_i\rangle$ for some $c_i \in \mathbb{C}$. Now we have

$$M|u\rangle = \sum_{i\in[h]} \mu_i c_i |u_i\rangle \in \mathcal{H} \tag{3.42}$$

as desired. $\square$

We now show that whenever $a \otimes b \in \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ is non-disturbing for S, so must be $a \otimes I$ and $I \otimes b$.

Lemma 3.21. *Let $a \otimes b \in \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ and $P = P_A \otimes P_B \in \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ be a projector onto a subspace $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. If $(a \otimes b)P \neq 0$ and $\mathcal{H}$ is $(a \otimes b)$-invariant, then $\mathcal{H}$ is also $(a \otimes I)$ and $(I \otimes b)$-invariant.*

Proof. Fix some orthonormal basis $\{|\alpha_i\rangle\}_{i\in[\dim \mathcal{H}_A]}$ of $\mathcal{H}_A$ and let $Q_A = \sum_i |\alpha_i\rangle\langle i|$, where $|i\rangle \in \mathbb{C}^{\mathrm{rank}(P_A)}$. Define Q_B similarly. Then $P_A = Q_A Q_A^\dagger$ and $P_B = Q_B Q_B^\dagger$. If $\mathcal{H}$ is $(a \otimes b)$-invariant, then

$$(a \otimes b)(Q_A \otimes Q_B) = (Q_A \otimes Q_B)X \tag{3.43}$$

for some square matrix X. Note that X is a tensor product, since $X = (Q_A^\dagger a Q_A) \otimes (Q_B^\dagger b Q_B)$. Since $(a \otimes b)P \neq 0$ we also have that $(a \otimes b)(Q_A \otimes Q_B) \neq 0$. Hence, Equation (3.43) together with the fact that X is a tensor product implies that

$$aQ_A = Q_A X_A \text{ and } bQ_B = Q_B X_B \tag{3.44}$$

for some X_A and X_B such that $X = X_A \otimes X_B$. The above implies that $\mathcal{H}_A$ is a-invariant and $\mathcal{H}_B$ is b-invariant. Since any subspace is invariant under the identity operation, the desired statement follows. $\square$

Having established the above lemma, we now give the proof of the main result of this section.

Proof of Theorem 3.18. To prove the theorem we use induction on $d_A + d_B$. Clearly the states in S can be discriminated with both finite and asymptotic LOCC if $d_A + d_B \leq 3$. We assume that the theorem statement holds for all values $d_A + d_B < m$ for some $m \in \mathbb{N}$.

Suppose that the n states from $S \subseteq \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ with $d_A + d_B \geq m$ can be discriminated with asymptotic LOCC. Then according to Theorem 3.15, for every $1/n \leq \chi \leq 1$ there exists a product operator $E = a \otimes b \in \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ satisfying the following three conditions:

1. $\sum_{\rho \in S} \mathrm{Tr}(E\rho) = 1$,
2. $\max_{\rho \in S} \mathrm{Tr}(E\rho) = \chi$,
3. $\mathrm{Tr}(E\rho E\sigma) = 0$ for all distinct $\rho, \sigma \in S$.

Note that condition (3) asserts that $a \otimes b$ is non-disturbing for S (see Definition 3.5 and Equation (3.13)). Our goal is to choose appropriate value of χ and use Lemma 3.21 to conclude that both $a \otimes I$ and $I \otimes b$ are non-disturbing for S.

Pick any $\chi \in (\frac{1}{n}, \frac{1}{n-1})$ and let $a \otimes b$ be the corresponding operator. Let us now check that $a \otimes b$ is nontrivial and satisfies the hypothesis of Lemma 3.21. The range of the allowed values of χ is chosen so that conditions (1) and (2) together imply that

- $a \otimes b$ is not proportional to the identity matrix (from now on we assume that a is not proportional to the identity matrix as the other case is similar);
- for all $\rho \in S$ we have $E\rho \neq 0$.

For all $i \in [n]$, let $\mathcal{H}^{(i)} = \mathcal{H}_A^{(i)} \otimes \mathcal{H}_B^{(i)}$ be the column space of ρ_i and P_i be the the projector onto it. Then the last item implies that $(a \otimes b)P_i \neq 0$. Since $a \otimes b$ is non-disturbing for S, Observation 3.19 lets us conclude that the subspace $\mathcal{H}^{(i)}$ is $(a \otimes b)$-invariant. Now, for each $i \in [n]$, we can apply Lemma 3.21 for the subspace $\mathcal{H}^{(i)}$ and the matrix $a \otimes b$ to conclude that the subspaces $\mathcal{H}^{(i)}$ are all $(a \otimes I)$-invariant.

Let a_λ be the projector onto the λ-eigenspace of a. Using the equivalence of (1) and (3) in Lemma 3.20, we get that the subspaces $\mathcal{H}^{(i)}$ are $(a_\lambda \otimes I)$-invariant for all $\lambda \in \mathrm{spec}(a)$. Hence, the nontrivial local projective measurement

$$\{a_\lambda : \lambda \in \mathrm{spec}(a)\} \otimes \mathcal{I}_B =: \mathcal{M}_A \otimes \mathcal{I}_B \tag{3.45}$$

is non-disturbing for S.

Suppose we measure the states in S using $\mathcal{M}_A \otimes \mathcal{I}_B$ and obtain outcome $\lambda \in \mathrm{spec}(a)$. If we restrict the unnormalized post-measurement states to the column space of $a_\lambda \otimes I$, we have the following set:

$$S_\lambda := \{(Q_\lambda \otimes I)^\dagger \rho_i (Q_\lambda \otimes I)\}_{i \in [n]} = \left\{(Q_\lambda^\dagger \tau_i Q_\lambda) \otimes \sigma_i\right\}_{i \in [n]} \subseteq \mathbb{C}^{\mathrm{rank}(a_\lambda)} \otimes \mathbb{C}^{d_B}. \tag{3.46}$$

Here,

$$Q_\lambda := \sum_{i \in [\mathrm{rank}(a_\lambda)]} |\lambda_i\rangle\langle i|, \tag{3.47}$$

and $|i\rangle \in \mathbb{C}^{\mathrm{rank}(a_\lambda)}$ and $\{|\lambda_i\rangle\}_i \subseteq \mathbb{C}^{d_A}$ is some orthonormal basis of the λ-eigenspace of a. We now want to use the induction hypothesis to conclude that the states in S_λ can be discriminated with finite LOCC. To do so, we have to check that S_λ is a set of mutually orthogonal states that can be discriminated with asymptotic LOCC and that $\sum_{\rho \in S_\lambda} \rho$ has full rank.

First, since $\sum_{i \in [n]} \rho_i$ is positive semidefinite and has full rank and Q_λ has full column rank, the matrix

$$\sum_{\rho \in S_\lambda} \rho = (Q_\lambda \otimes I)^\dagger \left(\sum_{i \in [n]} \rho_i \right) (Q_\lambda \otimes I) \tag{3.48}$$

has full rank. Suppose that a sequence $\mathcal{R}_1, \mathcal{R}_2, \ldots$ of finite LOCC protocols can be used to certify that the states in S can be discriminated with asymptotic LOCC. Let $\mathcal{R}_i'$ be the finite LOCC protocol in which Alice first embeds her input space $\mathbb{C}^{\mathrm{rank}(a_\lambda)}$ in $\mathbb{C}^{d_A}$ by applying the isometry Q_λ and then the two parties proceed with the protocol $\mathcal{R}_i$. After embedding Alice and Bob have the states $(a_\lambda \otimes I)\rho_i(a_\lambda \otimes I)$ up to normalization. Since the column space, $\mathcal{H}^{(i)}$, of ρ_i is $(a_\lambda \otimes I)$-invariant, the column space of $(a_\lambda \otimes I)$ is contained in $\mathcal{H}^{(i)}$. Therefore, the sequence $\mathcal{R}_1', \mathcal{R}_2', \ldots$ can be used to certify the asymptotic distinguishability of the states from S_λ.

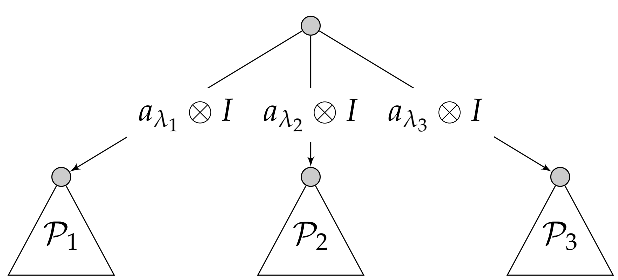


Figure 3.3: An example, where a has three distinct eigenvalues λ_1, λ_2 and λ_3. We first perform the measurement $\mathcal{M}_A \otimes \mathcal{I}_B$ and then, conditioned on the outcome λ_i, proceed with the protocol $\mathcal{P}_i$ that discriminates the states from S_{λ_i}.

Since $\mathcal{M}_A \otimes \mathcal{I}_B$ is non-disturbing for S, the states in S_λ are mutually orthogonal. Finally, as $\mathrm{rank}(a_\lambda) + d_B < d_A + d_B$, we can apply the induction hypothesis to conclude that the states from S_λ can be discriminated by finite LOCC. We are now done since the measurement $\mathcal{M}_A \otimes \mathcal{I}_B$ can be combined with the finite LOCC protocols for discriminating the states from S_λ to give a finite LOCC protocol for discriminating the states in S (See Figure 3.3). $\square$

We now slightly extend the result of Theorem 3.18 by lifting the tensor product requirement for one of the states in S.

Corollary 3.22. *Let $S = \{\rho_i\}_{i\in[n]} \subseteq \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ be a set of orthogonal states such that $\sum_i \rho_i$ has full rank and all the ρ_i except at most one can be expressed as $\rho_i = \sigma_i \otimes \tau_i$. If the states from S can be discriminated with asymptotic LOCC then they can be discriminated with finite LOCC.*

Proof. Suppose that ρ_1 is the state that is not a tensor product. The proof is similar to that of the previous theorem. The difference is that we cannot use Lemma 3.21 to conclude that $\mathcal{H}^{(1)}$ is $(a \otimes I)$ invariant because the lemma can only be applied to tensor product spaces. However, since the orthogonal subspaces $\mathcal{H}^{(2)}, \ldots, \mathcal{H}^{(n)}$ are all $(a \otimes I)$-invariant, $\bigoplus_{i\in\{2,\ldots,n\}} \mathcal{H}^{(i)} = \big(\mathcal{H}^{(1)}\big)^{\perp}$, and a is Hermitian, we can conclude that $\mathcal{H}$ is also $(a \otimes I)$-invariant. $\square$

Although Corollary 3.22 is only a slight generalization of Theorem 3.17, it provides answers to natural questions. For example, let us consider the domino basis, a product basis of $\mathbb{C}^3 \otimes \mathbb{C}^3$ first introduced by [BDF+99] (see Section 4.4 for explicit definitions of the states). Let $S := \{\rho_i\}_{i\in[8]}$, where ρ_8 is the uniform mixture of two fixed domino states and the remaining ρ_i correspond to the remaining six domino states. Since the domino states form an orthogonal product basis, the states ρ_i for $i \in [7]$ are tensor products and $\sum_{i\in[8]} \rho_i$ has full rank. It is not hard to show that for most choices of the two domino states comprising ρ_8, the states from S cannot be discriminated using finite LOCC. For such sets S we can apply Corollary 3.22 to show that perfect discrimination cannot be achieved even with asymptotic LOCC.

The main obstacle in generalizing Theorem 3.17 to all separable projective measurements is the lack of an analogue of Lemma 3.21 for separable projectors P. For example, consider

$$P := \big(|0\rangle\langle 0| \otimes |1\rangle\langle 1|\big) + \big(|1\rangle\langle 1| \otimes |-\rangle\langle -|\big) + \big(|-\rangle\langle -| \otimes |2\rangle\langle 2|\big) \tag{3.49}$$

and $a \otimes b := |1\rangle\langle 1| \otimes |-\rangle\langle -|$. Let $\mathcal{H}$ be the space onto which P projects. Although $\mathcal{H}$ is $(a \otimes b)$-invariant, it is neither $(a \otimes I)$- nor $(I \otimes b)$-invariant, since $(a \otimes I)|-,2\rangle \notin \mathcal{H}$ and $(I \otimes b)|0,1\rangle \notin \mathcal{H}$.

Therefore, although Theorem 3.17 makes partial progress, the general question of whether the set of POVM measurements implementable by $\mathrm{LOCC}_{\mathbb{N}}$ is closed, remains wide open.

Chapter 4

A framework for bounding nonlocality of state discrimination

4.1 Introduction

In this chapter we study the task of discriminating pure bipartite states from some known set S by LOCC. Building on the work in the paper *Quantum nonlocality without entanglement* [BDF+99], we provide a framework for lower bounding the error probability of any LOCC protocol aiming at discriminating the states from S. We apply our framework to an orthonormal product basis known as the domino states and obtain an alternative and simplified bound on how well these states can be discriminated using LOCC. We generalize this result for similar bases in larger dimensions, as well as the "rotated" domino states, resolving a long-standing open question [BDF+99].

This chapter is based on the results obtained in collaboration with Andrew Childs, Debbie Leung, and Maris Ozols [CLMO12]. We start by explaining our motivation in Section 4.1.1 and proceed by more formally stating our contributions and explaining the layout of the rest of this chapter in Section 4.1.2. The material in this chapter relies on the concepts introduced in Chapter 3. In particular, we analyze LOCC protocols using their protocol trees as described in Section 3.2.4.

4.1.1 Motivation

The 1999 paper *Quantum nonlocality without entanglement* [BDF+99] exhibits an orthonormal basis $S \subseteq \mathbb{C}^3 \otimes \mathbb{C}^3$ of product states, known as domino states, shared between two separated parties. When the parties are restricted to perform only local quantum operations and classical communication (LOCC), it is impossible to discriminate the domino states arbitrarily well [BDF+99]. In such cases we say that perfect discrimination cannot be achieved with *asymptotic* LOCC. Moreover, [BDF+99] also quantifies the extent to which any LOCC protocol falls short of perfect discrimination of the domino states.

This result spurred interest in state discrimination with LOCC. Several alternative proofs [GV01, WH02, Coh07] of the impossibility of perfect LOCC discrimination of the domino states were given along with many other results concerning perfect state discrimination (*e.g.*, [BDM+99, WSHV00, GKR+01, GV01, VSPM01, CY01, CY02, WH02, DMS+03, CL03, HSSH03, HM03, Fan04, GKRS04, Che04, CL04, JCY05, Wat05a, Nat05, NC06, DFJY07, FS09, DFXY09, DXY10]). However, the problem of asymptotic LOCC state discrimination has not received much attention since the initial study of nonlocality without entanglement [BDF+99].

The main motivation for our work is to better understand the phenomenon of quantum nonlocality without entanglement. More concretely, our goals are to

- simplify the original proof,
- render the technique applicable to a wider class of sets of bipartite states,
- exhibit new classes of product bases that cannot be *asymptotically* (as opposed to just perfectly) discriminated with LOCC, and
- investigate the possibility of larger gaps between the sets of LOCC and separable operations.

In particular, we seek to exhibit quantitative gaps between the classes of LOCC and separable operations. Separable operations often serve as a relaxation of LOCC operations and such gaps show how imprecise this relaxation can be. The rationale behind this relaxation is that separable operations have a clean mathematical description whereas LOCC operations can be much harder to understand.

There is also an operational motivation to quantify the difference between separable measurements and those implemented by asymptotic LOCC: the former are precisely the measurements that cannot generate entangled states, while the latter are those that do not require entanglement to implement [BDF+99, KTYI07, Koa09]. Thus, a separable measurement that cannot be implemented by asymptotic LOCC uses entanglement irreversibly.

4.1.2 Our contributions

In this chapter, we develop a framework for obtaining quantitative results on the hardness of quantum state discrimination by LOCC. More precisely, we provide a method for proving a lower bound on the error probability of any LOCC measurement for discriminating states from a given set S. Any strictly positive lower bound implies that the states from S cannot be even asymptotically discriminated with LOCC.

Our first main contribution (Theorem 4.7) is that any LOCC measurement for discriminating states from a set S errs with probability $p_{\text{error}} \geq \frac{2}{27}\frac{\eta^2}{|S|^5}$, where η is a constant that depends on S (see Definition 4.4). Intuitively, η measures the nonlocality of S.

Our second main contribution is a systematic method for bounding the nonlocality constant η for a large class of product bases. Together with the above theorem, this lets us quantify the hardness of LOCC discrimination for the following bases of product states:

1. *domino states*, the original set of nine states in 3×3 dimensions first considered in [BDF+99], have $p_{\text{error}} \geq 1.9 \times 10^{-8}$;
2. *domino-type states*, a generalization of domino states to higher dimensions corresponding to tilings of a rectangular $d_A \times d_B$ grid by tiles of size at most two, have $p_{\text{error}} \geq 1/(216D^2 d_A^5 d_B^5)$, where D is a property of the tiling that we call "diameter";
3. *θ-rotated domino states*, a 1-parameter family that includes the domino states and the standard basis as extreme cases, have $p_{\text{error}} \geq 2.4 \times 10^{-11} \sin^2 2\theta$ (determining whether these states can be discriminated perfectly by LOCC and finding a lower bound on the probability of error were left as open problems in [BDF+99]).

The rest of this chapter is organized as follows. In Section 4.2 we introduce our general framework for lower bounding the error probability of LOCC measurements, and in Section 4.2.5 we prove Theorem 4.7. In Section 4.3 we consider the case where S is a product basis and propose a method for bounding the nonlocality constant η by another quantity that we call "rigidity." Our approach is based on a description of sets of bipartite states in terms of tilings. In Section 4.4 we define the three classes of states mentioned above and prove a bound on the rigidity of the domino states; bounds on the rigidity of the domino-type states and the rotated domino states appear in Appendix A and B, respectively. Finally, we discuss limitations of our framework in Section 4.5 and conclude with a discussion of open problems in Section 4.6.

4.2 Framework

In this section we introduce a framework for proving lower bounds on the error probability of any LOCC measurement for discriminating bipartite states from a given set

$$S := \{|\psi_1\rangle, \ldots, |\psi_n\rangle\} \subseteq \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}. \tag{4.1}$$

We make no assumptions about the states $|\psi_i\rangle$. In particular, they need not be product states or be mutually orthogonal.

From now on, $\mathcal{P}$ denotes an arbitrary LOCC protocol for discriminating states from S. In rough outline our argument proceeds as follows:

1. We modify $\mathcal{P}$ so that it can be stopped when a specific amount of information ε has been obtained (see Section 4.2.1). This is done by terminating the protocol prematurely and possibly making the last measurement less informative (see Section 4.2.2).
2. When the information gain is ε, we lower bound a measure of disturbance (defined in Section 4.2.3) by $\eta\varepsilon$ for some constant η (see Section 4.2.4).
3. We show that at least two of the possible initial states have become nonorthogonal at this stage of the protocol, and we infer a lower bound on the error probability of $\mathcal{P}$ (see Section 4.2.5).

Our framework reuses some ideas of the original approach [BDF+99]. However, instead of mutual information, we quantify how much an LOCC protocol has learned about the state using error probability. This allows us to replace the long mutual information analysis in the original paper with a simple application of Helstrom's bound. The idea of relating information gain and disturbance also comes from [BDF+99]. Here, we analyze this tradeoff using the nonlocality constant (see Definition 4.4) which can be applied to any set of states. In Section 4.3 we give a method for lower bounding the nonlocality constant that applies specifically when S is an *orthonormal basis* of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. In Section 4.4 we apply this method for the domino states and some other related bases.

4.2.1 Interpolated LOCC protocol

Consider an arbitrary node in the tree representing the protocol $\mathcal{P}$. Let m be the corresponding measurement record and let $A \otimes B$ denote the Kraus operator that is applied to the initial state when this node is reached. Note that the output dimensions of operators A and B could be arbitrary.

