Tensors and the Entanglement of Pure Quantum States

Asger Kjærulff Jensen

PhD Thesis

Department of Mathematical Sciences
University of Copenhagen

30-April-2019

Asger Kjærulff Jensen
Department of Mathematical Sciences
University of Copenhagen
Universitetesparken 5
2100 København Ø
ISBN: 978-87-7078-890-8

This thesis has been submitted to the PhD School of the Faculty of Science, University of Copenhagen, Denmark in April 2019.

Advisor:

Matthias Christandl
Professor
University of Copenhagen

Assessment committee:

Mikael Rørdam	Alexander Holevo	Peter Harremoes
Professor	Professor	Lektor
University of Copenhagen	Steklov Mathematical Institute	Niels Brock

Abstract

Entanglement constitutes one of the important resources in quantum information theory. Accordingly, characterizing the entanglement content of a given quantum state is an important concept in quantum information. A natural approach to quantifying entanglement is to consider how an entangled quantum state can be transformed, by transformations that cannot generate entanglement in a system. A central class of such operations are so-called Local Operations and Classical Communication (LOCC).

This thesis deals with asymptotic conversion rates for pure quantum states under exact LOCC transformations, in particular by expressing such rates through various kinds of entanglement monotones.

Firstly, the hierarchy of multipartite states under stochastic LOCC (SLOCC) is studied via its equivalence to restrictions of tensors. The tensor rank is of particular interest, as it describes the cost of creating the associated state by Greenberger–Horne–Zeilinger (GHZ) states. The fact that the GHZ-cost is non-linear is equivalent to the non-multiplicativity of tensor rank under the Kronecker product. In order to better understand how and why this non-linearity occurs, we consider whether strict sub-multiplicativity of tensor rank stems entirely from the joining of tensor legs, or if it can happen without joining tensor legs. It is shown that strict sub-multiplicativity happens for both tensor rank and border rank when just taking the tensor product.

A tool for working with asymptotic tensor rank is that of an asymptotic spectrum of a preordered semiring. This theory reduces the question of asymptotic restrict-ability to majorization on the set of order preserving homomorphisms into the reals. This concept will be introduced and an example will be computed in the tripartite case, for the sub-semiring generated by the W and GHZ state together with an Einstein–Podolsky–Rosen (EPR) pair shared between two fixed parties.

As the ultimate goal should be characterizing multipartite entanglement through non-stochastic LOCC, the asymptotic spectrum method is applied to a refinement of the tensor semiring. This refinement keeps some control on the probability of successful outcomes of LOCC protocols, specifically yielding a set of monotones, describing asymptotic conversion rates given any converse error exponent, r. While this is still some distance from the ideal asymptotic regime, we see that in the bipartite case, in fact the conversion rate for success probability going to 1 is also captured by these monotones.

Finally, a formula for the pure, bipartite, exact, deterministic conversion rate is presented, as derived through type class arguments.

Resumé

Kvantesammenfiltring udgør en vigtig resource i kvanteinformationsteori. Derfor er beskrivelsen af kvantesammenfiltringen for en kvantetilstand en vigtig opgave indenfor feltet. En naturlig tilgang til denne opgave er at undersøge hvordan sammenfiltrede kvantetilstande transformeres via operationer der ikke kan generere sammenfiltring. En central klasse af sådanne transformationer er såkaldte Lokale Operationer og Klassisk Kommunikation (LOCC).

Denne afhandling omhandler asymptotiske konverteringsrater for rene kvantetilstande under LOCC transformationer, især ved at udtrykke sådanne rater ved forskellige former for sammenfiltringsmonotoner.

I første omgang studeres hierakiet af mangedelte kvantetilstande under stochastisk LOCC (SLOCC) via dets relation til restriktion af tensorer. Tensorrang er af særlig interesse, da rangen beskriver omkostningen ved produktion af den givne kvantetilstand målt i antal af Greenberger–Horne–Zeilinger (GHZ) tilstande. At GHZ-omkostning er ikke-lineær er ækvivalent med ikke-multiplikativitet af tensorrang under Kronecker produktet. For bedre at forstå hvordan denne ikke-linearitet opstår, stilles der spørgsmål ved hvorvidt streng submultiplikativitet af tensorrang er forårsaget udelukkende ved sammensætningen af tensordele, eller om det kan forekomme ved det almindelige tensor produkt. Det vises at både tensorrang og "borderrang" kan være strengt submultiplikativ under det almindelige tensorprodukt.

Et værktøj ved arbejde med asymptotisk tensorrang er det asymptotiske spektrum for en præordnet semiring. Teorien reducerer spørgsmål om asymptotisk restriktionsbarhed til evaluering af ordensbevarende homomorfier ind i de reelle tal. Det asymptotiske spektrum introduceres og et eksempel bliver udregnet for del-semiringen genereret af de tredelte tilstande W og GHZ samt et Einstein–Podolsky–Rosen (EPR) par delt mellem to bestemte lokationer.

Da karakterisering af mangedelt sammenfiltring under ikke-stokastik LOCC bør være det endelige mål, anvendes den ovennævnte metode på en forfining af semiringen af tensorer under restriktion. Denne forfining bibeholder en hvis kontrol med sandsynligheden for succesfuldt udfald af LOCC protokoller, hvilket fører til monotoner, der beskriver asymptotiske konverteringsrater, givet enhver omvendt fejleksponent, r. Selvom dette stadig er et stykke fra det ideelle asymptotiske regime, ser vi at konverteringsraten for konverteringssandsynlighed gående mod 1, i det todelte tilfælde, er beskrevet præcis ved disse monotoner.