The initial state $|\psi_k\rangle$ yields measurement record m with probability

$$p(m|\psi_k) := \operatorname{Tr}\big[(A \otimes B)^\dagger (A \otimes B)|\psi_k\rangle\langle\psi_k|\big] = \langle\psi_k|(a \otimes b)|\psi_k\rangle \tag{4.2}$$

where $a := A^\dagger A \in \operatorname{Pos}(\mathbb{C}^{d_A})$ and $b := B^\dagger B \in \operatorname{Pos}(\mathbb{C}^{d_B})$. Note that we need not concern ourselves with the arbitrary output dimensions of A and B from this point onward. We use Bayes's rule and the uniformity of the probabilities $p(\psi_k)$ to obtain the probability that the initial state was $|\psi_k\rangle$ conditioned on the measurement record being m:

$$p(\psi_k|m) = \frac{p(\psi_k)p(m|\psi_k)}{\sum_{j=1}^n p(\psi_j)p(m|\psi_j)} = \frac{\langle\psi_k|(a \otimes b)|\psi_k\rangle}{\sum_{j=1}^n \langle\psi_j|(a \otimes b)|\psi_j\rangle}. \tag{4.3}$$

At the root, the measurement record m is the empty string and $p(\psi_k|m) = \frac{1}{n}$ for all k. As we proceed toward the leaves, these probabilities fluctuate away from $\frac{1}{n}$. For example, if $\mathcal{P}$ discriminates the states perfectly, the distribution reaches a Kronecker delta function.

For a given node m let us define

$$p_{\max}(m) := \max_{k \in [n]} p(\psi_k|m). \tag{4.4}$$

Let $\varepsilon := p_{\max}(m) - \frac{1}{n}$. Then ε characterizes the non-uniformity of the distribution $p(\psi_k|m)$ and thus the amount of information learned about the input state. The next theorem shows that we can modify the protocol $\mathcal{P}$ so that it can be stopped when some but not too much information has been learned. While this idea originates from [BDF+99], we use a specific result from [KKB11].

Theorem 4.1 (Kleinmann, Kampermann, Bruß [KKB11])**.** *Let $\mathcal{P}$ be an LOCC protocol for discriminating states from a set S of size n. For any $\varepsilon > 0$ there exists an LOCC protocol $\mathcal{P}_\varepsilon$ that has the same success probability as $\mathcal{P}$, but each branch of $\mathcal{P}_\varepsilon$ has a node m such that either*

$$p_{\max}(m) = \frac{1}{n} + \varepsilon \qquad \text{or} \qquad p_{\max}(m) < \frac{1}{n} + \varepsilon \text{ and } m \text{ is a leaf of } \mathcal{P}. \tag{4.5}$$

Proof sketch. Fix $\varepsilon > 0$. Then in each branch of $\mathcal{P}$ either $p_{\max} \leq \frac{1}{n} + \varepsilon$ for all nodes in that branch or there exists a node at which $\frac{1}{n} + \varepsilon < p_{\max}$. For each branch of the latter kind we identify the closest node to the root, say u, that has a child for which $\frac{1}{n} + \varepsilon < p_{\max}$. To obtain the the interpolated protocol $\mathcal{P}_\varepsilon$, we modify the measurement at all such vertices u using the technique described in Section 3.7.1. We now outline the modification procedure.

Let $v_1, \ldots, v_m$ be all the children of u. Also let $\mathcal{M} = \{F_1, \ldots, F_m\}$ be the measurement at node u, where F_i leads to node v_i. Then for some $i \in [m]$ we have that

$$p_{\max}(u) < \frac{1}{n} + \varepsilon < p_{\max}(v_i), \tag{4.6}$$

which means that the measurement outcome corresponding to F_i is too informative. To rectify this, we break up the measurement $\mathcal{M}$ into two steps so that each individual measurement is less informative.

In the interpolated protocol $\mathcal{P}_\varepsilon$ we replace $\mathcal{M}$ with its pseudo-weak implementation $\mathcal{M}_1 = \{E_1, \ldots, E_m\}$ with interpolation parameters $c_1, \ldots, c_m$ to be choosen later (see Definition 3.14). To ensure equivalence with the original protocol, after measuring $\mathcal{M}_1$ and obtaining outcome i, we perform the recovery measurement $\mathcal{M}_2^{(i)}$ (see Section 3.7.1). We represent the outcomes of $\mathcal{M}_1$ by new nodes $\tilde{v}_1, \ldots, \tilde{v}_m$ while the outcomes of $\mathcal{M}_2^{(i)}$ lead to the original nodes $v_1, \ldots, v_m$ (see Figure 4.1). Note that $\mathcal{M}$ is of the form $\mathcal{M}_A \otimes I_B$ or $I_A \otimes \mathcal{M}_B$. As can be seen from Definition 3.14, the measurements $\mathcal{M}_1$ and $\mathcal{M}_2^{(i)}$ also take the same form and can thus be implemented locally.

Let us now explain how to choose the values of the interpolation parameters c_i. Recall from Definition 3.14 that the POVM elements of the pseudo-weak measurement are given by

$$E_i = c(c_i I + F_i), \tag{4.7}$$

where $c = (1 + \sum_i c_i)^{-1}$. We set $c_i = 0$ for all i such that $\frac{1}{n} + \varepsilon \geq p_{\max}(v_i)$ in the original protocol $\mathcal{P}$. We now argue that the remaining c_i can be chosen so that $p_{\max}(\tilde{v}_i) = \frac{1}{n} + \varepsilon$. Note that E_i changes continuously from cF_i to I as we change c_i from 0 to ∞ and keep the remaining c_j fixed. Since $p_{\max}$ is continuous on nonzero operators[1], the interpolation parameter c_i can be chosen to achieve any value strictly between $p_{\max}(v_i)$ and $p_{\max}(u)$. Since $\frac{1}{n} + \varepsilon$ lies strictly between the two values, c_i can be chosen so that $p_{\max}(\tilde{v}_i) = \frac{1}{n} + \varepsilon$.

[1] Although $p_{\max}$ takes nodes as its arguments, we can think of it as a function of positive semidefinite operators. This can be done by identifying any node u, with the POVM element $a \otimes b$ that has resulted since the start of the protocol upon reaching u.

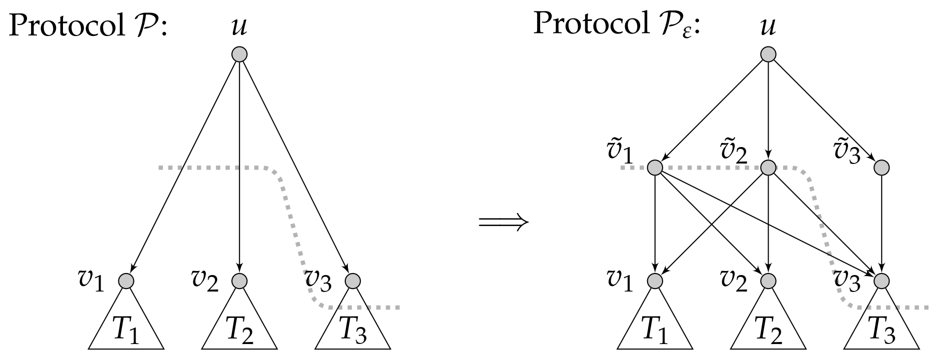


Figure 4.1: The protocol tree before (left) and after (right) splitting the measurement at node u into two steps. (The graph on the right has been condensed for clarity, but it can be expanded into a tree by making a new copy of subtree T_i for each incoming arc in v_i.) The amount of information learned in the first step is controlled by diluting the measurement operators, and the purpose of the second step is to complete the original measurement. The dotted line corresponds to the end of stage I (see Definition 4.2).

From Equations (4.3) and (4.4), we see that the value of $p_{\max}(m)$ does not change if the corresponding POVM element $a \otimes b$ is multiplied by a non-zero constant. Changing the interpolation parameters c_j for $j \neq i$ has exactly this effect on the operator E_i. Hence, we can independently adjust the interpolation parameters so that $p_{\max}(\tilde{v}_i) = \frac{1}{n} + \varepsilon$ for all i for which $\frac{1}{n} + \varepsilon < p_{\max}(v_i)$ in the original protocol $\mathcal{P}$.

After replacing the measurement at u with the above described two-stage measurement process, we proceed according to the original protocol as shown in Figure 4.1. As explained in Section 3.7.1, the above two-stage measurement process reproduces both the measurement statistics and the post-measurement states of the original measurement. Therefore, the interpolated protocol $\mathcal{P}_\varepsilon$ will have the same success probability as the original protocol $\mathcal{P}$. Also, the modification procedure has ensured that each branch of $\mathcal{P}_\varepsilon$ has a node satisfying condition (4.5). $\square$

In the context of discriminating states from an orthonormal basis, the possibility of interpolating a protocol to obtain some but not too much information is what distinguishes LOCC measurements from separable ones. In particular, a separable measurement for a set of states that cannot be distinguished by asymptotic LOCC cannot be divided into two steps, with the first yielding information precisely ε and the second completing the measurement (see Section 4.2.1 for further details).

4.2.2 Stopping condition

To control how much information the protocol has learned, we fix some $\varepsilon > 0$ and stop the execution of $\mathcal{P}_\varepsilon$ when we reach a node m that satisfies the conditions in Equation (4.5).

Definition 4.2. In any given execution of $\mathcal{P}_\varepsilon$, we say that *stage I* is complete at the earliest point when Equation (4.5) is satisfied.

We choose $\varepsilon < \frac{1}{n(n-1)}$ in our analysis. Operationally, this means that none of the n states has been eliminated at the end of stage I, since

$$\min_{k\in[n]} p(\psi_k|m) \geq 1 - (n-1)p_{\max}(m) \geq \frac{1}{n} - (n-1)\varepsilon > 0. \tag{4.8}$$

This allows us to use Helstrom's bound to lower bound the probability of error (see Section 4.2.5). It also ensures that the disturbance measure $\delta_S(a \otimes b)$ introduced in Section 4.2.3 is well defined at m. All constraints imposed on the distribution $p(\psi_k|m)$ are summarized in Figure 4.2.

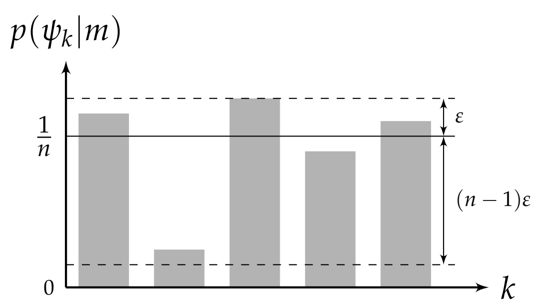


Figure 4.2: Probability distribution $p(\psi_k|m)$ at the end of stage I. For all k we have $\frac{1}{n} + \varepsilon \geq p(\psi_k|m) \geq \frac{1}{n} - (n-1)\varepsilon > 0$ where the first inequality is tight for some k.

Since the error probability of the protocol $\mathcal{P}_\varepsilon$ is a weighted average of error probabilities of individual branches, it suffices to lower bound these individual error probabilities. For any branch that terminates without a node satisfying

$$p_{\max}(m) = \frac{1}{n} + \varepsilon\,, \tag{4.9}$$

we can put a large lower bound on the error probability. In particular, for the optimal choice $\varepsilon = \frac{2}{3}\frac{1}{n(n-1)}$ of Theorem 4.7 with $n \geq 2$,

$$p_{\mathrm{error}}(m) \geq 1 - p_{\max}(m) > 1 - \left(\frac{1}{n} + \varepsilon\right) = 1 - \frac{1}{n} - \frac{2}{3}\frac{1}{n(n-1)} \geq \frac{1}{6}, \tag{4.10}$$

which is much higher than the lower bound we obtain for other branches. We now consider the remaining case where stage I ends with a node satisfying Equation (4.9).

4.2.3 Measure of disturbance

Now we show that at least two possible post-measurement states $(A \otimes B)|\psi_i\rangle$ and $(A \otimes B)|\psi_j\rangle$ are nonorthogonal at the end of stage I, and lower bound their overlap quantitatively. Assuming that the initial state was $|\psi_i\rangle \in S$, the normalized post-measurement state at a node with corresponding measurement operator $A \otimes B$ is

$$|\phi_i\rangle := \frac{(A \otimes B)|\psi_i\rangle}{\sqrt{\langle\psi_i|(a \otimes b)|\psi_i\rangle}} \tag{4.11}$$

where $a := A^\dagger A$ and $b := B^\dagger B$. We are interested in the overlaps $\langle\phi_i|\phi_j\rangle$ rather than the actual post-measurement states. Hence, from now on we use the POVM elements a and b instead of the measurement operators A and B.

Definition 4.3. The *disturbance* caused by the operator $a \otimes b$ on the set of states S is defined as

$$\delta_S(a \otimes b) := \max_{i \neq j} |\langle\phi_i|\phi_j\rangle| = \max_{i \neq j} \frac{|\langle\psi_i|(a \otimes b)|\psi_j\rangle|}{\sqrt{\langle\psi_i|(a \otimes b)|\psi_i\rangle\langle\psi_j|(a \otimes b)|\psi_j\rangle}}. \tag{4.12}$$

We use $\delta_S(a \otimes b)$ only in the context where $a \otimes b$ is the operator corresponding to an end node of stage I. In this case, from Equations (4.8) and (4.3) we get that $0 < \min_{k\in[n]} p(\psi_k|m) = \min_{k\in[n]} \frac{\langle\psi_k|(a\otimes b)|\psi_k\rangle}{\sum_{j=1}^{n}\langle\psi_j|(a\otimes b)|\psi_j\rangle}$. Hence $\langle\psi_i|(a \otimes b)|\psi_i\rangle > 0$ for all $i \in [n]$, so δ_S is well-defined.

Note that $\delta_S(a \otimes b)$ measures the nonorthogonality of the post-measurement states $|\phi_i\rangle$. If the initial states $|\psi_i\rangle$ are orthogonal, $\delta_S(a \otimes b)$ characterizes the disturbance imparted to the states at the end of stage I in the branch corresponding to $a \otimes b$.

4.2.4 Disturbance/information gain trade-off

Now we define the *nonlocality constant* η. It relates the disturbance caused at the end of stage I, minimized over all branches (see Definition 4.3), to the amount of information learned, ε.

Definition 4.4. The *nonlocality constant* η of S is the supremum over all η' such that

$$\eta' \cdot \left(\frac{\max_{k\in[n]} \langle \psi_k | (a \otimes b) | \psi_k \rangle}{\sum_{j\in[n]} \langle \psi_j | (a \otimes b) | \psi_j \rangle} - \frac{1}{n} \right) \leq \delta_S(a \otimes b) \tag{4.13}$$

for all $a \in \mathrm{Pos}(\mathbb{C}^{d_A}), b \in \mathrm{Pos}(\mathbb{C}^{d_B})$ for which $\langle \psi_i | (a \otimes b) | \psi_i \rangle \neq 0$ for all $i \in [n]$.

Equivalently, if $G_{ij} := \langle \psi_i | (a \otimes b) | \psi_j \rangle$ for $i, j \in [n]$ then

$$\eta := \inf_{a,b} \left\{ \frac{\max_{i\neq j} \dfrac{|G_{ij}|}{\sqrt{G_{ii} G_{jj}}}}{\dfrac{\max_k G_{kk}}{\sum_{j=1}^n G_{jj}} - \dfrac{1}{n}} \right\} \tag{4.14}$$

where the infimum is over all $a \in \mathrm{Pos}(\mathbb{C}^{d_A})$ and $b \in \mathrm{Pos}(\mathbb{C}^{d_B})$ such that $G_{ii} \neq 0$ for all $i \in [n]$.

The nonlocality constant η depends only on the set of states S and applies to all branches of the protocol.

When stage I ends, each branch satisfies the condition in Equation (4.9) with some $\varepsilon \in \left(0, \frac{1}{n(n-1)}\right)$. Consider a branch with end node m and corresponding measurement operator $a \otimes b$. We now use the shorthand δ to denote $\delta_S(a \otimes b)$, the disturbance caused in that branch.

Lemma 4.5 (Disturbance/information gain trade-off)**.** *For every branch, the amount of information ε learned at the end of stage I lower bounds the disturbance δ as*

$$\eta\, \varepsilon \leq \delta \tag{4.15}$$

where η is the nonlocality constant of S.

Proof. This immediately follows from the definitions of ε and η:

$$\eta\, \varepsilon = \eta \left(\max_{k\in[n]} p(\psi_k | m) - \frac{1}{n} \right) = \eta \left(\frac{\max_{k\in[n]} \langle \psi_k | (a \otimes b) | \psi_k \rangle}{\sum_{j=1}^n \langle \psi_j | (a \otimes b) | \psi_j \rangle} - \frac{1}{n} \right) \leq \delta \tag{4.16}$$

where we have used Equations (4.9), (4.3), and (4.13). □

4.2.5 Lower bounding the error probability

In this section we use Lemma 4.5 to lower bound the error probability of any LOCC measurement for discriminating states from the set S.

Equation (4.15) implies that at the end of stage I, for every branch, there are two distinct post-measurement states $|\phi_i\rangle$ and $|\phi_j\rangle$ such that

$$|\langle\phi_i|\phi_j\rangle| = \delta \geq \eta\,\varepsilon. \tag{4.17}$$

As discussed in Section 4.2.2, our choice of ε guarantees that $p(\psi_i|m)$ and $p(\psi_j|m)$ are both strictly positive. Thus we can use the following result to lower bound the error probability for each branch.

Fact (Helstrom bound [Hel76, p.113])**.** *Suppose we are given state $|\Phi_0\rangle$ with probability q_0 and state $|\Phi_1\rangle$ with probability $q_1 = 1 - q_0$. Any measurement trying to discriminate the two cases errs with probability at least*

$$Q(q_0, q_1, \delta) := \frac{1}{2}\Big(1 - \sqrt{1 - 4q_0q_1\delta^2}\Big) \geq q_0q_1\delta^2, \tag{4.18}$$

where $\delta := |\langle\Phi_0|\Phi_1\rangle|$ is the overlap between the two states, and the inequality follows from the fact that $1 - \sqrt{1 - x^2} \geq \frac{1}{2}x^2$ for $x \in [0, 1]$.

As ε increases, the error probability for a specific branch changes in two opposite ways. On one hand, two post-measurement states $|\phi_i\rangle$ and $|\phi_j\rangle$ have overlap $\delta \geq \eta\varepsilon$, and this lower bound increases with ε. On the other hand, the probabilities of these two states, $p(\psi_i|m)$ and $p(\psi_j|m)$, are lower bounded by a function that decreases with ε. Balancing these two effects, the choice $\varepsilon = \frac{2}{3}\frac{1}{n(n-1)}$ gives a lower bound on the error probability as follows.

Lemma 4.6. *Let S be a set of quantum states in $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ of size $n \geq 2$. For any LOCC measurement discriminating states drawn uniformly from S, each branch errs with probability*

$$p_{\text{error}} \geq \frac{2}{27}\frac{\eta^2}{n^5} \tag{4.19}$$

where η is the nonlocality constant of S.

Proof. At the end of stage I, for each branch, there are two post-measurement states $|\Phi_0\rangle$ and $|\Phi_1\rangle$ with overlap δ. Let p_0 and p_1 be the posterior probabilities of these states. To lower bound the error probability of $\mathcal{P}_\varepsilon$ (and thus that of $\mathcal{P}$), we give Alice and Bob extra power at this point:

- if the actual input state does not lead to $|\Phi_0\rangle$ or $|\Phi_1\rangle$, we assume that Alice and Bob succeed with certainty;
- otherwise Alice and Bob are allowed to perform the best joint measurement to discriminate the states $|\Phi_0\rangle$ and $|\Phi_1\rangle$.

For fixed ε and probabilities p_0 and p_1, we can lower bound the error probability by the following expression:

$$P(p_0, p_1, \varepsilon) := (p_0 + p_1) \cdot Q\Big(\tfrac{p_0}{p_0+p_1}, \tfrac{p_1}{p_0+p_1}, \delta\Big). \tag{4.20}$$

Using Equation (4.18) and the inequality $\delta \geq \eta\,\varepsilon$ from Lemma 4.5, we get that

$$P(p_0, p_1, \varepsilon) \geq \frac{p_0 p_1}{p_0 + p_1}(\eta\,\varepsilon)^2. \tag{4.21}$$

Recall that we stop the protocol at a point where we are guaranteed that $0 < \varepsilon < \frac{1}{n(n-1)}$ and, by Equations (4.8) and (4.9),

$$\frac{1}{n} - (n-1)\varepsilon \leq p_i \leq \frac{1}{n} + \varepsilon \tag{4.22}$$

for all i. Given these constraints on p_0 and p_1, we can choose the ε that maximizes $P(p_0, p_1, \varepsilon)$ and guarantee that the error probability in the branch of the LOCC protocol being considered satisfies

$$p_{\text{error}} \geq \max_{\varepsilon \in \left(0, \frac{1}{n(n-1)}\right)} \quad \min_{p_0, p_1 \in \left[\frac{1}{n} - (n-1)\varepsilon, \frac{1}{n} + \varepsilon\right]} \quad P(p_0, p_1, \varepsilon). \tag{4.23}$$

From Equation (4.21) we get

$$p_{\text{error}} \geq \max_{\varepsilon \in \left(0, \frac{1}{n(n-1)}\right)} \quad \min_{p_0, p_1 \in \left[\frac{1}{n} - (n-1)\varepsilon, \frac{1}{n} + \varepsilon\right]} \quad \frac{p_0 p_1}{p_0 + p_1}(\eta\,\varepsilon)^2. \tag{4.24}$$

The minimum is attained when $p_0 = p_1 = \frac{1}{n} - (n-1)\varepsilon$ (*i.e.*, the probabilities are equal and as small as possible), so the problem simplifies to

$$p_{\text{error}} \geq \max_{\varepsilon \in \left(0, \frac{1}{n(n-1)}\right)} \frac{1}{2}\left(\frac{1}{n} - (n-1)\varepsilon\right)(\eta\,\varepsilon)^2 \geq \frac{2}{27}\frac{\eta^2}{n^3(n-1)^2} \geq \frac{2}{27}\frac{\eta^2}{n^5} \tag{4.25}$$

where the value

$$\varepsilon = \frac{2}{3}\frac{1}{n(n-1)} \tag{4.26}$$

achieves the maximum. $\square$

The lower bound for the error probability for each branch implies the same lower bound for the LOCC measurement for discriminating states from the set S:

Theorem 4.7. *Let S be a set of quantum states in $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ of size $n \geq 2$. Any LOCC measurement for discriminating states drawn uniformly from S errs with probability*

$$p_{\text{error}} \geq \frac{2}{27} \frac{\eta^2}{n^5} \tag{4.27}$$

where η is the nonlocality constant of S.

Theorem 4.7 shows that any LOCC protocol for discriminating states from S errs with probability proportional to η^2, justifying the name "nonlocality constant."

4.3 Bounding the nonlocality constant

The framework described in Section 4.2 reduces the problem of bounding the error probability for discriminating bipartite states by LOCC to the one of bounding the nonlocality constant η (see Theorem 4.7). This reduction holds for any set of pure states S. In this section we assume that S is an *orthonormal basis* of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ and provide tools for bounding the nonlocality constant. In particular, we bound η in terms of another quantity that we call "rigidity".

For the remainder of the chapter we represent pure states from $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ using "tiles" in a $d_A \times d_B$ grid. We first introduce some notations related to tilings in Section 4.3.1. Then we define rigidity and relate it to the nonlocality constant η in Section 4.3.2. Section 4.3.3 provides a tool, the "pair of tiles" lemma, that we use to bound rigidity for specific sets of states in Section 4.4.

4.3.1 Definitions

Given a fixed orthonormal basis $\{|i\rangle : i \in [d]\}$, define the *support* of a pure state $|\psi\rangle \in \mathbb{C}^d$ as

$$\operatorname{supp}|\psi\rangle := \{i \in [d] : \langle i|\psi\rangle \neq 0\}. \tag{4.28}$$

If $|\psi\rangle \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ then $\operatorname{supp}|\psi\rangle \subseteq [d_A] \times [d_B]$. Consider $[d_A] \times [d_B]$ as a rectangular grid of size $d_A \times d_B$. Any region that corresponds to a submatrix of this grid is called a tile. More formally, a *tile* is a subset $T \subseteq [d_A] \times [d_B]$ such that $T = R \times C$ for some $R \subseteq [d_A]$ and $C \subseteq [d_B]$. (Note that a tile is not necessarily a contiguous region of the grid.) We use $\mathrm{rows}(T) = R$ and $\mathrm{cols}(T) = C$ to denote the *rows* and *columns* of this tile, respectively, and we use $|T|$ to denote the *size* or the *area* of T. If $|\psi\rangle = |\alpha\rangle|\beta\rangle$ is a product state, then $\mathrm{supp}\,|\psi\rangle = \mathrm{supp}\,|\alpha\rangle \times \mathrm{supp}\,|\beta\rangle$ and thus $\mathrm{supp}\,|\psi\rangle$ is a tile, which we call the *tile induced by* $|\psi\rangle$.

We say that an orthonormal set of product states $S \subseteq \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ *induces a tiling* of a $d_A \times d_B$ grid if the tiles induced by the states in S are either disjoint or identical. Note that if S is an orthonormal basis of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, then a tile of area L is induced by L states that form a *basis* of that tile. In a *domino-type* tiling, every tile has area 1 or 2 and the states inducing the tiles of size 2 are of the form $|i \pm j\rangle|k\rangle$ or $|k\rangle|i \pm j\rangle$, where $|i \pm j\rangle := (|i\rangle \pm |j\rangle)/\sqrt{2}$.

For a given tiling T of a $d_A \times d_B$ grid let us define the corresponding *row graph* as follows: its vertex set is $[d_A]$ with two vertices i and j adjacent if and only if there exists a column c such that (i, c) and (j, c) belong to the same tile. The *column graph* of a tiling is defined similarly. We say that a tiling is *irreducible* if its row graph and its column graph are both connected. The *diameter* of a tiling T is the maximum of the diameters of its row and column graphs. See Figure 4.3 for an example.

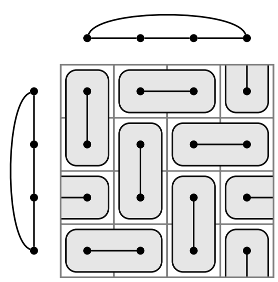


Figure 4.3: A domino-type tiling and the corresponding row and column graphs. This tiling is irreducible and has diameter two.

Without loss of generality we consider only irreducible tilings. Reducible tilings can be broken down into several smaller components without disturbing the underlying states. To do this, both parties simply perform a projective measurement with respect to the subspaces corresponding to the different components of the row and column graphs.

Note that in general, a tiling is not invariant under local unitaries. In particular, the irreducibility of the tiling induced by a given set of states is a basis-dependent property.

The most extreme example of this phenomenon is the case of the standard basis. It induces a completely reducible tiling that consists only of 1×1 tiles. However, if both parties apply a generic local unitary transformation, the resulting tiling consists only of a single tile of maximal size.

4.3.2 Lower bounding the nonlocality constant using rigidity

In this section we assume that S is an orthonormal basis of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ (so in particular, $n = d_A d_B$) and discuss a particular strategy for lower bounding η for such S. We apply this strategy to several sets of orthonormal product bases in Section 4.4.

We bound η (quantifying a disturbance/information gain tradeoff) by considering a quantitative property of the set S called *rigidity*. Intuitively, we call a measurement operator strong if it is far from being proportional to the identity matrix; a set of states S is rigid if there exists a strong measurement that leaves the set undisturbed. We formalize this as follows (recall that $\|\cdot\|_{\max}$ denotes the largest entry of a matrix in absolute value):

Definition 4.8. We say that an orthonormal basis S is *c-rigid*, or c is an upper bound on the rigidity of S, if $c \in \mathbb{R}$ is such that

$$\left\| \frac{a \otimes b}{\operatorname{Tr}(a \otimes b)} - \frac{I}{n} \right\|_{\max} \leq c \cdot \delta_S(a \otimes b), \tag{4.29}$$

for all $a \in \operatorname{Pos}(\mathbb{C}^{d_A})$, $b \in \operatorname{Pos}(\mathbb{C}^{d_B})$ for which $\langle \psi_i | (a \otimes b) | \psi_i \rangle \neq 0$ for all i.

When S is rigid, the states can remain unchanged despite application of a strong measurement. For example, a tensor product basis is not c-rigid for any finite c (i.e., such a basis is arbitrarily rigid). In contrast, if c is small, then *any* strong measurement disturbs the set S, and Equation (4.29) quantifies how weak a measurement operator $a \otimes b$ must be for the disturbance $\delta_S(a \otimes b)$ to be small.

We now relate upper bounds on the rigidity of S to lower bounds on its nonlocality constant:

Lemma 4.9. *Let S be an orthonormal basis of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. If S is c-rigid then*

$$\eta \geq \frac{1}{cL}. \tag{4.30}$$

where L is the size of the largest tile corresponding to states in S.

Proof. If S is c-rigid, then for any $a \in \operatorname{Pos}(\mathbb{C}^{d_A})$ and $b \in \operatorname{Pos}(\mathbb{C}^{d_B})$ (such that $\langle \psi_k | (a \otimes b) | \psi_k \rangle \neq 0$ for all $k \in [n]$), we have

$$\frac{a \otimes b}{\operatorname{Tr}(a \otimes b)} - \frac{I}{n} = cM \cdot \delta_S(a \otimes b) \tag{4.31}$$

for some Hermitian matrix $M \in \mathrm{L}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ with $\|M\|_{\max} \leq 1$. From this we get

$$\max_{k \in [n]} \langle \psi_k | \frac{a \otimes b}{\operatorname{Tr}(a \otimes b)} | \psi_k \rangle - \frac{1}{n} = c \max_{k \in [n]} \langle \psi_k | M | \psi_k \rangle \cdot \delta_S(a \otimes b) \tag{4.32}$$

$$\leq cL \cdot \delta_S(a \otimes b). \tag{4.33}$$

By the definition of η (Equation (4.13)) and the fact that $\operatorname{Tr}(a \otimes b) = \sum_{j \in [n]} \langle \psi_j | (a \otimes b) | \psi_j \rangle$ for any orthonormal basis S, we get the desired inequality. □

Putting Lemma 4.9 and Theorem 4.7 together gives the following:

Theorem 4.10. *Let S be an orthonormal basis of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. If S is c-rigid then any LOCC measurement for discriminating states from S errs with probability*

$$p_{\text{error}} \geq \frac{2}{27} \frac{1}{(cL)^2 n^5} \tag{4.34}$$

where L is the size of the largest tile of S.

4.3.3 The "pair of tiles" lemma

In this section we present a lemma that serves as our main tool for bounding rigidity.

Lemma 4.11. *Let $U \in \mathrm{U}(m), V \in \mathrm{U}(n)$, and define $|\varphi_i\rangle := U|i\rangle$ for $i \in [m]$ and $|\psi_j\rangle := V|j\rangle$ for $j \in [n]$. Then for any $M \in \mathrm{L}(\mathbb{C}^n, \mathbb{C}^m)$ we have*

$$\sqrt{mn} \cdot \max_{i,j} |\langle \varphi_i | M | \psi_j \rangle| \geq \max_{k,l} |M_{kl}|. \tag{4.35}$$

The main idea of the proof is that a unitary change of basis can only increase the largest entry of a vector by a multiplicative factor depending on the dimension of the vector.

Proof. Let us define a mapping vec: $\mathrm{L}(\mathbb{C}^n, \mathbb{C}^m) \to \mathbb{C}^n \otimes \mathbb{C}^m$ as

$$\mathrm{vec}\colon |i\rangle\langle j| \mapsto |i\rangle|j\rangle \tag{4.36}$$

for $i \in [m]$ and $j \in [n]$ and extend it by linearity over $\mathbb{C}$. One can check that $\mathrm{vec}(AXB) = (A \otimes B^{\mathsf{T}})\,\mathrm{vec}(X)$. Using this and basic inequalities between the 2-norm and the ∞-norm, we get

$$\begin{aligned}
\max_{i,j} |\langle \varphi_i | M | \psi_j \rangle| &= \left\| \mathrm{vec}\Big(\sum_{i,j} \langle \varphi_i | M | \psi_j \rangle |i\rangle\langle j| \Big) \right\|_\infty && (4.37)\\
&= \left\| \mathrm{vec}\Big(\sum_{i,j} \langle i | U^\dagger M V | j \rangle |i\rangle\langle j| \Big) \right\|_\infty && (4.38)\\
&= \left\| \mathrm{vec}(U^\dagger M V) \right\|_\infty && (4.39)\\
&= \left\| (U^\dagger \otimes V^{\mathsf{T}})\, \mathrm{vec}(M) \right\|_\infty && (4.40)\\
&\geq \frac{1}{\sqrt{mn}} \left\| (U^\dagger \otimes V^{\mathsf{T}})\, \mathrm{vec}(M) \right\|_2 && (4.41)\\
&= \frac{1}{\sqrt{mn}} \left\| \mathrm{vec}(M) \right\|_2 && (4.42)\\
&\geq \frac{1}{\sqrt{mn}} \left\| \mathrm{vec}(M) \right\|_\infty && (4.43)\\
&= \frac{1}{\sqrt{mn}} \max_{k,l} |M_{kl}|, && (4.44)
\end{aligned}$$

as desired. □

Let us restate Lemma 4.11 using the language of tilings:

Lemma 4.12. *Let $R_1, R_2 \subseteq [d_A] \times [d_B]$ be two arbitrary regions of a $d_A \times d_B$ grid, and $\{|\varphi_i\rangle\}_{i=1}^{|R_1|}$ and $\{|\psi_j\rangle\}_{j=1}^{|R_2|} \subseteq \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ be their bases (here $|\varphi_i\rangle$ and $|\psi_j\rangle$ need not be product states). Then for any matrices $a \in \mathrm{L}(\mathbb{C}^{d_A})$ and $b \in \mathrm{L}(\mathbb{C}^{d_B})$ we have*

$$\sqrt{|R_1| \cdot |R_2|}\ \max_{i,j} |\langle \varphi_i | (a \otimes b) | \psi_j \rangle| \ \geq \max_{\substack{(r_1,c_1) \in R_1 \\ (r_2,c_2) \in R_2}} |a_{r_1 r_2}| \cdot |b_{c_1 c_2}|. \tag{4.45}$$

This follows from Lemma 4.11 by restricting $|\varphi_i\rangle$ and $|\psi_j\rangle$ to regions R_1 and R_2, respectively, and choosing M to be a submatrix of $a \otimes b$ with rows determined by R_1 and columns by R_2.

Proof. For $t \in \{1, 2\}$ let us enumerate the cells of region R_t by integers from $\{1, \ldots, |R_t|\}$ arbitrarily, and let $(r_t(i), c_t(i))$ be the coordinates of the ith cell of region R_t. Let

$$\Pi_t := \sum_{i=1}^{|R_t|} |i\rangle\langle r_t(i), c_t(i)| \tag{4.46}$$

be a linear operator that restricts the space $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ to region R_t. Then $|\varphi'_i\rangle := \Pi_1|\varphi_i\rangle$ is the restriction of $|\varphi_i\rangle$ to region R_1 and $|\psi'_i\rangle := \Pi_2|\psi_i\rangle$ is the restriction of $|\psi_i\rangle$ to R_2. Also, let $M := \Pi_1(a \otimes b)\Pi_2^\dagger$.

Note that for all $i \in \{1, \ldots, |R_1|\}$ we have $\Pi_1^\dagger\Pi_1|\varphi_i\rangle = |\varphi_i\rangle$ since the support of $|\varphi_i\rangle$ lies entirely within region R_1 and $\Pi_1^\dagger\Pi_1$ is the projection onto R_1. Similarly, $\Pi_2^\dagger\Pi_2|\psi_j\rangle = |\psi_j\rangle$ for all $j \in \{1, \ldots, |R_2|\}$. Hence

$$\langle\varphi_i|(a \otimes b)|\psi_j\rangle = \langle\varphi_i|\Pi_1^\dagger\Pi_1(a \otimes b)\Pi_2^\dagger\Pi_2|\psi_j\rangle = \langle\varphi'_i|M|\psi'_j\rangle \tag{4.47}$$

for all i and j. Finally, we apply Lemma 4.11 to $\{|\varphi'_i\rangle\}_{i=1}^{|R_1|}$, $\{|\psi'_j\rangle\}_{j=1}^{|R_2|}$, and M:

$$\begin{aligned}
\sqrt{|R_1| \cdot |R_2|}\ \max_{i,j} |\langle\varphi_i|(a \otimes b)|\psi_j\rangle| &= \sqrt{|R_1| \cdot |R_2|}\ \max_{i,j} |\langle\varphi'_i|M|\psi'_j\rangle| \\
&\geq \max_{k,l} |M_{kl}| \\
&= \max_{k,l} |\langle k|\Pi_1(a \otimes b)\Pi_2^\dagger|l\rangle| \\
&= \max_{k,l} \big|\langle r_1(k)|\, a\, |r_2(l)\rangle\big| \cdot \big|\langle c_1(k)|\, b\, |c_2(l)\rangle\big| \\
&= \max_{\substack{(r_1,c_1)\in R_1 \\ (r_2,c_2)\in R_2}} |a_{r_1 r_2}| \cdot |b_{c_1 c_2}|
\end{aligned}$$

and the result follows. □

When regions R_1 and R_2 are two distinct tiles from the tiling induced by S, we can use Lemma 4.12 to get the following result:

Lemma 4.13 ("Pair of tiles" Lemma)**.** *Let T_1 and T_2 be two distinct tiles in the tiling induced by S, and let $a \in \mathrm{Pos}(\mathbb{C}^{d_A})$ and $b \in \mathrm{Pos}(\mathbb{C}^{d_B})$. Then*

$$\sqrt{|T_1| \cdot |T_2|}\ \delta_S(a \otimes b)\,\mathrm{Tr}(a \otimes b) \geq |a_{r_1 r_2}| \cdot |b_{c_1 c_2}| \tag{4.48}$$

for any $r_t \in \mathrm{rows}(T_t)$ and $c_t \in \mathrm{cols}(T_t)$ where $t \in \{1, 2\}$.

Proof. We relax the inequality in Lemma 4.12 by observing that

$$\delta_S(a \otimes b) \geq \frac{\max_{i \neq j} |\langle \psi_i | (a \otimes b) | \psi_j \rangle|}{\|a \otimes b\|_\infty} \geq \frac{\max_{i \neq j} |\langle \psi_i | (a \otimes b) | \psi_j \rangle|}{\operatorname{Tr}(a \otimes b)} \tag{4.49}$$

which easily follows from the definition of $\delta_S(a \otimes b)$ in Equation (4.12). □

Note that the tiles T_1 and T_2 in Lemma 4.13 have to be distinct since the maximization in the definition of $\delta_S(a \otimes b)$ is performed only over pairs of distinct states. This lemma will be used later to bound the off-diagonal entries of $a \otimes b$ (see Figure 4.4).

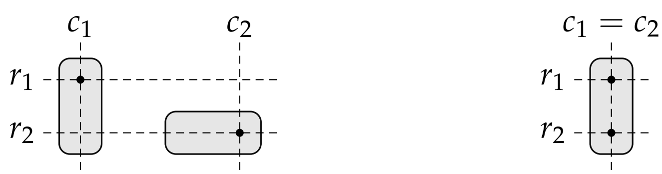


Figure 4.4: Whenever (r_1, c_1) and (r_2, c_2) belong to different tiles (left), Lemma 4.13 can be used to upper bound the off-diagonal entry $a_{r_1 r_2} \cdot b_{c_1 c_2}$ of $a \otimes b$. When both coordinates correspond to the same tile (right), this result cannot be applied directly.

4.4 Domino states

In this section we use the framework introduced earlier to give a lower bound on the error probability of any LOCC measurement for discriminating states from certain bipartite orthonormal product bases known as domino states. This provides an alternative proof of the quantitative separation between LOCC and separable measurements first given in [BDF+99] as well as generalizations to states corresponding to other domino-type tilings and a rotated version of the original domino states.