Endelig udledes en formel for konverteringsraten under eksakt, deterministisk konvertering af rene todelte kvantetilstande.

Acknowledgements

Firstly, I would like to thank my advisor, Matthias Christandl, for introducing me to tensor rank, which has been an exciting field to work with, and for always having good ideas for topics and questions to work on.

I thank my officemates over the years, Anton Samojlow, Sabiha Tokus, Birger Brietzke, Christian Majenz, Marie Fialová, Nicholas Gauguin Houghton-Larsen and Maxime Savoy for enduring my mess and keeping a positive atmosphere in the office.

I would also like to thank Fulvio Gesmundo, Jeroen Zuiddam, Christopher David Perry, Angelo Lucia, Daniel Stilck França, Alexander Müller-Hermes and Albert Werner for taking part in stimulating, weekly tensor meetings. An extra thank you to Albert Werner for feedback on my thesis.

I am grateful to Milán Mosonyi for making me feel welcome during my stay at BME.

Finally and especially, I thank Péter Vrana for stimulating discussions, helpful advice, critique on my thesis, for sharing his insights, answering questions and for listening to and understanding my mathematical intuitions and ideas.

I acknowledge financial support from the European Research Council (ERC Grant Agreement no. 337603) and VILLUM FONDEN via the QMATH Centre of Excellence (Grant no. 10059).

Contents

Notation

- $[d]$: For $d \in \mathbb{N}$, $[d] = \{1, \ldots, d\}$
- $\mathcal{S}(\mathcal{H})$: Density operators on Hilbert space $\mathcal{H}$, i.e. positive semidefinite operators with trace 1.
- K^*: Given a matrix K, K^* is the conjugate transpose of the matrix. In much quantum information literature, this is denoted $K^\dagger$.
- $\odot$: The flattened tensor product. See Section 1.1.1
- $\mathrm{supp}P$: Given a probability distribution P on a finite set $\mathcal{X}$. The support is the set $\mathrm{supp}P = \{x \in \mathcal{X} | P(x) \neq 0\}$.
- $\mathcal{X}^n$: sequences of length n with values in $\mathcal{X}$. In other words, maps $I : [n] \to \mathcal{X}$.
- $\mathbb{C}^{\mathcal{X}}$: The vector space of functions (families) $v : \mathcal{X} \to \mathbb{C}$. As a special case we have:
- $\mathbb{C}^d$: The vector space of functions (families) $v : [d] \to \mathbb{C}$ or $v : \{0, \ldots d-1\} \to \mathbb{C}$ depending on indexing.
- log: log will always be taken to mean $\log_2$, as is standard in much of quantum information theory.
- $|\psi_P\rangle$: Given a probability distribution $P : I \to [0, 1]$ on some finite set I, $|\psi_P\rangle = \sum_{i \in I} \sqrt{P(i)}\, |ii\rangle$ is the bipartite pure state with Schmidt coefficients $\sqrt{P}$.
- u_r: The unit tensor $|u_r\rangle = \sum_{i=1}^{r} |i \ldots i\rangle$. The order of u_r is inferred from context or specified with a superscript u_r^k.
- $|\mathrm{GHZ}_r\rangle$: The r-level, k-partite GHZ-state $|\mathrm{GHZ}_r\rangle = \frac{1}{\sqrt{r}}\, |u_r\rangle$. Again, the order is inferred from context or written with superscript.
- $P^{\otimes n}$: Given probability distribution $P : I \to [0, 1]$, $P^{\otimes n}(x) = \prod_{i=1}^{n} P(x_i)$ for $x \in I^n$.
- $n \gg 1$: For n sufficiently large.

Chapter 1

Introduction

A major concept setting quantum reality apart from the classical is the notion of entanglement. The fact that physical systems can be correlated in ways that exceed shared randomness is an important tool in many quantum computing protocols which exceed the best known classical protocols. But characterizing, quantifying and even defining entanglement in a multipartite scenario is tricky business. We can easily make sense of what it means for quantum systems to be entangled; they are correlated in a way that cannot be described as simply shared randomness. But if we then ask how entangled the systems are, the answer is unclear. Since entanglement between systems is only created when the systems come into contact with each other, we are comfortable in saying that any local action performed on an entangled system will always produce a less entangled state, at least on average. For this reason we quantify entanglement through so-called entanglement monotones; functions that assign a positive number to each possible state, in a way that respects whichever operations we consider to be entanglement-reducing.

Entanglement monotone is a notion from the resource theory of entanglement, arguably the most widely studied quantum resource theory. For a recent review of quantum resource theories in general, see [1]. In general, a resource theory consists of two things; resources and allowed operations. The resources of any quantum resource theory are quantum states and the allowed operations are generally some subset of completely positive maps between state spaces. Monotones in a resource theory are then maps that assign real numbers to states in a way that is monotone under allowed operations. Since the goal is to characterize entanglement, the allowed operations will be precisely the ones which can produce only classical correlations between separated systems. That is, local quantum operations and classical communication (LOCC). Given a resource theory, the mathematical tasks can generally be split into two categories. 1: Quantify the resource by determining all monotones with certain properties (e.g. additivity for entanglement monotones). 2: Characterize which conversions are possible, for instance by determining how much of a resource is needed to obtain another, either in a single-shot way or in some asymptotic way. Sometimes the answers to questions in the second category will be expressed in terms of some subset of monotones, and so we may sometimes reduce questions about convertibility to questions of determining monotones. This thesis is heavily focused on this kind of reduction.

The most famous entanglement montone is the entropy of entanglement, which to a bipartite pure state, $\rho_{AB} = |\psi\rangle\langle\psi|_{AB}$, associates the von Neumann entropy of either marginal, $-\operatorname{Tr}\rho_A \log \rho_A$. This monotone is by itself a good measure of pure bipartite entanglement as it characterizes the asymptotic conversion rate between any pair of bipartite pure states up to any arbitrarily small fixed error in terms of fidelity [2]. However, for bipartite mixed states or multipartite pure states, there is no single measure that perfectly quantifies entanglement. For bipartite states, the asymptotic entanglement cost of a state, ρ, is the number of maximally entangled qubit pairs needed per copy of ρ to create many copies of ρ. The asymptotic distillable entanglement is, conversely, the number of maximally entangled pairs that can be created per copy of ρ. For bipartite mixed states, there is a gap between asymptotic entanglement cost and asymptotic distillable entanglement, showing that asymptotic entanglement transformation is irreversible. An example of mixed states with such a gap are bound entangled states [3], which are entangled and therefore have positive entanglement cost, but with zero distillable entanglement. The multipartite case ($k \geq 3$) is also very difficult, even for pure states and has been studied extensively in recent decades (see e.g. [4, 5, 6, 7, 8, 9, 10] to mention a few). There are multiple reasons for considering entanglement transformations with fidelity loss. For one, the knowledge of a real physical state is only ever up to some approximation, so describing a state up to some error is physically realistic. But accepting some error can also make the mathematically difficult task of determining convertibility of entanglement easier, as witnessed by the case of bipartite pure states and the entropy of entanglement.

A different kind of relaxation which one can consider is that of probabilistic conversion. Rather than allowing for output states within some proximity of the target state, we might allow for only some chance of successful conversion. This leads to a different resource theory of entanglement, where the set of permissible operations increase. A benefit of this resource theory of entanglement under stochastic LOCC (SLOCC) is the simplicity of describing the channels, as they are merely represented by the tensor product of linear maps (Proposition 1.1.16). This thesis will be dealing entirely with asymptotic, exact LOCC conversions and mostly in a probabilistic regime. By exact, we mean that the output state of an (S)LOCC protocol must yield the target state exactly, with no loss in fidelity. It is the belief of this author that there is a relation between the LOCC spectrum $\Delta(\mathcal{S}_k)$ of Chapter 4 and the set of additive multipartite entanglement monotones describing asymptotic bound fidelity-loss entanglement transformations.

The overarching theme of this thesis is the transformation of many copies of a resource into many copies of some other resource of the same kind. The resources will be pure quantum states (Chapter 4), tensors (Chapter 2 and Chapter 3) and probability distributions (Chapter 5) and the set of allowed operations will vary, depending on probabilistic regime. The initial motivation is resource theories of quantum entanglement. By Proposition 1.1.16, the resource theory of pure multipartite quantum states under SLOCC is equivalent to the resource theory of tensors under restriction (see Definition 2.1.3), motivating the study of restrictions of tensors from a quantum information standpoint.

The only way the many-copy setting of a resource theory can differ from the single-copy setting is if the amount of [phenomenon] in a collection of resources is somehow different from the sum of the amounts of the members of the collection. For instance, in the entanglement resource theory of multipartite states, one might consider the GHZ state $\frac{1}{\sqrt{2}}(|0\dots0\rangle + |1\dots1\rangle)$ as the gold-standard of multipartite entanglement. It seems reasonable to quantify the amount of entanglement of a pure state, $|\psi\rangle$, as either the number of GHZ states needed to produce $|\psi\rangle$ (GHZ-cost) with allowed operations (e.g. LOCC or SLOCC) or as the number of GHZ states which can be produced from $|\psi\rangle$ (GHZ-distillation). Generally, the collection of multiple quantum states is described by taking the flattened tensor product (or Kronecker product, see Section 1.1.1) of the individual states. As it turns out, neither the GHZ-cost nor the GHZ-distillability of pure quantum states are additive under the flattened tensor product. For SLOCC, the GHZ-cost and GHZ-value of a pure state, $|\psi\rangle$, relates naturally to the logarithm of the rank and sub-rank of the corresponding tensor ψ (see Proposition 2.1.5 and the subsequent discussion). The fact that GHZ-cost is sub-additive is, by the state-tensor correspondence, equivalent to tensor rank being sub-multiplicative under the flattened tensor product. In order to better understand how this strict sub-multiplicativity occurs, it was asked and answered in [11], whether the strict sub-multiplicativity was dependent on the flattening. It turns out that the rank drop can happen both in the process of taking the non-flattened tensor product and with taking the flattening. The first examples found of strict sub-multiplicativity of tensor rank were found with tensors that have a gap between tensor rank, $R(\psi)$, and border rank $\underline{R}(\psi)$.

$R(\psi) > \underline{R}(\psi)$ means that $R(\psi)$ is on the boundary of the rank-$\underline{R}(\psi)$ tensors, implying that these examples of strict sub-multiplicativity are very non-generic. This motivated the question of whether also border rank can be strictly sub-multiplicative, which turned out also to be possible [12]. These results will be discussed in Chapter 2.

In Chapter 3 a tool, called the asymptotic spectrum of tensors, is described. It is a tool for working with many copies in the resource theory of tensors under restrictions. The asymptotic spectrum, which consists of certain homomorphic monotones, entirely determine the asymptotic conversion rates between tensors. The known spectral points from the literature [8, 13] are presented, together with some of the implications for conversion rates between pure states under SLOCC.