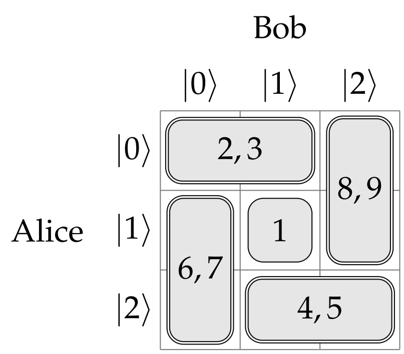


Figure 4.5: The tiling induced by the states from Equations (4.50)–(4.54).

4.4.1 Definition

The following orthonormal product basis is known as the *domino states*:

$$|\psi_1\rangle = |1\rangle|1\rangle, \tag{4.50}$$
$$|\psi_2\rangle = |0\rangle|0+1\rangle, \qquad |\psi_3\rangle = |0\rangle|0-1\rangle, \tag{4.51}$$
$$|\psi_4\rangle = |2\rangle|1+2\rangle, \qquad |\psi_5\rangle = |2\rangle|1-2\rangle, \tag{4.52}$$
$$|\psi_6\rangle = |1+2\rangle|0\rangle, \qquad |\psi_7\rangle = |1-2\rangle|0\rangle, \tag{4.53}$$
$$|\psi_8\rangle = |0+1\rangle|2\rangle, \qquad |\psi_9\rangle = |0-1\rangle|2\rangle, \tag{4.54}$$

where $|i \pm j\rangle := (|i\rangle \pm |j\rangle)/\sqrt{2}$. In [BDF+99] it was shown that any LOCC protocol for discriminating these states has information deficit at least 5.31×10^{-6} (out of $\log_2 9 \approx 3.17$) bits.

In [BDF+99] the authors also consider a family of orthonormal product bases, the so-called *rotated domino states*, which are parametrized by four angles $0 \leq \theta_1, \theta_2, \theta_3, \theta_4 \leq \pi/4$ and are defined as follows:

$$|\psi_1\rangle = |1\rangle|1\rangle, \tag{4.55}$$
$$|\psi_2\rangle = |0\rangle(\cos\theta_1|0\rangle + \sin\theta_1|1\rangle), \qquad |\psi_3\rangle = |0\rangle(-\sin\theta_1|0\rangle + \cos\theta_1|1\rangle), \tag{4.56}$$
$$|\psi_4\rangle = |2\rangle(\cos\theta_2|1\rangle + \sin\theta_2|2\rangle), \qquad |\psi_5\rangle = |2\rangle(-\sin\theta_2|1\rangle + \cos\theta_2|2\rangle), \tag{4.57}$$
$$|\psi_6\rangle = (\cos\theta_3|1\rangle + \sin\theta_3|2\rangle)|0\rangle, \qquad |\psi_7\rangle = (-\sin\theta_3|1\rangle + \cos\theta_3|2\rangle)|0\rangle, \tag{4.58}$$
$$|\psi_8\rangle = (\cos\theta_4|0\rangle + \sin\theta_4|1\rangle)|2\rangle, \qquad |\psi_9\rangle = (-\sin\theta_4|0\rangle + \cos\theta_4|1\rangle)|2\rangle. \tag{4.59}$$

Let $S_3(\theta_1, \theta_2, \theta_3, \theta_4)$ denote the rotated domino basis parametrized as above. Then the original domino basis is $S_3 := S_3(\pi/4, \pi/4, \pi/4, \pi/4)$.

Reference [BDF+99] shows that states from the domino basis S_3 cannot be perfectly discriminated by asymptotic LOCC and conjectures that the same holds for the rotated domino basis $S_3(\theta_1, \theta_2, \theta_3, \theta_4)$ for any $0 < \theta_1, \theta_2, \theta_3, \theta_4 \leq \pi/4$. In the next section we give an alternative proof that quantifies the nonlocality of the original domino states S_3 and then adapt the argument to the rotated domino states, thus resolving the conjecture.

4.4.2 Nonlocality of the domino states

To lower bound the nonlocality constant of the domino states S_3, we put an upper bound on their rigidity. In other words, we show that measurement operators that only slightly disturb these states are weak (approximately proportional to the identity operator). The key ingredient of the proof is Lemma 4.13 from Section 4.3.

Lemma 4.14. *The domino state basis S_3 is 4-rigid.*

Proof. The claimed result can be restated as follows (see Definition 4.8):

$$\left| a_{ii} b_{jj} - \frac{1}{9} \operatorname{Tr}(a \otimes b) \right| \leq 4\delta \operatorname{Tr}(a \otimes b), \tag{4.60}$$

$$|a_{ij} b_{kt}| \leq 4\delta \operatorname{Tr}(a \otimes b), \tag{4.61}$$

where $i, j, k, t \in \{0, 1, 2\}$ and $i \neq j$ or $k \neq t$ in the second equation. First we prove the bound for the diagonal elements and then we proceed to bound the off-diagonal ones.

Bounding the diagonal elements:

We start by bounding the differences of the diagonal elements of matrices a and b separately. Let us rewrite the definition of δ from Equation (4.12) in the case of product states $|\psi_i\rangle = |\alpha_i\rangle|\beta_i\rangle$:

$$\delta = \max_{i \neq j} \frac{|\langle\alpha_i|a|\alpha_j\rangle|}{\sqrt{\langle\alpha_i|a|\alpha_i\rangle\langle\alpha_j|a|\alpha_j\rangle}} \cdot \frac{|\langle\beta_i|b|\beta_j\rangle|}{\sqrt{\langle\beta_i|b|\beta_i\rangle\langle\beta_j|b|\beta_j\rangle}}. \tag{4.62}$$

If we consider the pair of states $|\psi_{2,3}\rangle = |0\rangle|0 \pm 1\rangle$, we get

$$
\begin{aligned}
\delta &\geq \frac{|a_{00}|}{|a_{00}|} \cdot \frac{|b_{00} - b_{01} + b_{10} - b_{11}|}{\sqrt{(b_{00} + b_{01} + b_{10} + b_{11})(b_{00} - b_{01} - b_{10} + b_{11})}} && (4.63)\\
&= \frac{|b_{00} - b_{11} + 2i \operatorname{Im} b_{10}|}{\sqrt{(b_{00} + b_{11})^2 - (b_{01} + b_{10})^2}} && (4.64)\\
&\geq \frac{|b_{00} - b_{11}|}{|b_{00} + b_{11}|} && (4.65)\\
&\geq \frac{|b_{00} - b_{11}|}{\operatorname{Tr}(b)}. && (4.66)
\end{aligned}
$$

Note that the cancellation of $|a_{00}|$ is valid since $a_{00} \neq 0$ by the definition of stage I. Applying a similar argument to the pairs of states from the other three tiles of size 2, we get that for any $i \in \{0, 2\}$,

$$
\delta \operatorname{Tr}(a) \geq |a_{11} - a_{ii}| \qquad \text{and} \qquad \delta \operatorname{Tr}(b) \geq |b_{11} - b_{ii}|. \tag{4.67}
$$

Using these bounds and the triangle inequality, we can bound the difference between the first and last diagonal elements:

$$
|a_{00} - a_{22}| \leq |a_{00} - a_{11}| + |a_{11} - a_{22}| \leq 2\delta \operatorname{Tr}(a) \tag{4.68}
$$

and similarly $|b_{00} - b_{22}| \leq 2\delta \operatorname{Tr}(b)$.

Next, we use the bounds on the differences of the diagonal elements of a and b to bound the differences of the diagonal elements of $a \otimes b$. For all $i, j, k, t \in \{0, 1, 2\}$ we have

$$
\begin{aligned}
|a_{ii} b_{jj} - a_{kk} b_{tt}| &\leq |a_{ii} b_{jj} - a_{kk} b_{jj}| + |a_{kk} b_{jj} - a_{kk} b_{tt}| && (4.69)\\
&= |b_{jj}| \cdot |a_{ii} - a_{kk}| + |a_{kk}| \cdot |b_{jj} - b_{tt}| && (4.70)\\
&\leq |b_{jj}| \cdot 2\delta \operatorname{Tr}(a) + |a_{kk}| \cdot 2\delta \operatorname{Tr}(b) && (4.71)\\
&\leq 4\delta \operatorname{Tr}(a \otimes b). && (4.72)
\end{aligned}
$$

Using this inequality we can obtain the desired bound (4.60) for the diagonal elements: for all $i, j \in \{0,1,2\}$ we have

$$\left| a_{ii}b_{jj} - \frac{1}{9}\operatorname{Tr}(a \otimes b) \right| = \left| a_{ii}b_{jj} - \frac{1}{9}\sum_{k,t\in\{0,1,2\}} a_{kk}b_{tt} \right| \tag{4.73}$$

$$\leq \frac{1}{9}\sum_{k,t\in\{0,1,2\}} |a_{ii}b_{jj} - a_{kk}b_{tt}| \tag{4.74}$$

$$\leq 4\delta\operatorname{Tr}(a \otimes b). \tag{4.75}$$

Bounding the off-diagonal elements:

From Lemma 4.13 we know that $\sqrt{|T_1|\cdot|T_2|}\,\delta\operatorname{Tr}(a\otimes b) \geq |a_{r_1r_2}|\cdot|b_{c_1c_2}|$, where T_1 and T_2 are two distinct tiles and $|T_t|$ is the area of the tile containing (r_t, c_t). For $(r_1, c_1) = (1,1)$ and any $(r_2, c_2) \neq (1,1)$ we get

$$\sqrt{2}\delta\operatorname{Tr}(a\otimes b) \geq |a_{r_1r_2}|\cdot|b_{c_1c_2}|. \tag{4.76}$$

Similarly, for any (r_1, c_1) and (r_2, c_2) that belong to distinct tiles of size two we get

$$2\delta\operatorname{Tr}(a\otimes b) \geq |a_{r_1r_2}|\cdot|b_{c_1c_2}|. \tag{4.77}$$

Now it only remains to bound the following four off-diagonal elements (each of which corresponds to one of the four tiles of size 2):

$$|a_{00}|\cdot|b_{01}|, \quad |a_{01}|\cdot|b_{22}|, \quad |a_{22}|\cdot|b_{12}|, \quad |a_{12}|\cdot|b_{00}|. \tag{4.78}$$

To bound $|a_{00}|\cdot|b_{01}|$, first choose $(r_2, c_2) = (1,0)$ and use Equation (4.76):

$$\sqrt{2}\delta\operatorname{Tr}(a\otimes b) \geq |a_{11}|\cdot|b_{10}| = |a_{11}|\cdot|b_{01}|. \tag{4.79}$$

Now it only remains to replace a_{11} by a_{00}. Notice from Equation (4.67) that $\delta\operatorname{Tr}(a) \geq |a_{11} - a_{00}| \geq |a_{00}| - |a_{11}|$, so

$$\sqrt{2}\delta\operatorname{Tr}(a\otimes b) \geq |a_{11}|\cdot|b_{01}| \geq \big(|a_{00}| - \delta\operatorname{Tr}(a)\big)\cdot|b_{01}| \tag{4.80}$$

$$\geq |a_{00}|\cdot|b_{01}| - \delta\operatorname{Tr}(a\otimes b) \tag{4.81}$$

where the last inequality holds since $|b_{01}| \leq \max\{b_{00}, b_{11}\} \leq \operatorname{Tr}(b)$ as b is positive semidefinite. After rearranging the previous expression we obtain

$$(1+\sqrt{2})\delta\operatorname{Tr}(a\otimes b) \geq |a_{00}|\cdot|b_{01}|. \tag{4.82}$$

By appropriately choosing the value of (r_2, c_2) and using a similar argument, we get the same upper bound for the remaining three off-diagonal elements listed in Equation (4.78). Since the constants obtained in bounds (4.76), (4.77), and (4.82) satisfy $\max\{\sqrt{2}, 2, 1+\sqrt{2}\} \leq 4$, we have shown that Equation (4.61) holds for all off-diagonal elements of $a \otimes b$. $\square$

Combining the above lemma with Equation (4.30) we obtain that the nonlocality constant for the domino states is $\eta \geq 1/8$. To get an explicit value for the lower bound on the error probability, we use Theorem 4.10 with $n = 9$, $L = 2$, and $c = 4$.

Corollary 4.15. *Any LOCC measurement for discriminating the domino states S_3 errs with probability*

$$p_{\text{error}} \geq 1.9 \times 10^{-8}. \tag{4.83}$$

4.4.3 Nonlocality of irreducible domino-type tilings

Lemma 4.14 can be easily generalized to product bases that are similar to domino states on larger quantum systems. The main ideas of the proof are essentially the same as in Lemma 4.14, but the argument must be adapted to accommodate tiles in arbitrary positions. A complete proof can be found in Appendix A.

Lemma 4.16. *Let $d_A, d_B \geq 3$ and let S be an orthonormal product basis of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. If S induces an irreducible domino-type tiling of diameter D then S is $2D$-rigid (see Section 4.3.1 for terminology).*

To bound the error probability, we use Theorem 4.10 with $n = d_A d_B$, $L = 2$, and $c = 2D$.

Corollary 4.17. *Any LOCC measurement for discriminating states from an orthonormal product basis of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ that induces an irreducible domino-type tiling of diameter D errs with probability*

$$p_{\text{error}} \geq \frac{1}{216 D^2 (d_A d_B)^5}. \tag{4.84}$$

4.4.4 Nonlocality of the rotated domino states

The following is an analog of Lemma 4.14 for rotated domino states.

Lemma 4.18. *The rotated domino basis $S_3(\theta_1, \theta_2, \theta_3, \theta_4)$ is $\frac{C}{\sin 2\theta}$-rigid where*

$$C := 6\Big(1 + 6\sqrt{2} + 2\sqrt{3(6+\sqrt{2})}\Big) \leq 114 \tag{4.85}$$

and $\theta := \min\{\theta_1, \theta_2, \theta_3, \theta_4\}$.

The proof appears in Appendix B.

Again, we use Theorem 4.10 to lower bound the error probability. Here the parameters are $n = 9$, $L = 2$, and $c = 114/\sin(2\theta)$.

Corollary 4.19. *Any LOCC measurement for discriminating $S_3(\theta_1, \theta_2, \theta_3, \theta_4)$, the set of rotated domino states, errs with probability*

$$p_{\text{error}} \geq 2.4 \times 10^{-11} \sin^2(2\theta), \tag{4.86}$$

where $\theta := \min\{\theta_1, \theta_2, \theta_3, \theta_4\}$.

Note that as θ approaches zero, the rigidity bound tends to infinity and the bound on the error probability goes to zero. As the original domino basis is transformed continuously to the standard basis, the nonlocality decreases to zero. Moreover, since any orthonormal product basis of $\mathbb{C}^3 \otimes \mathbb{C}^3$ is equivalent to $S_3(\theta_1, \theta_2, \theta_3, \theta_4)$ (up to local unitary transformations) for some angles θ_i [FS09], Corollary 4.19 effectively covers all product bases of $\mathbb{C}^3 \otimes \mathbb{C}^3$.

4.5 Limitations of the framework

4.5.1 Dependence of the nonlocality constant on n

Recall that in Theorem 4.7 we established the lower bound $p_{\text{error}} \geq \frac{2}{27}\frac{\eta^2}{n^5}$ on the error probability, where η is the nonlocality constant and n is the number of states. Intuitively it seems that it should be possible to prove a stronger lower bound on p_{error} as

n increases. However, to lower bound p_{error} by a fixed constant in *any* dimension using our framework, one would have to prove a lower bound on η that increases with n.

Let us consider the problem of discriminating orthonormal product states. In the next lemma we show that it is not possible to obtain such strong error bounds using our framework in its present form. We do this by proving a fixed upper bound on the nonlocality constant in any dimension.

Lemma 4.20. *Let S be a set of orthonormal product states in $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. The nonlocality constant of S satisfies $\eta \leq 2$.*

Proof. Let $n = |S|$ and $|\psi_i\rangle = |\alpha_i\rangle|\beta_i\rangle$. Fix some small $\epsilon > 0$, choose any $i \in [n]$, and define

$$a = |\alpha_i\rangle\langle\alpha_i| + \epsilon I_{d_A}, \qquad\qquad b = |\beta_i\rangle\langle\beta_i| + \epsilon I_{d_B}. \tag{4.87}$$

Note that a and b have full rank and are positive semidefinite. We can easily check that

$$\mathrm{Tr}(a) = 1 + \epsilon d_A, \qquad \mathrm{Tr}(b) = 1 + \epsilon d_B, \qquad \max_{k\in[n]} \langle\psi_k|(a\otimes b)|\psi_k\rangle = (1+\epsilon)^2. \tag{4.88}$$

Using these observations together with the definition of η in Equation (4.13), we get

$$\begin{aligned}
\eta\left(\frac{(1+\epsilon)^2}{(1+\epsilon d_A)(1+\epsilon d_B)} - \frac{1}{n}\right) &\leq \eta\left(\frac{\max_{k\in[n]}\langle\psi_k|(a\otimes b)|\psi_k\rangle}{\sum_{j\in[n]}\langle\psi_j|(a\otimes b)|\psi_j\rangle} - \frac{1}{n}\right) && (4.89)\\
&\leq \delta_S(a\otimes b) && (4.90)\\
&\leq 1, && (4.91)
\end{aligned}$$

where the last inequality follows directly from Definition 4.3. As $\epsilon \to 0$, the left-hand side goes to $\eta(1 - \frac{1}{n})$. We can choose ϵ arbitrarily small, so $\eta(1 - \frac{1}{n}) \leq 1$ and $\eta \leq \frac{n}{n-1} = 1 + \frac{1}{n-1}$. Since $n \geq 2$, we get $\eta \leq 2$. □

Note that from the above proof it follows that as $|S| \to \infty$ the upper bound on η tends to one.

4.5.2 Comparison to the result of Kleinmann, Kampermann, and Bruß

The main application of the framework introduced in this chapter is to show the impossibility of asymptotically discriminating a set of states S with LOCC. We do this by showing that the nonlocality constant of S is strictly positive. In other words, the nonlocality constant being zero is a necessary condition for the sates in S to be asymptotically distinguishable with LOCC. Another necessary condition is given by Theorem 3.15 from [KKB11] which we stated and discussed in Section 3.7.2.

One should note however that in contrast to the *qualitative* result of [KKB11], our framework can be applied to any set of orthogonal pure states (with no restriction on $\bigcap_i \ker \rho_i$) and can be used to obtain explicit lower bounds on the error probability. It is an open question whether our necessary condition ("the nonlocality constant of S is zero") or that of the above theorem is also sufficient. The lemma below shows that if our necessary condition is also sufficient then so is that of [KKB11].

Lemma 4.21. *Let $S = \{|\psi_i\rangle\}_{i\in[n]}$ be a set of orthogonal pure states such that $\bigcap_i \ker(|\psi_i\rangle\langle\psi_i|)$ does not contain any nonzero product vector. If for all χ with $1/n \leq \chi \leq 1$ there exists a positive semidefinite product operator E satisfying conditions (1)–(3) from Theorem 3.15, then the nonlocality constant η of S is zero.*

Proof. Consider $\chi \in \left(\frac{1}{n}, \frac{1}{n-1}\right)$ and a positive semidefinite product operator E_χ satisfying conditions (1)–(3). Conditions (1) and (2) imply that $\langle\psi_i|E_\chi|\psi_i\rangle > 0$ thus making $\delta_S(E_\chi)$ well defined (see Definition 4.3). Moreover, by condition (3) we have that $|\langle\psi_i|E_\chi|\psi_j\rangle|^2 = 0$ for all $i \neq j$. Hence $\delta_S(E_\chi) = 0$ according to Definition 4.3. Finally, from conditions (1) and (2), we get that

$$\frac{\max_i \langle\psi_i|E_\chi|\psi_i\rangle}{\sum_j \langle\psi_j|E_\chi|\psi_j\rangle} = \frac{\max_i \operatorname{Tr}(E\rho_i)}{\sum_j \operatorname{Tr}(E\rho_j)} = \chi. \tag{4.92}$$

Using these observations we can rewrite Equation (4.13) in the definition of η as

$$\eta\left(\chi - \frac{1}{n}\right) \leq 0. \tag{4.93}$$

Since $\chi > \frac{1}{n}$ it follows from the above inequality that $\eta = 0$. □

4.6 Discussion and open problems

We have developed a framework for quantifying the hardness of distinguishing sets of bipartite pure states with LOCC. Using this framework, we proved lower bounds on the error probability of distinguishing several sets of states, as summarized in Table 4.1.