In an effort to move from SLOCC to LOCC, [14] constructed a refinement of the asymptotic spectrum of tensors, which retains some control on the asymptotic behavior of the stochastic part (the **S**) in SLOCC. This refinement is called the asymptotic spectrum of LOCC transformations, and is introduced in Chapter 4. Concretely, the asymptotic spectrum of tensors from [15] deals with pure states under SLOCC, while the refinement deals with pure states under unnormalized LOCC transformations. Just like the asymptotic spectrum of tensors encodes all information on conversion rates under SLOCC, so does the LOCC spectrum describe the conversion rates under LOCC in the regime of converse error exponents (see Eq. (4.2)). A characterization of LOCC spectral points is given in the multipartite case. In the bipartite case, the entire LOCC spectrum is determined and a concrete formula for rates, given converse error exponents, is given (see Section 4.3).

Though entropy of entanglement is a good measure of pure bipartite entanglement, it deals with approximate entanglement transformations, as is appropriate when describing physics. For exact transformations, which is the topic of this thesis, the formula from Section 4.3, mentioned above, hints at a formula for conversion rates for exact LOCC transformations of bipartite states, both for success probability going to 1 (Theorem 5.3.3) and for deterministic conversion (Theorem 5.2.9, conjectured in [16]). These formulas are shown in [14] and [17] respectively. In this thesis they are both shown in Chapter 5.

The majority of this thesis is a recap of large parts of the original work in [11], [12], [14] and [17], which has been carried out in collaboration with co-authors M. Christandl ([11], [12]), F.

Gesmundo ([12]), J. Zuiddam ([11]) and P. Vrana ([14]). The results on non-multiplicativity of tensor rank and border rank in Chapter 2 are from [11] and [12], respectively, while the results in Chapter 4 and Chapter 5 are from [14] and [17], respectively, with the exception that a section from [14] has been moved to Section 5.3 as this section depends on results in [17]. The contents of Section 3.4 is work done in collaboration with Péter Vrana.

1.1. Introduction to LOCC

The fundamental objects of study in this thesis are quantum states spread out on multiple quantum systems, also known as multipartite quantum states. In this thesis, a quantum system A is represented by a finite dimensional Hilbert space $\mathcal{H}_A$, known as the state space. Given multiple state spaces, $\mathcal{H}_1, \ldots, \mathcal{H}_k$, the composite system is represented by the k-partite state space $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k$. This thesis will make use of Dirac notation (see Appendix A.2), in particular inner products will be conjugate linear in the first variable and linear in the second. A state of the system A is represented by a positive semidefinite operator ρ on $\mathcal{H}_A$ with $\mathrm{Tr}(\rho) = 1$, known as a density operator. The set of density operators on the state space $\mathcal{H}_A$ will be denoted by $\mathcal{S}(\mathcal{H}_A)$. Of particular interest are the pure states, represented by one-dimensional projections, $|\psi\rangle\langle\psi| \in \mathcal{S}(\mathcal{H}_A)$.

Given a quantum state we wish to describe the ways in which it can be manipulated. If a party has access to the entire system $\mathcal{H}$, they may apply quantum channels to the state. In reality there will be limitations on which quantum channels can practically be implemented, but we idealize and imagine that any channel can be implemented.

Definition 1.1.1. *A linear map* $\Lambda : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{H}')$ *is said to be* completely positive *if the map* $\Lambda \otimes \mathrm{Id}_{\mathrm{End}(\mathbb{C}^n)} : \mathrm{End}(\mathcal{H}) \otimes \mathrm{End}(\mathbb{C}^n) \to \mathrm{End}(\mathcal{H}') \otimes \mathrm{End}(\mathbb{C}^n)$ *is positive (sends positive semidefinite operators to positive semidefinite operators) for all* $n \in \mathbb{N}$*.*

Definition 1.1.2. *Given two state spaces* $\mathcal{H}$ *and* $\mathcal{H}'$*, a quantum channel* Λ *is a completely positive, trace preserving (commonly abbreviated CPTP) linear map* $\Lambda : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}')$ *[18]. If* Λ *is completely positive and trace non-increasing, then we say that* Λ *is an unnormalized quantum channel.*

By linear, we mean that Λ extends to a linear map $\mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{H}')$. Note that the positive and trace preserving maps $\mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{H}')$ are exactly the maps that map density operators to density operators. We imagine that the k-partite state $\rho \in \mathcal{S}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k)$ is split between k spatially separated locations. At each location, one might manipulate the state by performing some quantum operation, represented by locally applying a quantum channel. By complete positivity it also makes sense to talk of local application of quantum channels in the following sense:

Definition 1.1.3. *Given two k-partite state spaces* $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_i \otimes \cdots \otimes \mathcal{H}_k$ *and* $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}'_i \otimes \cdots \otimes \mathcal{H}_k$*, whose i'th system are possibly different, a* local *quantum channel on the i'th system, is a quantum* *channel that can be written as a tensor product*

$$\mathrm{Id}_1 \otimes \cdots \otimes \Lambda \otimes \cdots \otimes \mathrm{Id}_k : \mathcal{S}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_i \otimes \cdots \otimes \mathcal{H}_k) \to \mathcal{S}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_i' \otimes \cdots \otimes \mathcal{H}_k),$$

for some quantum channel $\Lambda : \mathcal{S}(\mathcal{H}_i) \to \mathcal{S}(\mathcal{H}_i')$*.*

As is common in quantum theory, when $\Lambda : \mathcal{S}(\mathcal{H}_i) \to \mathcal{S}(\mathcal{H}_i')$ and $\rho \in \mathcal{S}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k)$, then we shall understand $\Lambda\rho = \Lambda(\rho)$ to mean $(\mathrm{Id}_1 \otimes \cdots \otimes \Lambda \otimes \cdots \otimes \mathrm{Id}_k)(\rho)$. Similarly, when $K : \mathcal{H}_i \to \mathcal{H}_i'$ is a linear map and $|\psi\rangle \in \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k$, we shall understand $K|\psi\rangle$ to mean $(I_{\mathcal{H}_1} \otimes \cdots \otimes K \otimes \cdots \otimes I_{\mathcal{H}_k})|\psi\rangle$. If we imagine that the k parties sharing a state ρ can only manipulate their own part of the system and have no way of sending information, either classical or quantum, between each other, then compositions of the above operations are the only channels available. We might call compositions of above channels LO-channels, where LO is short for Local Operations.

A state $\rho \in M_d(\mathbb{C})$ is called classical, if ρ is diagonal. Given a probability distribution $P : [d] \to [0, 1]$, the diagonal density operator $\rho = \sum_{i=1}^{d} P(i) |i\rangle\langle i|$ represents a physical system, with d possible states, which is in state i with probability $P(i)$. One might reasonably say that a classical system is always in some definite state, and that describing the state of a classical system probabilistically is weird. As such the density operator ρ, does not really describe the intrinsic state of the system, but rather describes an observers knowledge of the system, or perhaps an observers knowledge of a future state of a system, given some stochastic process which is to be applied. In my view, these are the appropriate ways to think of a state, quantum or classical. There are multiple ways that scientists interpret the physical meaning of a density operator, so the above description should merely be thought of as my own personal interpretation and a suggestion to the uninitiated, rather than a postulate on how one should interpret the physicality. Any reader with a separate interpretation of the physical meaning of a quantum state is therefore welcome to apply their own understanding to what follows. The physical interpretation of quantum mechanics is not the topic of this thesis, and will not be further discussed.

Since the ultimate goal is to understand and classify entanglement, we need to distinguish between shared entanglement and shared classical randomness. For this reason we want to work with a resource theory where shared classical randomness is free, which motivates the extension of allowed operations from just local quantum operations to combinations of local quantum operations and sharing of classical information (LOCC). An LOCC channel is the result of applying an LOCC protocol (precisely defined in Definition 1.1.10 below). An LOCC protocol is a protocol where the parties in turn perform local quantum operations, known to the other parties, and share their measurement results. In order to formalize the notion of sharing classical information, we first define a quantum-classical state: A quantum-classical state is a bipartite state $\rho \in \mathcal{S}(\mathcal{H} \otimes \mathbb{C}^{\mathcal{X}})$ of the form

$$\sum_{x \in \mathcal{X}} P(x) \rho_x \otimes |x\rangle\langle x| \,, \tag{1.1}$$

where $\mathcal{X}$ is some finite set, each $\rho_i \in \mathcal{S}(\mathcal{H})$ and $P : \mathcal{X} \to [0,1]$ is a probability distribution on $\mathcal{X}$; i.e. $\sum_{x \in \mathcal{X}} P(x) = 1$. The quantum classical states are precisely the intersection of $\mathcal{S}(\mathcal{H} \otimes \mathbb{C}^{\mathcal{X}})$ and $\mathrm{End}(\mathcal{H}) \otimes \mathrm{Diag}(\mathbb{C}^{\mathcal{X}})$, where $\mathrm{Diag}(\mathbb{C}^{\mathcal{X}})$ denotes the set of diagonal matrices w.r.t. the basis $(|x\rangle)_{x \in \mathcal{X}}$. A quantum-classical state represents the joint state of a system which has both a quantum and a classical part. For instance, this could be the joint system of some quantum system in a laboratory together with either a classical register storing the results of measurements, or a monitor which a scientist uses to read off the results of measurements. The scientist reading off the result would then correspond to a measurement of the classical register in the computational basis. As such, the state (1.1) may be interpreted as the quantum part being in the state ρ_x with probability $P(x)$, while the classical part flags the state of the quantum system. If we want to model a channel that performs measurements and stores the measurement result in a classical register, we start by adding a classical register to the output system. If we demand that a channel $\Lambda : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}' \otimes \mathbb{C}^d)$ only outputs quantum-classical states, then one sees that Λ can be written as

$$\Lambda : \rho \mapsto \sum_{j \in [d]} \mathcal{E}_j(\rho) \otimes |j\rangle\langle j| \,, \tag{1.2}$$

where $(\mathcal{E}_j)_{j \in [d]}$ is a family of unnormalized quantum channels $\mathcal{E}_j : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}')$ with $\sum_j \mathcal{E}_j$ trace preserving. Note that from Eq. (1.2) $\mathrm{Tr} \circ \sum_j \mathcal{E}_j = \mathrm{Tr} \circ \Lambda$. The collection $\mathcal{E} = (\mathcal{E}_j)_{j \in [d]}$ is called a *quantum instrument* and we interpret the map (1.2) as an operation involving measurements of the quantum part with a total of d different possible outcomes. The resulting quantum state after applying a quantum instrument is thus $\mathcal{E}_j(\rho)$ with probability $\mathrm{Tr}\, \mathcal{E}_j(\rho)$ and the information about the resulting state is stored in the classical register as $|j\rangle\langle j|$. The maps $\mathcal{E}_j$ defining the instrument $\mathcal{E}$ will be called the components of $\mathcal{E}$. If $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k$ is a composite quantum system and each $\mathcal{E}_j$ is local with respect to the same subsystem, $\mathcal{H}_i$, in the sense of Definition 1.1.3, we say that $(\mathcal{E}_j)_{j \in [d]}$ is a *local quantum instrument.*