Set of states	c	η	p_{error}
Domino states	4	$\frac{1}{8}$	1.96×10^{-8}
Domino-type states	$2D$	$\frac{1}{4D}$	$\frac{1}{216D^2(d_A d_B)^5}$
θ-rotated domino states	$\frac{114}{\sin 2\theta}$	$\frac{\sin 2\theta}{227}$	$2.41 \times 10^{-11} \sin^2(2\theta)$

Table 4.1: Rigidity c and lower bounds on the nonlocality constant η and error probability p_{error} for various states.

This work raises several open problems. While we were able to lower bound the nonlocality constant η in many cases, it could be useful to develop more generic approaches to computing or lower bounding this quantity. We are also interested in applying our method to other sets of states. For example, we would like to apply the method when S is an incomplete orthonormal set (*e.g.*, the domino basis with the middle tile omitted) or a product basis with tiles of size larger than two (see Figure 4.6 for concrete examples of such tilings where no bounds on p_{error} are known) or both (*e.g.*, the states called GenTiles in [DMS⁺03]). It is unknown whether there exists a set S of 2-qubit states that can be perfectly discriminated with separable operations, but for which any LOCC protocol has $p_{\text{error}}(S) > 0$ (see [DFXY09] for all possible candidate sets). Finally, it would be interesting to consider random product bases, since this would tell us how generic the phenomenon of nonlocality without entanglement is.

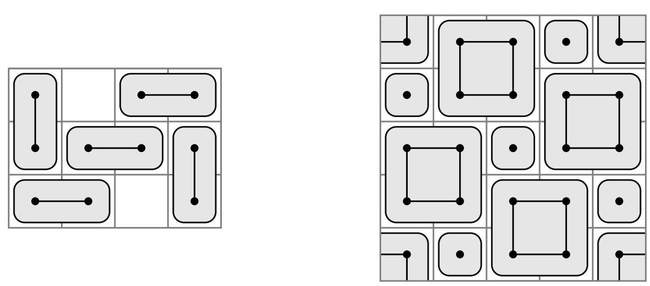


Figure 4.6: Tilings corresponding to an incomplete orthonormal set in $\mathbb{C}^3 \otimes \mathbb{C}^4$ (left) and a product basis of $\mathbb{C}^5 \otimes \mathbb{C}^5$ with larger tiles (right). On the right, the tiles of size four are induced by states of the form $|\pm\rangle|\pm\rangle$ and one of the tiles corresponds to the four corners of the grid.

We discussed some limitations of our framework in Section 4.5, but we would like to better understand how broadly the framework can be applied. In particular, can it always be used to obtain a lower bound on p_{error} whenever such a bound exists? For example, from Section 4.4.4 we know that the answer to this question is "yes" for orthonormal product bases on two qutrits.

Finally, the gaps between the classes of separable and LOCC operations exhibited by our framework are rather small (see Table 4.1). One cannot hope to do significantly better within our framework, as shown in Section 4.5.1. Is this due to limitations of our framework or because orthonormal product states in general can be discriminated well by LOCC?

Along these lines, a major open question raised by our work is the following: does there exist a sequence $S_1, S_2, S_3, \ldots$ of sets of orthonormal product states such that

$$\lim_{l \to \infty} p_{\text{error}}^{LOCC}(S_l) = 1?$$

Existence of such a sequence would give a strong separation between the classes of separable and LOCC measurements. Note that the local standard basis measurement followed by guessing gives the correct answer with probability at least $1/L_l$, where L_l is the maximum number of states within a tile in the tiling induced by S_l. Thus for any such sequence, the value of L_l must grow with l. In particular, the number of states (and hence the local dimensions) must also grow with l.

Chapter 5

Interpolability distinguishes LOCC from separable von Neumann measurements

5.1 Introduction

In this chapter we ask what determines whether a given separable operation can or cannot be implemented using LOCC. We show that, in the case of bipartite von Neumann measurements, the ability to interpolate measurements is an operational principle that sets apart LOCC from all separable operations.

5.1.1 Motivation

In contrast to the class LOCC, the class of separable operations, which is easily seen to encompass all LOCC operations, has a succinct and easy-to-use mathematical description. However, unlike LOCC, the class of separable operations does not have a natural operational interpretation.

Despite known quantitative separations between the two classes [BDF+99, CLMO12, KTYI07, Koa09, CCL12a, CCL12b, CH13], it is not understood what determines whether a given separable operation can or cannot be implemented with LOCC. Here, we draw our intuition from the proof in [BDF+99] and answer the above question for separable von Neumann measurements. In [BDF+99], the authors divided any LOCC measurement into two stages but did not distinguish between LOCC and separable measurements in any other way. Here, we show that the possibility to interpolate a measurement to obtain partial information is intrinsic to LOCC but not separable von Neumann measurements. More precisely, a separable von Neumann measurement can be interpolated only if it can be decomposed into two nontrivial steps, the first of which can be performed by a finite LOCC protocol. Therefore, the ability to interpolate is an operational principle that sets apart LOCC from separable von Neumann measurements.

5.1.2 Overview

This chapter is organized as follows. We start in Section 5.1.3 with a review of the concepts used in this chapter. In Section 5.2 we define interpolation, the central notion of this work, and provide a sufficient condition for interpolatability of arbitrary measurements (see Theorem 5.5). Our main result regarding interpolatability of separable and LOCC measurements is presented in Section 5.3 (see Theorem 5.10). We conclude in Section 5.4.

For simplicity, we present our results for the bipartite case. However, the same proof applies to an arbitrary number of parties.

5.1.3 Preliminaries

When discussing measurements, we use the terminology introduced in Section 1.3.3 and Section 3.2. We also employ the notation from Section 2.2.2 for the different classes of LOCC instruments ($\mathrm{LOCC}_{\mathbb{N}}$, $\overline{\mathrm{LOCC}}$ *etc.*).

We now give a definition of coarse graining specialized for the case of POVMs. This definition is in agreement with its more general counterpart from Chapter 2.

Definition 5.1 (Coarse graining)**.** Let $\mathcal{M} = \{E_1, \dots, E_k\}$ and $\mathcal{M}' = \{F_1, \dots, F_m\}$ be two measurements. We say that $\mathcal{M}'$ is a *coarse graining* of $\mathcal{M}$ if there exists a partition $(\Lambda_1, \dots, \Lambda_m)$ of $[k]$ such that $F_i = \sum_{j \in \Lambda_i} E_j$ for all $i \in [m]$.

Any LOCC protocol $\mathcal{P}$ implements a quantum operation of the form

$$\rho \mapsto \sum_{m \in \Gamma} |m\rangle\langle m| \otimes (A_m \otimes B_m)\rho(A_m^\dagger \otimes B_m^\dagger), \tag{5.1}$$

where Γ is the set of all terminating classical measurement records and $A_m \otimes B_m$ is the Kraus operator corresponding to record m (see Section 3.2.2 for more details). We refer to the operators $(A_m^\dagger A_m) \otimes (B_m^\dagger B_m)$ as the POVM elements of the protocol $\mathcal{P}$. We say that $\mathcal{P}$ implements a measurement $\mathcal{M} = \{E_i\}_i$ if the set Γ can be partitioned into parts Γ_i such that

$$E_i = \sum_{j \in \Gamma_i} (A_j^\dagger A_j) \otimes (B_j^\dagger B_j). \tag{5.2}$$

Operationally the partition corresponds to classical post-processing after the execution of $\mathcal{P}$ that coarse grains all outcomes in Γ_i for each i. When $\mathcal{M}$ is a von Neumann measurement, E_i in Equation (5.2) is rank one for each i. Thus $(A_j^\dagger A_j) \otimes (B_j^\dagger B_j)$ is proportional to E_i for all $j \in \Gamma_i$, and hence E_i is necessarily a tensor product operator. Therefore, if a von Neumann measurement in basis S can be implemented with LOCC then S consists only of tensor product vectors. We call such bases *product bases*.

Since non-orthogonal states can never be perfectly distinguished, to discriminate the states from an orthogonal set S with certainty one can only apply non-disturbing measurements. Recall from Definition 3.5 that we call a measurement $\mathcal{M}$ non-disturbing for an orthogonal set S, if $\langle\psi|E|\phi\rangle = 0$ for all $E \in \mathcal{M}$ and all distinct $|\psi\rangle, |\phi\rangle \in S$.

From Lemma 3.3 we see that the problem of implementing a measurement $\mathcal{M}$ in basis S with finite LOCC is equivalent to the problem of perfectly discriminating the states from S with finite LOCC. Throughout this chapter, we use the two perspectives interchangeably.

5.2 Interpolation of measurements

In this section we consider the problem of implementing a von Neumann measurement in two stages, *i.e.*, as a sequence of two measurements followed by coarse graining. In addition, we want to control how much progress is made during the first stage.

5.2.1 Progress function

To quantify the progress of the first measurement, we introduce a function that assigns numerical values to POVM elements. We take its range to be $[0, \infty)$. Each value indicates how much progress is made when a particular measurement outcome occurs; a larger value corresponds to more progress.

Any such progress function must satisfy some operationally-motivated properties. First, it must be continuous. Second, it must vanish on POVM elements that are non-informative (*i.e.*, proportional to the identity matrix). Third, as we want to measure the progress *conditioned* on having obtained a particular outcome, the progress function must be *scale-invariant* (*i.e.*, it remains the same when the POVM element is multiplied by a positive scalar). Fourth, since coarse graining corresponds to discarding classical information, the progress achieved by a coarse-grained operator $\sum_i E_i$ must not exceed that of the most informative E_i. We call the last condition *quasiconvexity*.

Definition 5.2 (Progress function)**.** A continuous function $\mu : \mathrm{Pos}(\mathbb{C}^n) \setminus \{0\} \to [0, \infty)$ such that $\mu(I) = 0$, $\mu(tE) = \mu(E)$ for all $t > 0$, and $\mu(E + F) \leq \max\{\mu(E), \mu(F)\}$ for all $E, F \in \mathrm{Pos}(\mathbb{C}^n) \setminus \{0\}$ is called a *progress function*.

5.2.2 Interpolation

We are interested in measurements $\mathcal{M}$ whose outcome statistics can be reproduced by a two-stage process: first perform some measurement $\mathcal{M}_1$ and then, conditioned on the outcome i, perform some other measurement $\mathcal{M}_2^{(i)}$. More formally:

Definition 5.3 (Composition of measurements)**.** Let $\mathcal{M}_1 = \{E_1, \ldots, E_k\}$ and $\mathcal{M}_2 = \{|i\rangle\langle i| \otimes E_j^{(i)}\}_{ij}$ be measurements such that $\mathcal{M}_2^{(i)} := \{E_j^{(i)}\}_j$ is also a measurement for each i. We say that a measurement $\mathcal{M}$ is a *composition* of $\mathcal{M}_1$ and $\mathcal{M}_2$, and write $\mathcal{M} = \mathcal{M}_2 \circ \mathcal{M}_1$, if $\mathcal{M}$ is a coarse graining (see Definition 5.1) of a POVM with elements

$$\{E_i^{\frac{1}{2}} E_j^{(i)} E_i^{\frac{1}{2}}\}_{ij}. \tag{5.3}$$

As a shorthand, we denote the second measurement $\mathcal{M}_2 = \bigoplus_{i \in [k]} \mathcal{M}_2^{(i)}$.

Note that due to coarse graining, the POVM elements in Equation (5.3) that correspond to some POVM element $E \in \mathcal{M}$ need not all be proportional to E. In such a case $\mathcal{M}_2 \circ \mathcal{M}_1$ does not reproduce the post-measurement state of $\mathcal{M}$ for the outcome corresponding to E. However, if $\mathcal{M}$ is a von Neumann measurement, then each $E \in \mathcal{M}$ is rank one, so the POVM elements in Equation (5.3) that correspond to E must be proportional to E. Therefore, any $\mathcal{M}_2 \circ \mathcal{M}_1$ that reproduces the measurement statistics of a von Neumann measurement $\mathcal{M}$ also reproduces its post-measurement states.

A progress function together with the ability to compose measurements allows us to speak of measurement interpolation, a two-stage implementation of a measurement where the amount of progress achieved in the first stage can be controlled. In general, some measurement outcomes might be more informative than others. In an ε-interpolation, the progress after the first measurement is at most ε regardless of the outcome obtained.

Definition 5.4 (ε-interpolation)**.** Let $\varepsilon \geq 0$. An *ε-interpolation* of a measurement $\mathcal{M}$ with respect to a progress function μ is a pair of measurements $\mathcal{M}_1 = \{E_1, \ldots, E_k\}$ and $\mathcal{M}_2$ such that

- $\max_i \mu(E_i) = \varepsilon$ and
- $\mathcal{M} = \mathcal{M}_2 \circ \mathcal{M}_1$ where $\mathcal{M}_2 = \bigoplus_{i \in [k]} \mathcal{M}_2^{(i)}$ for some measurements $\mathcal{M}_2^{(i)}$.

The following theorem from [KKB11] (whose idea originates in [BDF+99]) shows that any measurement can be ε-interpolated. Note that this theorem does not require the progress function to be quasiconvex.

Theorem 5.5 ([KKB11])**.** *Let μ be any progress function (see Definition 5.2). Then any measurement $\mathcal{M} = \{F_1, \ldots, F_k\} \subseteq \mathrm{Pos}(\mathbb{C}^n)$ can be ε-interpolated with respect to μ for any $\varepsilon \in [0, \lambda]$ where $\lambda := \max_i \mu(F_i)$.*

Proof. Recall from Section 3.7.1 that the pseudo-weak implementation of $\mathcal{M}$ with interpolation parameters $c_1, \ldots, c_k \geq 0$ is the measurement $\mathcal{M}_1 = \{E_1, \ldots, E_k\}$ where

$$E_i = c\,(c_i I + F_i) \tag{5.4}$$

and $c := (1 + \sum_i c_i)^{-1}$. Also recall that the corresponding recovery measurements $\mathcal{M}_2^{(i)} = \{E_1^{(i)}, \ldots, E_k^{(i)}\}$ for $i \in [k]$ are specified by

$$E_j^{(i)} := \begin{cases} \delta_{ij} I & \text{if } c_i = 0, \\ c\,(c_i + \delta_{ij})\, E_i^{-\frac{1}{2}} F_j E_i^{-\frac{1}{2}} & \text{otherwise.} \end{cases} \tag{5.5}$$

As explained in Section 3.7.1, performing $\mathcal{M}_1$ followed by an appropriate recovery measurement implements measurement $\mathcal{M}$. In other words, $\mathcal{M} = \mathcal{M}_2 \circ \mathcal{M}_1$, where $\mathcal{M}_2 = \bigoplus_{i \in [k]} \mathcal{M}_2^{(i)}$.

We now show that the measurements $\mathcal{M}_1$ and $\mathcal{M}_2^{(i)}$ satisfy the two conditions of ε-interpolation in Definition 5.4. First, note that E_i changes continuously from $\tilde{F}_i := (1 + \sum_{k \neq i} c_k)^{-1} F_i$ to I as c_i changes from 0 to ∞. Since the progress function μ is continuous on nonzero operators, the parameter c_i can be chosen so that $\mu(E_i)$ achieves any value between $\mu(\tilde{F}_i) = \mu(F_i)$ and $\mu(I) = 0$. Hence, for any $\varepsilon \in [0, \lambda]$ we can choose c_i so that $\mu(E_i) = \min\{\varepsilon, \mu(F_i)\}$. Since μ is scale-invariant, changing c_i does not affect the value of $\mu(E_j)$ for $j \neq i$. Hence, $\mu(E_i)$ can be adjusted independently for each i and the parameters c_i can be chosen so that $\max_i \mu(E_i) = \varepsilon$ for any $\varepsilon \in [0, \lambda]$. □

Theorem 5.5 states that any measurement can be ε-interpolated for small enough $\varepsilon > 0$ when the type of measurement in the interpolation is not restricted. In Theorem 5.10 we will see that this is not the case for interpolation with a restricted type of measurement.

5.2.3 Interpolation in SEP

When interpolating a separable measurement, we demand that the first-stage measurement $\mathcal{M}_1$ is separable, *i.e.*, $E = \sum_j a_j \otimes b_j$ for all $E \in \mathcal{M}_1$. Here the coarse graining over index j can be viewed as giving away the information about j to the environment. We wish to measure the achieved progress by taking into account all extracted classical information, even if it is held by the environment. Therefore, when interpolating within SEP we modify Definition 5.4.

Definition 5.6 (ε-interpolation in SEP)**.** Let $\mathcal{M} \subseteq \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ be a separable measurement. Given $\varepsilon \leq 0$, we say that $\mathcal{M}$ can be *ε-interpolated in* SEP if $\mathcal{M}$ admits an ε-interpolation $\mathcal{M}_2 \circ \mathcal{M}_1$ such that

- the measurement $\mathcal{M}_1 = \{E_1, \dots, E_k\}$ is separable, and
- $\max\limits_i \tilde{\mu}(E_i) = \varepsilon$, where $\tilde{\mu}$ is obtained by minimizing over all product decompositions:

$$\tilde{\mu}(E) := \min\Big\{ \max_j \mu(a_j \otimes b_j) : E = \sum_j a_j \otimes b_j \text{ where}$$
$$a_j \in \mathrm{Pos}(\mathbb{C}^{d_A}) \setminus \{0\}, b_j \in \mathrm{Pos}(\mathbb{C}^{d_B}) \setminus \{0\} \Big\}. \quad (5.6)$$

Note that the minimum in the definition of $\tilde{\mu}$ is always achieved: μ is scale invariant and we can use Carathéodory's theorem to bound the number of terms in the sum $\sum_j a_j \otimes b_j$.

Observe that Definition 5.4 is equivalent to Definition 5.6 without the constraint of product measurement operators. To see this, suppose we replace $a_j \otimes b_j$ with a general $F_j \in \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ in Equation (5.6). Then $E = F_1$ is a valid decomposition of E, so $\tilde{\mu}(E) \leq \mu(E)$. On the other hand, $\mu(E) \leq \tilde{\mu}(E)$ because $\mu(E) = \mu(\sum_j F_j) \leq \max_j \mu(F_j)$ by quasiconvexity of μ. Therefore, the requirement $\max_i \mu(E_i) = \varepsilon$ in Definition 5.4 is equivalent to $\max_i \tilde{\mu}(E_i) = \varepsilon$ in Definition 5.6 (without the product constraint).

If $\mathcal{M}$ can be ε-interpolated in SEP, then the first-stage measurement $\mathcal{M}_1$ can be chosen to have tensor product POVM elements. This is because any POVM element $E = \sum_j a_j \otimes b_j$ of the first-stage measurement can be replaced with its fine-grained product operators $a_j \otimes b_j$ achieving the minimum in the definition of $\tilde{\mu}$. This fact is crucial for the proof of Lemma 5.9.

Finally, observe that an ε-interpolation whose first-stage measurement $\mathcal{M}_1$ has only tensor product operators is an ε-interpolation in SEP. We call such interpolations *product interpolations*.