Local quantum instruments form the basis for the concept of LOCC conversion. If k parties share a state $\rho \in \mathcal{S}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k)$, and they may share classical information, then we simply assume that the results of quantum operations are public knowledge. We model this by adding a public classical register, $\mathbb{C}^d$, which each party has access to. The state space, including register, is then

$$\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k \otimes \mathbb{C}^d, \tag{1.3}$$

and a state of the full system is a quantum-classical state, i.e. an element $\rho \in \mathcal{S}\left(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k \otimes \mathbb{C}^d\right)$ of the form

$$\rho = \sum_{j=1}^{d} P(j)\rho_j \otimes |j\rangle\langle j| \,, \tag{1.4}$$

where $\rho_j \in \mathcal{S}\left(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k\right)$ and P is some probability distribution on $[d]$. The i'th party may choose which quantum instrument to apply, depending on the information in the public register. Formally that is, they can choose to apply different quantum instruments to each ρ_j:

Definition 1.1.4. *Given a composite quantum-classical system $\mathcal{S}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k) \otimes \mathrm{Diag}(\mathbb{C}^{\mathcal{X}})$, we say that a* conditional application of local quantum instruments *on the i'th system is a channel $\Lambda = \mathcal{E}(J, f)$ given as*

$$\mathcal{E}(J,f) \overset{def}{=} \Lambda : \sum_{x\in\mathcal{X}} \rho_x \otimes |x\rangle\langle x| \mapsto \sum_{x\in\mathcal{X}} \sum_{j\in f^{-1}(x)} \mathcal{E}_j(\rho_x) \otimes |j\rangle\langle j| = \sum_{j\in J} \mathcal{E}_j(\rho_{f(j)}) \otimes |j\rangle\langle j| \,. \tag{1.5}$$

Here $f : J \to \mathcal{X}$ is a map and for all $x \in \mathcal{X}$, $\mathcal{E}^x = (\mathcal{E}_j)_{j\in f^{-1}(x)}$ is a local quantum instrument on the i'th system. J is a set of unique labels flagging the outcome of application of the quantum instruments.

The above definition corresponds to applying the instrument $\mathcal{E}^x = (\mathcal{E}_j)_{j\in f^{-1}(x)}$ conditioned on the register reading x. The resulting state is then $\sum_{j\in J} \mathcal{E}_j(\rho_{f(j)}) \otimes |j\rangle\langle j|$.

Remark 1.1.5. *The map Λ from Definition 1.1.4 is indeed a quantum channel on the surrounding system $\mathcal{S}(\mathcal{H}_1 \otimes \cdots \mathcal{H}_k \otimes \mathbb{C}^{\mathcal{X}})$, since it extends to*

$$\Lambda : \rho \mapsto \sum_{j\in J} \left[\mathcal{E}_j \otimes \Lambda_{|j\rangle\langle f(j)|}\right] \rho.$$

Here $\Lambda_{|j\rangle\langle f(j)|}$ is the map $\Lambda_{|j\rangle\langle f(j)|} : \rho \mapsto |j\rangle\langle f(j)| \, \rho \, |f(j)\rangle\langle j|$, which is completely positive, and complete positivity is stable under both tensor product and sum, so Λ is completely positive. Furthermore

$$\begin{aligned} \operatorname{Tr} \sum_{j\in J} \left[\mathcal{E}_j \otimes \Lambda_{|j\rangle\langle f(j)|}\right] \rho &= \sum_{j\in J} \operatorname{Tr} \mathcal{E}_j \left\langle f(j)\right| \rho \left|f(j)\right\rangle = \sum_{x\in\mathcal{X}} \operatorname{Tr} \sum_{j\in f^{-1}(x)} \mathcal{E}_j \left\langle x\right| \rho \left|x\right\rangle \\ &= \sum_{x\in\mathcal{X}} \operatorname{Tr} \left\langle x\right| \rho \left|x\right\rangle = \operatorname{Tr} \rho, \end{aligned} \tag{1.6}$$

where we have used the fact that $\sum_{j\in f^{-1}(x)} \mathcal{E}_j$ is trace preserving for all $x \in \mathcal{X}$.

Given a quantum-classical state $\rho = \sum_j \rho_j \otimes |j\rangle\langle j|$, the parties may choose to "delete some of the classical information" or "join some of the conditional states". Formally we call this a coarse-graining and it is a channel $\Pi_g : \mathcal{S}(\mathcal{H}) \otimes \mathrm{Diag}(\mathbb{C}^J) \to \mathcal{S}(\mathcal{H}) \otimes \mathrm{Diag}(\mathbb{C}^{\mathcal{X}})$ of the form

$$\Pi_g : \sum_{j \in J} \rho_j \otimes |j\rangle\langle j| \mapsto \sum_{j \in J} \rho_j \otimes |g(j)\rangle\langle g(j)| \tag{1.7}$$

where $g : J \mapsto \mathcal{X}$ is some map into the output register $\mathcal{X}$.

Example 1.1.6. *When $g : J \to \{0\}$,*

$$\Pi_g : \sum_{j \in J} \rho_j \otimes |j\rangle\langle j| \mapsto \left(\sum_{j \in J} \rho_j \right) \otimes |0\rangle\langle 0| . \tag{1.8}$$

If we make the association $\mathcal{H} \otimes \mathbb{C} \simeq \mathcal{H}$, Π_g corresponds to taking the partial trace of the register system. In this case we shall write $\Pi_g = \mathrm{Tr}_{reg}$.