Lemma 5.7. *Let $\mathcal{M}_2 \circ \mathcal{M}_1$ be a product ε-interpolation of a separable measurement $\mathcal{M} \subseteq \mathrm{Pos}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$. Then $\mathcal{M}_2 \circ \mathcal{M}_1$ is an ε-interpolation of $\mathcal{M}$ in* SEP*.*

Proof. Let $\mathcal{M}_1 = \{c_1 \otimes d_1, \ldots, c_k \otimes d_k\}$. According to Definition 5.4, $\max_i \mu(c_i \otimes d_i) = \varepsilon$. To see that $\mathcal{M}_2 \circ \mathcal{M}_1$ is also an ε-interpolation of $\mathcal{M}$ in SEP, we need to show that $\max_i \tilde{\mu}(c_i \otimes d_i) = \varepsilon$. By the quasiconvexity of μ, for any $c \otimes d$ and any decomposition $\sum_j a_j \otimes b_j = c \otimes d$, we have $\mu(c \otimes d) \leq \max_j \mu(a_j \otimes b_j)$. Therefore $\mu(c_i \otimes d_i) = \tilde{\mu}(c_i \otimes d_i)$ and the lemma follows. $\square$

5.3 Interpolatability distinguishes LOCC from SEP

In this section we prove our main result concerning ε-interpolation of von Neumann measurements within SEP. Recall from Lemma 3.3 that performing a von Neumann measurement is equivalent to discriminating states from some orthonormal basis $S = \{|\psi_j\rangle\}_j$. Note that the POVM elements E for which $\langle\psi|E|\psi\rangle = 0$ for some $|\psi\rangle \in S$ eliminate $|\psi\rangle$. We focus on progress functions where the progress in such cases cannot be arbitrarily small, since intuitively such an outcome represents partial progress towards discriminating the states from S.

Definition 5.8 (Threshold)**.** A progress function μ (see Definition 5.2) has *threshold* $\mu_0 > 0$ with respect to an orthonormal basis S if for all nonzero $E \in \mathrm{Pos}(\mathbb{C}^n)$ for which $\langle\psi|E|\psi\rangle = 0$ for some $|\psi\rangle \in S$, we have $\mu(E) \geq \mu_0$.

As a concrete example, consider the following progress function from Chapter 4.

Example. Let $S \subseteq \mathbb{C}^n$ be an orthonormal basis. Consider $\mu : \mathrm{Pos}(\mathbb{C}^n) \to [0, \infty)$ given by

$$\mu(E) := \frac{\max_{|\psi\rangle \in S} \langle\psi|E|\psi\rangle}{\mathrm{Tr}(E)} - \frac{1}{|S|}. \tag{5.7}$$

The first term in Equation (5.7) is the maximum probability of making a correct guess if the outcome corresponds to E, so μ measures the deviation of the best guess from a uniformly random guess. It can be verified that μ satisfies the conditions from Definition 5.2 and hence is a valid progress function. If $\langle\psi|E|\psi\rangle = 0$ for some $|\psi\rangle \in S$, then the first term is at least $\frac{1}{n-1}$, so μ has threshold $\mu_0 = \frac{1}{n(n-1)}$.

The following lemma shows that if a separable von Neumann measurement $\mathcal{M}$ in basis S can be ε-interpolated in SEP for some small ε, then there exists a nontrivial *local* measurement that is non-disturbing for S (see Definition 3.5). Intuitively this means that some part of the measurement $\mathcal{M}$ can be implemented by LOCC (we formalize this intuition later in Theorem 5.10, our main result).

Lemma 5.9. *Let $\mathcal{M} \in$ SEP be a von Neumann measurement in basis $S \subseteq \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ and let μ be a progress function with threshold μ_0 with respect to S (see Definition 5.8). If $\mathcal{M}$ can be ε-interpolated in* SEP *for some $\varepsilon \in (0, \mu_0)$, then there exists a projective measurement $\mathcal{L}$ of the form $\mathcal{A} \otimes \mathcal{I}$ or $\mathcal{I} \otimes \mathcal{B}$ that is non-disturbing for S and achieves progress $\mu(E) \geq \mu_0$ for all $E \in \mathcal{L}$.*

Proof. Assume that $\mathcal{M}$ admits an ε-interpolation in SEP for some $\varepsilon \in (0, \mu_0)$. As explained above, the first-stage measurement $\mathcal{M}_1 = \{E_i\}_i \in$ SEP can be chosen to have tensor product POVM elements, *i.e.*, each $E_i = a_i \otimes b_i$. Since the measurement $\mathcal{M}$ perfectly discriminates the states from S, the first-stage measurement $\mathcal{M}_1$ in the implementation of $\mathcal{M}$ must be non-disturbing, *i.e.*,

$$\langle\psi_j|E_i|\psi_k\rangle = 0 \tag{5.8}$$

for all $E_i \in \mathcal{M}_1$ and all distinct $j, k \in [d_A d_B]$. It follows that each E_i is diagonal in the basis S. Thus, for each i and k there exists $\lambda_{ik} \geq 0$ such that

$$E_i|\psi_k\rangle = \lambda_{ik}|\psi_k\rangle. \tag{5.9}$$

If any $\lambda_{ik} = 0$, then $\langle\psi_k|E_i|\psi_k\rangle = 0$ and hence $\mu(E_i) \geq \mu_0$. Yet this contradicts the interpolation condition requiring that $\mu(E_i) \leq \varepsilon < \mu_0$. Thus $\lambda_{ik} > 0$ for all i, k.

Now, using the fact that each $E_i = a_i \otimes b_i$ and each $|\psi_k\rangle = |\alpha_k\rangle \otimes |\beta_k\rangle$ for a product basis S, we rewrite Equation (5.9) as

$$(a_i \otimes b_i)|\alpha_k\rangle \otimes |\beta_k\rangle = \lambda_{ik}|\alpha_k\rangle \otimes |\beta_k\rangle. \tag{5.10}$$

Strict positivity of λ_{ik} implies $b_i|\beta_k\rangle \neq 0$ and thus

$$a_i|\alpha_k\rangle = \eta_{ik}|\alpha_k\rangle, \tag{5.11}$$

where $\eta_{ik} = \lambda_{ik}/\||b_i|\beta_k\rangle\|_2 > 0$. Thus $(a_i \otimes I_B)|\psi_k\rangle = a_i|\alpha_k\rangle \otimes I_B|\beta_k\rangle = \eta_{ik}|\psi_k\rangle$, so

$$\langle\psi_j|(a_i \otimes I_B)|\psi_k\rangle = 0 \tag{5.12}$$

for all distinct j, k. Thus the matrix $a_i \otimes I_B$ is diagonal in basis S, and so is each $\Pi_{i,\eta} \otimes I_B$, where $\Pi_{i,\eta}$ is the projector onto the eigenspace of a_i with eigenvalue η. Hence

$$\langle\psi_j|(\Pi_{i,\eta} \otimes I_B)|\psi_k\rangle = 0 \tag{5.13}$$

for all distinct j, k and all $\eta \in \operatorname{spec}(a_i)$. If $\mathcal{A}_i := \{\Pi_{i,\eta} : \eta \in \operatorname{spec}(a_i)\}$ is the projective measurement onto the eigenspaces of a_i, then the joint measurement $\mathcal{A}_i \otimes \mathcal{I}$ is non-disturbing for S according to Equation (5.13). Note that unless $a_i = I_A$, we have $\langle\psi|(\Pi_{i,\eta} \otimes I_B)|\psi\rangle = 0$ for some $|\psi\rangle \in S$. In this case $\mu(E) \geq \mu_0$ for all $E \in \mathcal{A}_i \otimes \mathcal{I}$. The same holds for $\mathcal{I} \otimes \mathcal{B}_i$ which can be defined similarly.

It remains to show that for some i at least one of $\mathcal{A}_i \otimes \mathcal{I}$ and $\mathcal{I} \otimes \mathcal{B}_i$ is nontrivial. Consider an i such that $\mu(E_i) = \varepsilon$. Since $\varepsilon > 0$, $a_i \otimes b_i$ is not proportional to the identity matrix. Thus either a_i is not proportional to the identity matrix and hence $\mathcal{A}_i \otimes \mathcal{I}$ is nontrivial, or b_i is not proportional to the identity matrix and $\mathcal{I} \otimes \mathcal{B}_i$ is nontrivial. □

Now we are ready to prove our main theorem, establishing interpolatability as an operational principle that distinguishes LOCC and separable von Neumann measurements.

Theorem 5.10. *Let $\mathcal{M} \in$ SEP be a von Neumann measurement in basis $S \subseteq \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ and let μ be a progress function with threshold μ_0 with respect to S (see Definition 5.8). Then $\mathcal{M}$ can be ε-interpolated in* SEP *for some $\varepsilon \in (0, \mu_0)$ if and only if $\mathcal{M} = \mathcal{M}_2 \circ \mathcal{M}_1$ for some $\mathcal{M}_1 \in \mathrm{LOCC}_{\mathbb{N}}$ that achieves progress $\mu(E) \geq \mu_0$ for all $E \in \mathcal{M}_1$.*

Proof. ($\Rightarrow$) Assume that $\mathcal{M}$ can be ε-interpolated for some $\varepsilon \in (0, \mu_0)$. Then according to Lemma 5.9 there exists a local k-outcome measurement $\mathcal{A}$ on one of the parties, say Alice, such that $\mathcal{A} \otimes \mathcal{I}$ is non-disturbing for S and achieves progress $\mu(E) \geq \mu_0$ for all $E \in \mathcal{A} \otimes \mathcal{I}$. Let us choose $\mathcal{M}_1 = \mathcal{A} \otimes \mathcal{I}$ and $\mathcal{M}_2 = \bigoplus_{i\in[k]} \mathcal{M}$. Since $\mathcal{A} \otimes \mathcal{I}$ is non-disturbing for S, coarse-graining according to the outcomes of $\mathcal{M}_2$ implements the original measurement $\mathcal{M}$ in basis S. Hence $\mathcal{M} = \mathcal{M}_2 \circ \mathcal{M}_1$ where $\mathcal{M}_1 \in \mathrm{LOCC}_{\mathbb{N}}$ and $\mu(E) \geq \mu_0$ for all $E \in \mathcal{M}_1$.

($\Leftarrow$) Assume that $\mathcal{M} = \mathcal{M}_2 \circ \mathcal{M}_1$ for some $\mathcal{M}_1 \in \mathrm{LOCC}_{\mathbb{N}}$ that achieves progress $\mu(E) \geq \mu_0$ for all $E \in \mathcal{M}_1$. To obtain the desired ε-interpolation of $\mathcal{M}$, in an LOCC protocol for implementing $\mathcal{M}_1$ we find the earliest measurement with an outcome achieving nonzero progress. Interpolating this local measurement for small $\varepsilon > 0$, gives an ε-interpolation of $\mathcal{M}_1$ in SEP and hence also of $\mathcal{M}$. We now formalize this idea.

Consider an LOCC protocol for implementing $\mathcal{M}_1$. We can naturally represent this protocol as a rooted tree $\mathcal{T}$, where the nodes in each level correspond to measurements performed in the corresponding round of the protocol (see Section 3.2.4 for more explanation). We define a subtree $\mathcal{T}'$ of $\mathcal{T}$ recursively as follows (see Figure 5.1 for an example). First, we include the root of $\mathcal{T}$ in $\mathcal{T}'$. Next, if a vertex v is in $\mathcal{T}'$ and all children of v have zero progress, then we include the children of v in $\mathcal{T}'$ as well. We obtain the desired ε-interpolation of $\mathcal{M}$ by interpolating the measurement at some leaf v' of $\mathcal{T}'$.

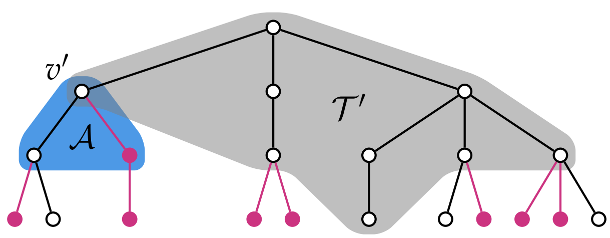


Figure 5.1: An example of a protocol tree $\mathcal{T}$ and its corresponding subtree $\mathcal{T}'$. The boundary around $\mathcal{T}'$ is marked by a dashed line. We use black edges and empty nodes to indicate that zero progress is made at that point of the protocol. Purple edges and solid nodes indicate nonzero progress. Since the marked node v' has a child with nonzero progress, we can ε-interpolate the local measurement at v' for some nonzero ε.

Since $\mu(E) > 0$ for all $E \in \mathcal{M}_1$ and μ is quasiconvex, μ must be nonzero at some leaf of $\mathcal{T}$. By construction, $\mathcal{T}'$ has some vertex v' that is a leaf in $\mathcal{T}'$ but has a child in $\mathcal{T}$ with nonzero progress. Consider the tree obtained from $\mathcal{T}'$ by adding all the children of v' and let $\mathcal{M}_1'$ be the corresponding measurement. Note that $\mathcal{M}_1$ can be expressed as $\mathcal{M}_1 = \mathcal{M}_1'' \circ \mathcal{M}_1'$, where $\mathcal{M}_1'' = \bigoplus_j \mathcal{M}_1^{(j)}$ and the $\mathcal{M}_1^{(j)}$ are some finite LOCC measurements.

Assume without loss of generality that Alice is the party performing a local measurement at v' and denote that measurement by $\mathcal{A}$. In analogy to Equation (5.3), define a function μ' on Alice's space via

$$\mu'(a) := \mu\big((\sqrt{a'}a\sqrt{a'}) \otimes b'\big), \tag{5.14}$$

where $a' \otimes b'$ is the POVM element that has been applied upon reaching node v'. Note that μ' is a valid progress function as it inherits all the properties required in Definition 5.2 from μ (e.g., $\mu'(I_A) = \mu(a' \otimes b') = 0$ by construction). Let $\lambda := \max_{a \in \mathcal{A}} \mu'(a)$ and note that $\lambda > 0$ according to our assumption that v' has children with nonzero progress.

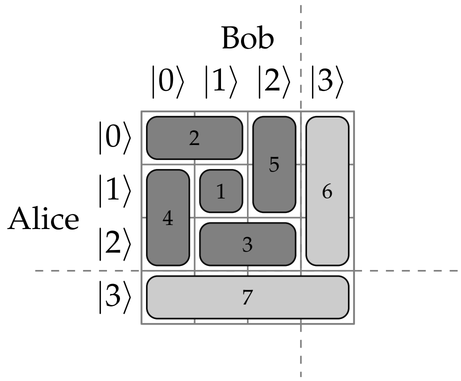
Now, using Theorem 5.5, we can ε-interpolate $\mathcal{A}$ with respect to μ' for any $\varepsilon \in (0, \min\{\lambda, \mu_0\}) \subseteq [0, \lambda]$. Any such ε-interpolation of $\mathcal{A}$ with respect to μ' gives a product ε-interpolation of $\mathcal{M}_1'$ with respect to μ. Since $\mathcal{M} = \mathcal{M}_2 \circ \mathcal{M}_1 = \mathcal{M}_2 \circ \mathcal{M}_1'' \circ \mathcal{M}_1'$, any product ε-interpolation of $\mathcal{M}_1'$ is also a product ε-interpolation of $\mathcal{M}$. Finally, applying Lemma 5.7 yields an ε-interpolation of $\mathcal{M}$ in SEP. □

To describe the consequences of Theorem 5.10, let us first consider an example.

Example. Let $\mathcal{M} \subseteq \mathrm{Pos}(\mathbb{C}^4 \otimes \mathbb{C}^4)$ be the von Neumann measurement corresponding to the product basis shown in Figure 5.2. Let $\mathcal{M}_{\mathrm{LOCC}} \in \mathrm{LOCC}_{\mathbb{N}}$ be a measurement implemented by the following two-step protocol (intuitively, it "peels off" the two extra tiles):

1. Alice performs a two-outcome measurement $\{I - |3\rangle\langle 3|, |3\rangle\langle 3|\}$ and sends the outcome to Bob.
2. If Alice got the first outcome, Bob applies the same measurement; otherwise he does nothing.

Note that $\mathcal{M}_{\mathrm{LOCC}}$ in this example is non-disturbing, so it can be completed by some measurement $\mathcal{M}'$ to obtain a decomposition $\mathcal{M} = \mathcal{M}' \circ \mathcal{M}_{\mathrm{LOCC}}$ as in Theorem 5.10. We can specify $\mathcal{M}'$ more precisely by describing the measurement associated to each outcome of $\mathcal{M}_{\mathrm{LOCC}}$. If either of the parties obtains $|3\rangle\langle 3|$, they are left with one of the two long tiles and the protocol can be easily completed by a local measurement in an appropriate basis. Otherwise they are left with the problem of discriminating the domino states. Then no nontrivial non-disturbing local measurement is possible [GV01, WH02, Coh07], so Alice and Bob cannot proceed any further by using only LOCC. We call the remaining measurement *purely separable* since it can be completed using only separable operations, but no further progress can be made by LOCC without ruining the orthogonality of the states.



$$
\begin{aligned}
|\psi_1\rangle &= |1\rangle|1\rangle \\
|\psi_2^{\pm}\rangle &= |0\rangle|0 \pm 1\rangle \\
|\psi_3^{\pm}\rangle &= |2\rangle|1 \pm 2\rangle \\
|\psi_4^{\pm}\rangle &= |1 \pm 2\rangle|0\rangle \\
|\psi_5^{\pm}\rangle &= |0 \pm 1\rangle|2\rangle \\
|\psi_6^{i}\rangle &= (U_3|i\rangle)|3\rangle \quad i \in \{0,1,2\} \\
|\psi_7^{j}\rangle &= |3\rangle(U_4|j\rangle) \quad j \in \{0,1,2,3\}
\end{aligned}
$$

Figure 5.2: A product basis corresponding to domino states (dark gray) augmented with two extra tiles (light gray). A tile of size l represents l states that are supported only on that tile (see Section 4.3.1 for more details). All 16 states are listed on the right, where $|x \pm y\rangle := (|x\rangle \pm |y\rangle)/\sqrt{2}$. The light gray tiles are generated by unitaries U_3 and U_4 of size 3×3 and 4×4, respectively, that have no zero entries in the computational basis. For concreteness, U_n could be the quantum Fourier transform modulo n.

5.4 Discussion

It is known that all LOCC measurements are separable but that some separable measurements are not even in the closure of LOCC [BDF+99, CLMO12]. Nevertheless, some separable measurements can be *partially* implemented by LOCC. Purely separable measurements are those that cannot even be partially implemented by LOCC. The resulting hierarchy is shown in Figure 5.3.

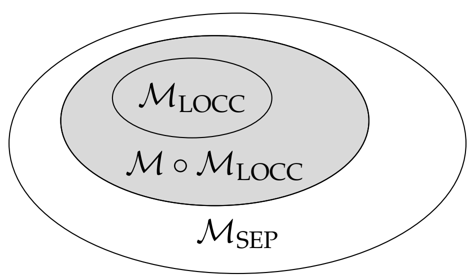


Figure 5.3: Subclasses of separable von Neumann measurements. The innermost region corresponds to LOCC measurements. The shaded region corresponds to measurements that can be partially implemented by LOCC, *i.e.*, decomposed as $\mathcal{M} \circ \mathcal{M}_{\text{LOCC}}$ as in Theorem 5.10. The white region corresponds to purely separable measurements.

Our main result (Theorem 5.10) characterizes purely separable measurements as precisely those for which ε-interpolation in SEP is not possible for any sufficiently small $\varepsilon > 0$. We conclude that ε-interpolatability in SEP for small $\varepsilon > 0$ is the key feature that distinguishes $\mathrm{LOCC}_{\mathbb{N}}$ from purely separable von Neumann measurements. In fact, this observation can be boosted to $\overline{\mathrm{LOCC}}$ (the closure of LOCC). This follows by combining Theorem 5.10 with Theorem 3.17.

Our results suggest several open problems. One possible research direction is to extend our results beyond von Neumann measurements. For example, can one generalize the notion of a progress function and prove an analogue of Theorem 5.10 for general POVMs or for the task of discriminating orthonormal states from an incomplete product basis?

Taking the idea of interpolation further, it might also be fruitful to find a continuous-time description of LOCC protocols. Such a description could give a new perspective on LOCC and a new tool for analyzing it. In particular, is it possible that the optimal protocol for some task is intrinsically continuous-time?

APPENDICES

Appendix A

Rigidity of domino-type states (Lemma 4.16)

Lemma 4.16. *Let $d_A, d_B \geq 3$ and let S be an orthonormal product basis of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. If S induces an irreducible domino-type tiling of diameter D then S is $2D$-rigid (see Section 4.3.1 for terminology).*

Proof. We mimic the proof of Lemma 4.14 and make the appropriate generalizations when necessary. We want to show that

$$\left| a_{ii} b_{jj} - \frac{1}{d_A d_B} \operatorname{Tr}(a \otimes b) \right| \leq 2D\delta \operatorname{Tr}(a \otimes b), \tag{A.1}$$

$$|a_{ij} b_{kt}| \leq 2D\delta \operatorname{Tr}(a \otimes b), \tag{A.2}$$

where $i \neq j$ or $k \neq t$ in the second inequality.

Bounding the diagonal elements:

Using the calculation in Equations (4.63–4.66) we can bound the difference of diagonal entries of a and b. Whenever there is a 2×1 tile that connects rows i and j, we get that

$$|a_{ii} - a_{jj}| \leq \delta \operatorname{Tr}(a). \tag{A.3}$$

A similar equation holds for b whenever there is a 1×2 tile that connects columns i and j.

Since T is irreducible, the row graph of T is connected. Moreover, any two vertices of this graph are connected by a path of length at most D. We apply the triangle inequality along this path in the same way as in Equation (4.68). After at most $D-1$ repetitions we get that for any i and j,

$$|a_{ii} - a_{jj}| \leq D\delta \operatorname{Tr}(a). \tag{A.4}$$

A similar equation holds for b. When we repeat the calculation in Equations (4.69–4.72), we get that for any i, j, k, t,

$$|a_{ii}b_{jj} - a_{kk}b_{tt}| \leq 2D\delta \operatorname{Tr}(a \otimes b). \tag{A.5}$$

Finally, we repeat the calculation in Equations (4.73–4.75) and get the desired bound stated in Equation (A.1).

Bounding the off-diagonal elements:

From Lemma 4.13 we get that

$$2\delta \operatorname{Tr}(a \otimes b) \geq |a_{r_1 r_2}| \cdot |b_{c_1 c_2}| \tag{A.6}$$

for all $(r_1, c_1) \neq (r_2, c_2)$, except when (r_1, c_1) and (r_2, c_2) belong to the same tile of size two.

Suppose that we want to bound $|a_{rr}| \cdot |b_{c_1 c_2}|$ where (r, c_1) and (r, c_2) belong to the same 1×2 tile. Since T is irreducible, we can find a row r' such that (r', c_1) and (r', c_2) belong to different tiles (if $\{r'\} \times \{c_1, c_2\}$ is a tile for each r' then $\{c_1, c_2\}$ is a connected component of the column graph of T, contradicting the irreducibility of T). From Equation (4.77) we get

$$2\delta \operatorname{Tr}(a \otimes b) \geq |a_{r'r'}| \cdot |b_{c_1 c_2}|. \tag{A.7}$$

According to Equation (A.4), $D\delta \operatorname{Tr}(a) \geq |a_{rr} - a_{r'r'}| \geq |a_{rr}| - |a_{r'r'}|$. Using this observation we repeat the calculation in Equations (4.80–4.81) and obtain

$$2\delta \operatorname{Tr}(a \otimes b) \geq |a_{rr}| \cdot |b_{c_1 c_2}| - D\delta \operatorname{Tr}(a \otimes b). \tag{A.8}$$

After rearranging terms we get

$$(D+2)\delta \operatorname{Tr}(a \otimes b) \geq |a_{rr}| \cdot |b_{c_1 c_2}|. \tag{A.9}$$

The same bound also holds for entries corresponding to 2×1 tiles. Together with Equation (A.6) this establishes the desired bound in Equation (A.2). □

Note that the above proof uses the assumption that the size of any tile is at most two in an essential way. Without this assumption, it might not be possible to find a row r' such that (r', c_1) and (r', c_2) belong to different tiles when bounding the off-diagonal element $|a_{rr}| \cdot |b_{c_1 c_2}|$.

Appendix B

Rigidity of rotated domino states (Lemma 4.18)

In this section we establish an analog of Lemma 4.14 for the rotated domino states $S_3(\theta_1, \theta_2, \theta_3, \theta_4)$. For simplicity we consider only the set $S_3(\theta) := S_3(\theta, \theta, \theta, \theta)$ and obtain a bound on the its rigidity in terms of θ. In the more general case one can choose $\theta := \min\{\theta_1, \theta_2, \theta_3, \theta_4\}$ and use the same bound.

Lemma B.1. *For $j \in \{0, 2\}$ we have*

$$|b_{11} - b_{jj}| \leq \frac{2}{\sin 2\theta}\big(\delta\|b\|_\infty + \big|\mathrm{Re}\, b_{j1}\big|\big). \tag{B.1}$$

The same inequality holds for a.

Proof. We show how to get the bound on b for $j = 0$. The remaining three cases are similar.

We use the states $|\psi_2\rangle$ and $|\psi_3\rangle$ from Equation (4.56) in the definition of δ in Equation (4.62):

$$\delta\|b\|_\infty \geq |\langle \beta_2|b|\beta_3\rangle| \tag{B.2}$$

$$= \left|\begin{pmatrix}\cos\theta & \sin\theta\end{pmatrix}\begin{pmatrix}b_{00} & b_{01}\\ b_{10} & b_{11}\end{pmatrix}\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}\right| \tag{B.3}$$

$$= \left|(b_{11}-b_{00})\sin\theta\cos\theta - b_{10}\sin^2\theta + b_{01}\cos^2\theta\right| \tag{B.4}$$

$$= \left|(b_{11}-b_{00})\sin\theta\cos\theta + \operatorname{Re} b_{01}(\cos^2\theta - \sin^2\theta) + i\operatorname{Im} b_{01}\right| \tag{B.5}$$

$$\geq |\frac{b_{11}-b_{00}}{2}\sin 2\theta + \operatorname{Re} b_{01}\cos 2\theta| \tag{B.6}$$

$$\geq \frac{|b_{11}-b_{00}|}{2}\sin 2\theta - |\operatorname{Re} b_{01}|. \tag{B.7}$$

By rearranging terms we get the desired bound. □

Lemma B.2. *If $a_{11} \geq \frac{1}{s}\|a\|_\infty$ for some $s > 0$ then for $j \in \{0,2\}$ we have*

$$|b_{j1}| \leq \sqrt{2}s\delta\|b\|_\infty, \qquad |b_{11}-b_{jj}| \leq 2(1+\sqrt{2}s)\frac{\delta}{\sin 2\theta}\|b\|_\infty. \tag{B.8}$$

The same statement holds when the roles of a and b are exchanged.

Proof. We show how to get bounds on b for $j=0$. The remaining three cases are identical, except one has to use states from different tiles.

We use Lemma 4.13 with tiles corresponding to states $|\psi_{6,7}\rangle$ and $|\psi_1\rangle$:

$$\sqrt{2}\delta\|a\otimes b\|_\infty \geq |a_{11}b_{01}| \geq \frac{1}{s}\|a\|_\infty|b_{01}|, \tag{B.9}$$

where the second inequality follows from our assumption $|a_{11}| \geq \frac{1}{s}\|a\|_\infty$. By rewriting this we get the first bound:

$$|b_{01}| \leq \sqrt{2}s\delta\|b\|_\infty. \tag{B.10}$$

Since $|\operatorname{Re} b_{01}| \leq |b_{01}| \leq \sqrt{2}s\delta\|b\|_\infty$, we get the second bound from Lemma B.1. □

Lemma B.3. *If $a_{11} \geq \frac{1}{s}\|a\|_\infty$ and $b_{11} \geq \frac{1}{s}\|b\|_\infty$ for some $s > 0$ then*

$$\|\frac{a\otimes b}{\operatorname{Tr}(a\otimes b)} - \frac{I}{9}\|_{\max} \leq 8(1+\sqrt{2}s)\frac{\delta}{\sin 2\theta}. \tag{B.11}$$

Proof. We follow the proof of Lemma 4.14 and show the following generalizations of Equations (4.60) and (4.61):

$$\left| a_{ii}b_{jj} - \frac{1}{9}\operatorname{Tr}(a \otimes b) \right| \leq 8(1+\sqrt{2}s)\frac{\delta}{\sin 2\theta}\|a \otimes b\|_\infty, \tag{B.12}$$

$$|a_{ij}b_{kt}| \leq \max\{\sqrt{2}, 2, \sqrt{2}s\}\delta\|a \otimes b\|_\infty. \tag{B.13}$$

Note that the second inequality is stronger than we need, since $1/\sin 2\theta \geq 1$.

First, we use Lemma B.2 to upper bound the difference of diagonal entries of a and b. We use these bounds in the same way as in Lemma 4.14 to upper bound the differences of diagonal entries of $a \otimes b$ and to get Equation (B.12). Finally, we use Lemma 4.13 to upper bound most of the off-diagonal entries of $a \otimes b$ and Lemma B.2 to upper bound the remaining ones. This gives us Equation (B.13).

Bounding the diagonal elements:

From Lemma B.2 we get bounds on $|b_{11} - b_{ii}|$ and $|a_{11} - a_{ii}|$ for $i \in \{0, 2\}$. Using the triangle inequality, we get

$$|a_{ii} - a_{jj}| \leq 4(1+\sqrt{2}s)\frac{\delta}{\sin 2\theta}\|a\|_\infty \tag{B.14}$$

for any $i, j \in \{0, 1, 2\}$ (and the same for b). Using the triangle inequality once more we can bound the difference of any two diagonal entries of $a \otimes b$:

$$|a_{ii}b_{jj} - a_{kk}b_{tt}| \leq 8(1+\sqrt{2}s)\frac{\delta}{\sin 2\theta}\|a \otimes b\|_\infty. \tag{B.15}$$

From this we obtain Equation (B.12) in the same way as in Lemma 4.14.

Bounding the off-diagonal elements:

Equation (B.13) can be obtained from Lemma 4.13. For most of the entries the constant is either $\sqrt{2}$ or 2, depending on the sizes of the tiles. For the remaining four entries, listed in Equation (4.78), we proceed in a slightly different way. For example, for $a_{00}b_{01}$ we use Equation (B.10) to see that

$$|a_{00}| \cdot |b_{01}| \leq \|a\|_\infty \cdot \sqrt{2}s\delta\|b\|_\infty. \tag{B.16}$$

A similar strategy works for the remaining three entries. □

Lemma B.4. *Fix any $s \geq 3$ and let*

$$\frac{1}{r(s)} := \min\{\frac{1}{14}\left(\frac{1}{3} - \frac{1}{s}\right), \frac{1}{2(1+\sqrt{2}s)}\left(\frac{1}{3} - \frac{1}{s}\right)\}. \tag{B.17}$$

If $\frac{\delta}{\sin 2\theta} \leq \frac{1}{r(s)}$ then $a_{11} \geq \frac{1}{s}\|a\|_\infty$ and $b_{11} \geq \frac{1}{s}\|b\|_\infty$.

Proof. We get one of the two lower bounds almost for free. We combine this with Lemma 4.13 and the triangle inequality to get the other lower bound.

If $\max_i a_{ii} = a_{11}$ then $a_{11} \geq \frac{1}{3}\operatorname{Tr}(a) \geq \frac{1}{3}\|a\|_\infty \geq \frac{1}{s}\|a\|_\infty$ and we are done with a. Similarly, if $\max_i b_{ii} = b_{11}$ then $b_{11} \geq \frac{1}{s}\|b\|_\infty$. Thus it only remains to consider the cases when $\max_i a_{ii} \in \{a_{00}, a_{22}\}$ and $\max_i b_{ii} \in \{b_{00}, b_{22}\}$. By symmetry, it suffices to consider the case where $\max_i a_{ii} = a_{22}$ and $\max_i b_{ii} = b_{00}$. The remaining three cases are similar.

Using the tiles that correspond to states $|\psi_{6,7}\rangle$ and $|\psi_{4,5}\rangle$, we get

$$2\delta\|a \otimes b\|_\infty \geq |a_{22}b_{01}| \geq \frac{1}{3}\|a\|_\infty |b_{01}|. \tag{B.18}$$

Thus $|\operatorname{Re} b_{01}| \leq |b_{01}| \leq 6\delta\|b\|_\infty$ and using Lemma B.1, we get

$$\begin{aligned} b_{00} - b_{11} &\leq |b_{11} - b_{00}| && \text{(B.19)}\\ &\leq \frac{2}{\sin 2\theta}(\delta\|b\|_\infty + |\operatorname{Re} b_{01}|) && \text{(B.20)}\\ &\leq 14\frac{\delta}{\sin 2\theta}\|b\|_\infty. && \text{(B.21)} \end{aligned}$$

We assumed that $\max_i b_{ii} = b_{00}$, so

$$\frac{1}{3}\|b\|_\infty \leq b_{00} \leq b_{11} + 14\frac{\delta}{\sin 2\theta}\|b\|_\infty. \tag{B.22}$$

By assumption, $\frac{\delta}{\sin 2\theta} \leq \frac{1}{r(s)} \leq \frac{1}{14}\left(\frac{1}{3} - \frac{1}{s}\right)$, so we get the desired bound $b_{11} \geq \frac{1}{s}\|b\|_\infty$.

As we have a lower bound on b_{11}, we can use Lemma B.2 and get

$$|a_{11} - a_{22}| \leq 2(1+\sqrt{2}s)\frac{\delta}{\sin 2\theta}\|a\|_\infty. \tag{B.23}$$

We assumed that $\max_i a_{ii} = a_{22}$, so we can rewrite this as

$$\frac{1}{3}\|a\|_\infty \leq a_{22} \leq a_{11} + 2(1+\sqrt{2}s)\frac{\delta}{\sin 2\theta}\|a\|_\infty. \tag{B.24}$$

By assumption, $\frac{\delta}{\sin 2\theta} \leq \frac{1}{r(s)} \leq \frac{1}{2(1+\sqrt{2}s)}\left(\frac{1}{3} - \frac{1}{s}\right)$, so we get the desired bound $a_{11} \geq \frac{1}{s}\|a\|_\infty$. $\square$

Lemma B.5. *For any fixed $s \geq 3$ we have the following:*

- *if* $\frac{\delta}{\sin 2\theta} \leq \frac{1}{r(s)}$ *then* $\|\frac{a\otimes b}{\mathrm{Tr}(a\otimes b)} - \frac{I}{9}\|_{\max} \leq 8(1+\sqrt{2}s)\frac{\delta}{\sin 2\theta}$,
- *if* $\frac{\delta}{\sin 2\theta} \geq \frac{1}{r(s)}$ *then* $\|\frac{a\otimes b}{\mathrm{Tr}(a\otimes b)} - \frac{I}{9}\|_{\max} \leq r(s)\frac{\delta}{\sin 2\theta}$,

where $r(s)$ is defined in Equation (B.17).

Proof. The first part follows by combining Lemmas B.3 and B.4. To obtain the second part, notice that all diagonal entries of $\frac{a\otimes b}{\mathrm{Tr}(a\otimes b)}$ are at most 1. Since this matrix is positive semidefinite, the off-diagonal entries are also at most 1, so the bound follows. $\square$

Lemma 4.18. *The rotated domino basis $S_3(\theta_1,\theta_2,\theta_3,\theta_4)$ is $\frac{C}{\sin 2\theta}$-rigid where*

$$C := 6\left(1+6\sqrt{2}+2\sqrt{3(6+\sqrt{2})}\right) \leq 114 \tag{4.85}$$

and $\theta := \min\{\theta_1,\theta_2,\theta_3,\theta_4\}$.

Proof. Let us denote the largest of the two constants in Lemma B.5 by

$$C(s) := \max\{8(1+\sqrt{2}s), r(s)\} \tag{B.25}$$

$$= \max\{8(1+\sqrt{2}s), 14\frac{3s}{s-3}, 2(1+\sqrt{2}s)\frac{3s}{s-3}\}, \tag{B.26}$$

where we substituted $r(s)$ from Equation (B.17). We want to make this constant as small as possible, so the best possible value is

$$C = \min_{s\geq 3} C(s) \tag{B.27}$$

$$= \min_{s\geq 3} 2(1+\sqrt{2}s)\frac{3s}{s-3} \tag{B.28}$$

$$= 6\left(1+6\sqrt{2}+2\sqrt{3(6+\sqrt{2})}\right), \tag{B.29}$$

where the minimum is reached at $s = 3+\sqrt{9+3/\sqrt{2}} \approx 6.33$. $\square$

References

[ABD+08] Scott Aaronson, Salman Beigi, Andrew Drucker, Bill Fefferman, and Peter Shor. The power of unentanglement. In *Proceedings of the 2008 IEEE 23rd Annual Conference on Computational Complexity*, CCC '08, pages 223–236, 2008. arXiv:0804.0802, doi:10.1109/CCC.2008.5. 3

[BaCY11] Fernando G.S.L. Brandão, Matthias Christandl, and Jon Yard. Faithful squashed entanglement. *Communications in Mathematical Physics*, 306(3):805–830, 2011. arXiv:1010.1750, doi:10.1007/s00220-011-1302-1. 3

[BBC+93] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70(13):1895–1899, Mar 1993. doi:10.1103/PhysRevLett.70.1895. 1, 23

[BBKW09] Somshubhro Bandyopadhyay, Gilles Brassard, Shelby Kimmel, and William K. Wootters. Entanglement cost of nonlocal measurements. *Phys. Rev. A*, 80:012313, Jul 2009. arXiv:0809.2264, doi:10.1103/PhysRevA.80.012313. 46

[BBPS96] Charles H. Bennett, Herbert Bernstein, Sandu Popescu, and Benjamin Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev. A*, 53(4):2046–2052, Apr 1996. arXiv:quant-ph/9511030, doi:10.1103/PhysRevA.53.2046. 2

[BCH+02] Dagmar Bruß, J. Ignacio Cirac, Pawel Horodecki, Florian Hulpke, Barbara Kraus, Maciej Lewenstein, and Anna Sanpera. Reflections upon separability and distillability. *J. Mod. Optics*, 49(8):1399–1418, 2002. arXiv:quant-ph/0110081, doi:10.1080/09500340110105975. 2

[BDF+99] Charles H. Bennett, David P. DiVincenzo, Christopher A. Fuchs, Tal Mor, Eric Rains, Peter W. Shor, John A. Smolin, and William K. Wootters. Quantum nonlocality without entanglement. *Phys. Rev. A*, 59:1070–1091, Feb 1999. arXiv:quant-ph/9804053, doi:10.1103/PhysRevA.59.1070. 2, 3, 4, 10, 22, 42, 44, 46, 62, 64, 65, 66, 67, 68, 82, 83, 84, 93, 94, 97, 104

[BDM+99] Charles H. Bennett, David P. DiVincenzo, Tal Mor, Peter W. Shor, John A. Smolin, and Barbara M. Terhal. Unextendible product bases and bound entanglement. *Phys. Rev. Lett.*, 82:5385–5388, Jun 1999. arXiv:quant-ph/9808030, doi:10.1103/PhysRevLett.82.5385. 22, 45, 65

[BDSW96] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54(5):3824–3851, Nov 1996. arXiv:quant-ph/9604024, doi:10.1103/PhysRevA.54.3824. 2, 22

[BN00] George Bachman and Lawrence Narici. *Functional Analysis*. Courier Dover Publications, 2000. URL: http://books.google.ca/books?id=wCHtLumoGY4C&pg=122. 13

[BPR+00] Charles H. Bennett, Sandu Popescu, Daniel Rohrlich, John A. Smolin, and Ashish V. Thapliyal. Exact and asymptotic measures of multipartite pure state entanglement. *Phys. Rev. A*, 63(1):012307, Dec 2000. arXiv:quant-ph/9908073, doi:10.1103/PhysRevA.63.012307. 2

[BT10] Aharon Brodutch and Daniel R. Terno. Quantum discord, local operations, and Maxwell's demons. *Phys. Rev. A*, 81:062103, Jun 2010. arXiv:1002.4913, doi:10.1103/PhysRevA.81.062103. 46

[CCL11] Wei Cui, Eric Chitambar, and Hoi-Kwong Lo. Randomly distilling W-class states into general configurations of two-party entanglement. *Phys. Rev. A*, 84(5):052301, Nov 2011. arXiv:1106.1209, doi:10.1103/PhysRevA.84.052301. 22