Definition 1.1.7 (LOCC)**.** *We define a one-step LOCC channel to be a channel Λ which is the composition of a conditional application of local quantum instruments at a subsystem i, followed by a coarse-graining.*

An LOCC protocol is a finite sequence of composable one-step LOCC channels and the composition is an LOCC channel.

One might ask why we allow for coarse-graining after each local application of quantum instruments. All the coarse-graining does is throw away information, which can hardly be useful in achieving a conversion task. This intuition is valid as we shall now see, that any LOCC channel may be implemented in such a manner that any coarse-graining is deferred to the end of the protocol.

Lemma 1.1.8. *Let $\Lambda = \mathcal{E}(J, f) \circ \Pi_g$ be an LOCC channel acting on the space $\mathcal{S}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k)$ $\otimes$ $\mathrm{Diag}(\mathbb{C}^{\mathcal{X}})$. Π_g is a coarse-graining given by a map $g : \mathcal{X} \to \mathcal{Y}$. $\mathcal{E}(J, f)$ is a conditional application of local quantum instruments $f : J \to \mathcal{Y}$.*

Then there exists a conditional application of local quantum instruments $\tilde{\mathcal{E}}(\tilde{J}, \tilde{f})$, $f : \tilde{J} \to \mathcal{X}$ and $\Pi_{\tilde{g}}$, $g : J \to \tilde{J}$ such that

$$\Lambda = \Pi_{\tilde{g}} \circ \tilde{\mathcal{E}}(\tilde{J}, \tilde{f}).$$

Proof. Let $\rho = \sum_{x \in \mathcal{X}} \rho_x \otimes |x\rangle\langle x|$, then

$$\Lambda : \rho \overset{\Pi_g}{\mapsto} \sum_{x \in \mathcal{X}} \rho_x \otimes |g(x)\rangle\langle g(x)| \overset{\mathcal{E}(J,f)}{\mapsto} \sum_{x \in \mathcal{X}} \sum_{j \in f^{-1}(g(x))} \mathcal{E}_j(\rho_x) \otimes |j\rangle\langle j| = \sum_{j \in J} \sum_{\substack{x \in \mathcal{X} \\ g(x) = f(j)}} \mathcal{E}_j(\rho_x) \otimes |j\rangle\langle j| . \tag{1.9}$$

Let $\tilde{J} = \{(j,x) \in J \times \mathcal{X} | f(j) = g(x)\}$ and $\tilde{f} : \tilde{J} \to \mathcal{X}$ be the map $(j,x) \mapsto x$. Define the conditional quantum instrument $\tilde{\mathcal{E}}(\tilde{J}, \tilde{f})$ by $\tilde{\mathcal{E}}_{(j,x)} = \mathcal{E}_j$. Let $\Pi_{\tilde{g}}$ be given by the map $\tilde{g} : \tilde{J} \to J$ defined as $\tilde{g} : (j,x) \mapsto j$. Then

$$\begin{aligned}\Pi_{\tilde{g}} \circ \tilde{\mathcal{E}}(\tilde{J}, \tilde{f}) : \rho &\overset{\tilde{\mathcal{E}}(\tilde{J},\tilde{f})}{\mapsto} \sum_{(j,x) \in \tilde{J}} \tilde{\mathcal{E}}_{(j,x)} \left(\rho_{\tilde{f}(j,x)}\right) \otimes |j,x\rangle\langle j,x| = \sum_{(j,x) \in \tilde{J}} \mathcal{E}_j\left(\rho_x\right) \otimes |j,x\rangle\langle j,x| \\ &\overset{\Pi_{\tilde{g}}}{\mapsto} \sum_{(j,x) \in \tilde{J}} \mathcal{E}_j\left(\rho_x\right) \otimes |j\rangle\langle j| .\end{aligned} \tag{1.10}$$

Note that the right-hand-side of (1.9) and (1.10) are equal. □

Proposition 1.1.9. *Let Λ be an LOCC channel. Then Λ can be written as a composition of conditional applications of local quantum instruments, followed by a single coarse-graining.*

Proof. By Definition 1.1.7, Λ is of the form

$$\Lambda = \Pi_{g_n} \mathcal{E}^n(J_n, f_n) \cdots \Pi_{g_1} \mathcal{E}^1(J_1, f_1), \tag{1.11}$$

for some choices of conditional instruments and coarse-grainings. Since the composition of two coarse-grainings is again a coarse-graining, we apply Lemma 1.1.8 $n-1$ times to obtain

$$\Lambda = \Pi_{\tilde{g}} \tilde{\mathcal{E}}^n(\tilde{J}_n, \tilde{f}_n) \cdots \tilde{\mathcal{E}}^1(\tilde{J}_1, \tilde{f}_1). \tag{1.12}$$

□

Since

$$\mathcal{S}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k) \overset{\rho \leftrightarrow \rho \otimes |0\rangle\langle 0|}{\simeq} \mathcal{S}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k) \otimes \mathrm{Diag}(\mathbb{C}) \tag{1.13}$$

we may associate any k-partite state on the left-hand-side, with the quantum-classical state on the right-hand-side. This corresponds to considering a k-partite quantum state versus considering the same quantum state together with some shared classical system in a default state. Using this association, it makes sense to apply LOCC channels to entirely quantum states. Given two k-partite states ρ and σ we write

$$\rho \xrightarrow{\mathrm{LOCC}} \sigma \overset{def}{\iff} \exists \Lambda \in \mathrm{LOCC} : \Lambda(\rho) = \sigma. \tag{1.14}$$