[CCL12a] Eric Chitambar, Wei Cui, and Hoi-Kwong Lo. Entanglement monotones for W-type states. *Phys. Rev. A*, 85(6):062316, Jun 2012. arXiv:1106.1208, doi:10.1103/PhysRevA.85.062316. 3, 9, 22, 46, 93

[CCL12b] Eric Chitambar, Wei Cui, and Hoi-Kwong Lo. Increasing entanglement monotones by separable operations. *Phys. Rev. Lett.*, 108(24):240504, Jun

2012. arXiv:1106.1208, doi:10.1103/PhysRevLett.108.240504. 3, 9, 22, 46, 93

[CD09] Eric Chitambar and Runyao Duan. Nonlocal entanglement transformations achievable by separable operations. *Phys. Rev. Lett.*, 103:110502, Sep 2009. arXiv:0811.3739, doi:10.1103/PhysRevLett.103.110502. 22, 35

[CDKL01] J. I. Cirac, W. Dür, B. Kraus, and M. Lewenstein. Entangling operations and their implementation using a small amount of entanglement. *Phys. Rev. Lett.*, 86(3):544–547, 2001. 18

[CH13] Eric Chitambar and Min-Hsiu Hsieh. A return to the optimal detection of quantum information. 2013. arXiv:1304.1555. 93

[Che04] Anthony Chefles. Condition for unambiguous state discrimination using local operations and classical communication. *Phys. Rev. A*, 69:050307, May 2004. doi:10.1103/PhysRevA.69.050307. 19, 65

[Chi11] Eric Chitambar. Local quantum transformations requiring infinite rounds of classical communication. *Phys. Rev. Lett.*, 107(19):190502, Nov 2011. arXiv:1105.3451, doi:10.1103/PhysRevLett.107.190502. 10, 16, 22

[Cho75] Man-Duen Choi. Completely positive linear maps on complex matrices. *Liner Alg. Appl.*, 10(3):285–290, Jun 1975. doi:10.1016/0024-3795(75)90075-0. 18

[CL03] Ping-Xing Chen and Cheng-Zu Li. Orthogonality and distinguishability: Criterion for local distinguishability of arbitrary orthogonal states. *Phys. Rev. A*, 68:062107, Dec 2003. arXiv:quant-ph/0209048, doi:10.1103/PhysRevA.68.062107. 65

[CL04] Ping-Xing Chen and Cheng-Zu Li. Distinguishing the elements of a full product basis set needs only projective measurements and classical communication. *Phys. Rev. A*, 70:022306, Aug 2004. doi:10.1103/PhysRevA.70.022306. 45, 65

[CLM+12] Eric Chitambar, Debbie Leung, Laura Mančinska, Maris Ozols, and Andreas Winter. Everything you always wanted to know about LOCC (but were afraid to ask). 2012. arXiv:1210.4583. 3, 4, 9, 10, 22

[CLMO12] Andrew M. Childs, Debbie Leung, Laura Mančinska, and Maris Ozols. A framework for bounding nonlocality of state discrimination. 2012. arXiv:1206.5822. 4, 16, 22, 64, 93, 104

[CLMO13] Andrew M. Childs, Debbie Leung, Laura Mančinska, and Maris Ozols. Interpolatability distinguishes LOCC from separable von Neumann measurements. 2013. arXiv:1306.5992. 4, 5

[Coh07] Scott M. Cohen. Local distinguishability with preservation of entanglement. *Phys. Rev. A*, 75:052313, May 2007. arXiv:quant-ph/0602026, doi:10.1103/PhysRevA.75.052313. 22, 65, 103

[Coh08] Scott M. Cohen. Understanding entanglement as resource: Locally distinguishing unextendible product bases. *Phys. Rev. A*, 77:012304, Jan 2008. arXiv:0708.2396, doi:10.1103/PhysRevA.77.012304. 46

[Coh11] Scott M. Cohen. When a quantum measurement can be implemented locally, and when it cannot. *Phys. Rev. A*, 84:052322, Nov 2011. arXiv:0912.1607, doi:10.1103/PhysRevA.84.052322. 46

[Cro12] Sarah Croke. There is no non-local information in a single qubit. In *APS Meeting Abstracts*, page 30011, Feb 2012. 44

[CY01] Yi-Xin Chen and Dong Yang. Optimal conclusive discrimination of two nonorthogonal pure product multipartite states through local operations. *Phys. Rev. A*, 64:064303, Nov 2001. doi:10.1103/PhysRevA.64.064303. 44, 65

[CY02] Yi-Xin Chen and Dong Yang. Optimally conclusive discrimination of nonorthogonal entangled states by local operations and classical communications. *Phys. Rev. A*, 65:022320, Jan 2002. doi:10.1103/PhysRevA.65.022320. 44, 65

[DFJY07] Runyao Duan, Yuan Feng, Zhengfeng Ji, and Mingsheng Ying. Distinguishing arbitrary multipartite basis unambiguously using local operations and classical communication. *Phys. Rev. Lett.*, 98:230502, Jun 2007. arXiv:quant-ph/0612034, doi:10.1103/PhysRevLett.98.230502. 65

[DFXY09] Runyao Duan, Yuan Feng, Yu Xin, and Mingsheng Ying. Distinguishability of quantum states by separable operations. *IEEE Trans. Inf. Theor.*, 55:1320–1330, March 2009. arXiv:0705.0795, doi:10.1109/TIT.2008.2011524. 22, 45, 46, 65, 91

[DHR02] Matthew J. Donald, Michał Horodecki, and Oliver Rudolph. The uniqueness theorem for entanglement measures. *J. Math. Phys.*, 43:4252–4272, 2002. arXiv:quant-ph/0105017, doi:10.1063/1.1495917. 10

[DL70] E. Brian Davies and John T. Lewis. An operational approach to quantum probability. *Comm. Math. Phys.*, 17(3):239–260, 1970. URL: http://projecteuclid.org/euclid.cmp/1103842336, doi:10.1007/BF01647093. 11

[DLT02] David P. DiVincenzo, Debbie W. Leung, and Barbara M. Terhal. Quantum data hiding. *IEEE Trans. Inf. Theory*, 48(3):580–598, Mar 2002. arXiv:quant-ph/0103098, doi:10.1109/18.985948. 3, 18

[DMS+03] David P. DiVincenzo, Tal Mor, Peter W. Shor, John A. Smolin, and Barbara M. Terhal. Unextendible product bases, uncompletable product bases and bound entanglement. *Communications in Mathematical Physics*, 238:379–410, 2003. arXiv:quant-ph/9908070, doi:10.1007/s00220-003-0877-6. 45, 65, 91

[DR04] Sergio De Rinaldis. Distinguishability of complete and unextendible product bases. *Phys. Rev. A*, 70:022309, Aug 2004. arXiv:quant-ph/0304027, doi:10.1103/PhysRevA.70.022309. 45, 46, 50

[DVC00] Wolfgang Dür, Guifre Vidal, and J. Ignacio Cirac. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A*, 62(6):062314, Nov 2000. arXiv:quant-ph/0005115, doi:10.1103/PhysRevA.62.062314. 23

[DXY10] Runyao Duan, Yu Xin, and Mingsheng Ying. Locally indistinguishable subspaces spanned by three-qubit unextendible product bases. *Phys. Rev. A*, 81:032329, Mar 2010. arXiv:0708.3559, doi:10.1103/PhysRevA.81.032329. 45, 65

[EW02] Tilo Eggeling and Reinhard F. Werner. Hiding classical data in multipartite quantum states. *Phys. Rev. Lett.*, 89(9):097905, Aug 2002. arXiv:quant-ph/0203004, doi:10.1103/PhysRevLett.89.097905. 3

[Fan04] Heng Fan. Distinguishability and indistinguishability by local operations and classical communication. *Phys. Rev. Lett.*, 92:177905, Apr 2004. arXiv:quant-ph/0311026, doi:10.1103/PhysRevLett.92.177905. 45, 65

[FS09] Yuan Feng and Yaoyun Shi. Characterizing locally indistinguishable orthogonal product states. *IEEE Trans. Inf. Theor.*, 55:2799–2806, June 2009. arXiv:0707.3581, doi:10.1109/TIT.2009.2018330. 45, 65, 88

[GB03] Leonid Gurvits and Howard Barnum. Separable balls around the maximally mixed multipartite quantum states. *Phys. Rev. A*, 68(4):042312, Oct 2003. arXiv:quant-ph/0302102, doi:10.1103/PhysRevA.68.042312. 26

[GKR+01] Sibasish Ghosh, Guruprasad Kar, Anirban Roy, Aditi Sen(De), and Ujjwal Sen. Distinguishability of Bell states. *Phys. Rev. Lett.*, 87:277902, Dec 2001. arXiv:quant-ph/0106148, doi:10.1103/PhysRevLett.87.277902. 45, 65

[GKR08] Shafi Goldwasser, YaelTauman Kalai, and GuyN. Rothblum. One-time programs. In *Advances in Cryptology — CRYPTO 2008*, volume 5157 of *Lecture Notes in Computer Science*, pages 39–56. Springer Berlin Heidelberg, 2008. doi:10.1007/978-3-540-85174-5_3. 3

[GKRS04] Sibasish Ghosh, Guruprasad Kar, Anirban Roy, and Debasis Sarkar. Distinguishability of maximally entangled states. *Phys. Rev. A*, 70:022304, Aug 2004. arXiv:quant-ph/0205105, doi:10.1103/PhysRevA.70.022304. 65

[GL03] Daniel Gottesman and Hoi-Kwong Lo. Proof of security of quantum key distribution with two-way classical communications. *IEEE Trans. Inf. Theory*, 49(2):457–475, Feb 2003. arXiv:quant-ph/0105121, doi:10.1109/TIT.2002.807289. 22

[GV01] Berry Groisman and Lev Vaidman. Nonlocal variables with product-state eigenstates. *Journal of Physics A: Mathematical and General*, 34(35):6881, 2001. arXiv:quant-ph/0103084, doi:10.1088/0305-4470/34/35/313. 42, 45, 65, 103

[Hel76] Carl W. Helstrom. *Quantum Detection and Estimation Theory*. Mathematics in science and engineering. Academic Press, 1976. URL: http://books.google.ca/books?id=Ne3iT_QLcsMC&pg=PA113. 74

[HHH96] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Separability of mixed states: necessary and sufficient conditions. *Phys. Lett. A*, 223(1–2):1–8, 1996. arXiv:quant-ph/9605038, doi:10.1016/S0375-9601(96)00706-2. 2

[HHH98] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Mixed-state entanglement and distillation: is there a "bound" entanglement in nature? *Phys. Rev. Lett.*, 80(24):5239–5242, Jun 1998. arXiv:quant-ph/9801069, doi:10.1103/PhysRevLett.80.5239. 17, 19

[HHH00] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Limits for entanglement measures. *Phys. Rev. Lett.*, 84(9):2014–2017, Feb 2000. arXiv:quant-ph/9908065, doi:10.1103/PhysRevLett.84.2014. 2

[HHHH09] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009. arXiv:0702225, doi:10.1103/RevModPhys.81.865. 2

[HM03] Mark Hillery and Jihane Mimih. Distinguishing two-qubit states using local measurements and restricted classical communication. *Phys. Rev. A*, 67:042304, Apr 2003. arXiv:quant-ph/0210179, doi:10.1103/PhysRevA.67.042304. 65

[HMM+06] Masahito Hayashi, Damian Markham, Mio Murao, Masaki Owari, and Shashank Virmani. Bounds on multipartite entangled orthogonal state discrimination using local operations and classical communication. *Phys. Rev. Lett.*, 96:040501, Feb 2006. arXiv:quant-ph/0506170, doi:10.1103/PhysRevLett.96.040501. 45

[Hor97] Paweł Horodecki. Separability criterion and inseparable mixed states with positive partial transposition. *Phys. Lett. A*, 232(5):333–339, 1997. arXiv:quant-ph/9703004, doi:10.1016/S0375-9601(97)00416-7. 22

[HSSH03] Michał Horodecki, Aditi Sen(De), Ujjwal Sen, and Karol Horodecki. Local indistinguishability: More nonlocality with less entanglement. *Phys. Rev. Lett.*, 90:047902, Jan 2003. arXiv:quant-ph/0301106, doi:10.1103/PhysRevLett.90.047902. 45, 65

[Jam72] Andrzej Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Rep. Math. Phys.*, 3(4):275–278, Dec 1972. doi:10.1016/0034-4877(72)90011-0. 18

[JCY05] Zhengfeng Ji, Hongen Cao, and Mingsheng Ying. Optimal conclusive discrimination of two states can be achieved locally. *Phys. Rev. A*, 71:032323, Mar 2005. arXiv:quant-ph/0407120, doi:10.1103/PhysRevA.71.032323. 44, 65

[KKB11] Matthias Kleinmann, Hermann Kampermann, and Dagmar Bruß. Asymptotically perfect discrimination in the local-operation-and-classical-communication paradigm. *Phys. Rev. A*, 84:042326, Oct 2011. arXiv:1105.5132, doi:10.1103/PhysRevA.84.042326. 4, 10, 16, 38, 46, 53, 56, 68, 90, 97

[KMY03] Hirotada Kobayashi, Keiji Matsumoto, and Tomoyuki Yamakami. Quantum Merlin-Arthur proof systems: Are multiple Merlins more helpful to Arthur? In *Algorithms and Computation*, volume 2906 of *Lecture Notes in Computer Science*, pages 189–198. Springer Berlin Heidelberg, 2003. arXiv:0306051, doi:10.1007/978-3-540-24587-2_21. 3

[Koa09] Masato Koashi. On the irreversibility of measurements of correlations. *Journal of Physics: Conference Series*, 143(1):012007, 2009. doi:10.1088/1742-6596/143/1/012007. 46, 65, 93

[KSV02] Alexei Yu. Kitaev, Alexander Shen, and Mikhail N. Vyalyi. *Classical and Quantum Computation*, volume 47 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002. URL: http://books.google.com/books?id=qYHTvHPvmG8C. 13

[KTYI07] Masato Koashi, Fumitaka Takenaga, Takashi Yamamoto, and Nobuyuki Imoto. Quantum nonlocality without entanglement in a pair of qubits. 2007. arXiv:0709.3196. 22, 46, 65, 93

[Liu13] Yi-Kai Liu. Building one-time memories from isolated qubits. 2013. arXiv:1304.5007. 3

[MP95] Serge Massar and Sandu Popescu. Optimal extraction of information from finite quantum ensembles. *Phys. Rev. Lett.*, 74(8):1259–1263, Feb 1995. doi:10.1103/PhysRevLett.74.1259. 2

[Nat05] Michael Nathanson. Distinguishing bipartitite orthogonal states using LOCC: Best and worst cases. *Journal of Mathematical Physics*, 46(6):062103, 2005. arXiv:quant-ph/0411110, doi:10.1063/1.1914731. 45, 65

[NC06] Julien Niset and Nicolas J. Cerf. Multipartite nonlocality without entanglement in many dimensions. *Phys. Rev. A*, 74:052103, Nov 2006. arXiv:quant-ph/0606227, doi:10.1103/PhysRevA.74.052103. 45, 65

[Nie99] Michael A. Nielsen. Conditions for a class of entanglement transformations. *Phys. Rev. Lett.*, 83(2):436–439, Jul 1999. arXiv:quant-ph/9811053, doi:10.1103/PhysRevLett.83.436. 2

[OH08] Masaki Owari and Masahito Hayashi. Two-way classical communication remarkably improves local distinguishability. *New J. Phys.*, 10(1):013006, 2008. arXiv:0708.3154, doi:10.1088/1367-2630/10/1/013006. 22

[PV07] Martin B. Plenio and Shashank Virmani. An introduction to entanglement measures. *Quant. Inf. Comput.*, 7(1&2):1–51, Jan 2007. URL: http://www.rintonpress.com/xqic7/qic-7-12/001-051.pdf, arXiv:quant-ph/0504163. 2

[PW91] Asher Peres and William K. Wootters. Optimal detection of quantum information. *Phys. Rev. Lett.*, 66(9):1119–1122, Mar 1991. doi:10.1103/PhysRevLett.66.1119. 2

[Rai97] Eric M. Rains. Entanglement purification via separable superoperators, 1997. arXiv:quant-ph/9707002. 18, 19

[Rai99] E. M. Rains. Rigorous treatment of distillable entanglement. *Phys. Rev. A*, 60(1):173–178, Jul 1999. arXiv:quant-ph/9809078, doi:10.1103/PhysRevA.60.173. 18

[Rai01] Eric M. Rains. A semidefinite program for distillable entanglement. *IEEE Trans. Inf. Theory*, 47(7):2921–2933, Nov 2001. arXiv:quant-ph/0008047, doi:10.1109/18.959270. 19

[Rin04] Sergio De Rinaldis. Distinguishability of complete and unextendible product bases. *Phys. Rev. A*, 70(2):022309, Aug 2004. arXiv:quant-ph/0304027, doi:10.1103/PhysRevA.70.022309. 10

[Roc96] R. Tyrell Rockafellar. *Convex Analysis*. Princeton Mathematical Series. Princeton University Press, 1996. URL: http://books.google.com/books?id=1TiOka9bx3sC. 19

[TDL01] Barbara M. Terhal, David P. DiVincenzo, and Debbie W. Leung. Hiding bits in Bell states. *Phys. Rev. Lett.*, 86(25):5807–5810, Jun 2001. arXiv:quant-ph/0011042, doi:10.1103/PhysRevLett.86.5807. 3, 19

[VPRK97] Vlatko Vedral, Martin B. Plenio, Michael A. Rippin, and Peter L. Knight. Quantifying entanglement. *Phys. Rev. Lett.*, 78(12):2275–2279, Mar 1997. arXiv:quant-ph/9702027, doi:10.1103/PhysRevLett.78.2275. 2

[VSPM01] Shashank Virmani, Massimiliano F. Sacchi, Martin B. Plenio, and Damian Markham. Optimal local discrimination of two multipartite pure states. *Physics Letters A*, 288(2):62–68, 2001. arXiv:quant-ph/0102073, doi:10.1016/S0375-9601(01)00484-4. 44, 65

[Wat05a] John Watrous. Bipartite subspaces having no bases distinguishable by local operations and classical communication. *Phys. Rev. Lett.*, 95:080505, Aug 2005. arXiv:quant-ph/0411077, doi:10.1103/PhysRevLett.95.080505. 65

[Wat05b] John Watrous. Notes on super-operator norms induced by Schatten norms. *Quant. Inf. Comput.*, 5(1):58–68, Jan 2005. URL: http://www.rintonpress.com/xqic5/qic-5-1/058-068.pdf, arXiv:quant-ph/0411077. 13

[Wat12] John Watrous. Quantum computational complexity. In *Computational Complexity*, pages 2361–2387. Springer New York, 2012. arXiv:0804.3401, doi:10.1007/978-1-4614-1800-9_147. 3

[WBC11] Christopher Wood, Jacob Biamonte, and David Corey. A return to the optimal detection of quantum information. 2011. arXiv:1111.6950. 7

[Wer89] Reinhard F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40(8):4277–4281, Oct 1989. doi:10.1103/PhysRevA.40.4277. 2

[WH02] Jonathan Walgate and Lucien Hardy. Nonlocality, asymmetry, and distinguishing bipartite states. *Phys. Rev. Lett.*, 89:147901, Sep 2002. arXiv:quant-ph/0202034, doi:10.1103/PhysRevLett.89.147901. 42, 45, 65, 103

[WSHV00] Jonathan Walgate, Anthony J. Short, Lucien Hardy, and Vlatko Vedral. Local distinguishability of multipartite orthogonal quantum states. *Phys. Rev. Lett.*, 85:4972–4975, Dec 2000. arXiv:quant-ph/0007098, doi:10.1103/PhysRevLett.85.4972. 44, 65

[XD08] Yu Xin and Runyao Duan. Local distinguishability of orthogonal $2 \otimes 3$ pure states. *Phys. Rev. A*, 77(1):012315, Jan 2008. arXiv:0709.1651, doi:10.1103/PhysRevA.77.012315. 22