The above tranformation of ρ to σ is an exact, deterministic LOCC-tranformation, in the sense that the output is exactly equal to σ and the channel outputs σ with certainty. As mentioned, we shall also consider probabilistic LOCC-conversion. Given a k-partite state ρ and some LOCC-channel Λ as defined in Definition 1.1.7, the output is of the form

$$\sum_{x \in \mathcal{X}} P(x) \rho_x \otimes |x\rangle\langle x| \tag{1.15}$$

for some family $(\rho_x)_{x\in\mathcal{X}}$ and probability distribution $P : \mathcal{X} \to [0,1]$, corresponding to the quantum part being in state ρ_x with probability $P(x)$, flagged by the classical register. This formalizes what we mean by ρ being LOCC-convertible to ρ_x with some probability $P(x)$. A convenient way of characterising probabilistic conversion is by the use of trace non-increasing LOCC channels.

Definition 1.1.10 (Unnormalized LOCC)**.** *A trace non-increasing one-step LOCC channel is defined exactly as the usual one-step LOCC channels in Definition 1.1.7, except that instead of conditional application of local quantum instruments, we allow for conditional application of local trace-non-increacing quantum instruments. That is, instruments $(\mathcal{E}_j)_{j\in f^{-1}(x)}$ such that $\sum_{j\in f^{-1}(x)}\mathcal{E}_j$ is trace-non-increasing.*

A trace-non-increasing LOCC channel is a finite composition of trace non-increasing one-step LOCC channels.

Remark 1.1.11. *Lemma 1.1.8 and Proposition 1.1.9 also work for trace non-increasing LOCC channels, by the exact same proofs.*

Given a trace non-increasing conditional instrument $\mathcal{E}(J,f)$ acting on $\mathcal{S}(\mathcal{H}_1\otimes\cdots\otimes\mathcal{H}_k)\otimes\mathrm{Diag}(\mathbb{C}^{\mathcal{X}})$

$$\Lambda : \sum_{x\in\mathcal{X}}\rho_x\otimes|x\rangle\langle x| \mapsto \sum_{j\in J}\mathcal{E}_j\left(\rho_{f(j)}\right)\otimes|j\rangle\langle j|\,, \tag{1.16}$$

we may extend to a trace preserving LOCC channel by introducing the instrument $\tilde{\mathcal{E}}(\tilde{J},\tilde{f})$, where $\tilde{J} = J\sqcup\mathcal{X}$ and $\tilde{f}$ extends f by $\tilde{f}(x) = x$ for $x\in\mathcal{X}\subset\tilde{J}$. $\tilde{\mathcal{E}}_j = \mathcal{E}_j$ for $j\in J\subset\tilde{J}$ and

$$\tilde{\mathcal{E}}_x : \rho_x \mapsto 1 - \mathrm{Tr}\left(\sum_{j\in f^{-1}(x)}\mathcal{E}_j(\rho_x)\right)\rho_0, \tag{1.17}$$

where ρ_0 is some fixed state, which we might consider as a failure output. In order to group the failures together we post-compose with the coarse-graining Π_g where $g : \tilde{J}\to J\oplus\{\perp\}$ is the identity on $J\subset\tilde{J}$ and sends $\mathcal{X}\subset\tilde{J}$ to $\perp$. The resulting channel $\tilde{\Lambda} = \Pi_g\tilde{\mathcal{E}}(\tilde{J},\tilde{f})$ is trace preserving and acts as

$$\tilde{\Lambda} = \Pi_g\tilde{\mathcal{E}}(\tilde{J},\tilde{f}) : \sum_{x\in\mathcal{X}}\rho_x\otimes|x\rangle\langle x| \mapsto \sum_{j\in J}\mathcal{E}_j(\rho_{f(j)})\otimes|j\rangle\langle j| + C_\rho\rho_0\otimes|\perp\rangle\langle\perp|\,, \tag{1.18}$$

where the scalar $C_\rho = 1-\sum_{x\in\mathcal{X}}\mathrm{Tr}\,\Lambda\rho_x$. The convenience of introducing trace non-increasing LOCC channels is captured by the following proposition:

Proposition 1.1.12. *A k-partite state ρ can be LOCC-transformed into σ with probability p if and only if*

$$\Lambda(\rho) = p\sigma \tag{1.19}$$

for some trace-non-increasing LOCC channel Λ. When this is the case we write $\rho \xrightarrow{\text{LOCC}} p\sigma$.

Proof. The "only-if"-part is the easy implication: Assume that ρ can be LOCC-transformed into σ with probability p. That is, for some trace-preserving LOCC channel Λ,

$$\Lambda(\rho) = \sum_{x \in \mathcal{X}} P(x) \rho_x \otimes |x\rangle\langle x| \,, \tag{1.20}$$

where $\rho_{x_0} = \sigma$ and $P(x_0) = p$ for some $x_0 \in \mathcal{X}$. Now let Λ_0 be the one-step trace non-increasing LOCC channel, acting on the output system of Λ;

$$\Lambda_0 : \sum_{x \in \mathcal{X}} \rho_x \otimes |x\rangle\langle x| \mapsto \rho_{x_0} \otimes |0\rangle\langle 0| \,. \tag{1.21}$$

Then $\Lambda_0 \circ \Lambda$ witnesses $\rho \xrightarrow{\text{LOCC}} p\sigma$.

Conversely, suppose $\Lambda(\rho) = p\sigma$ for some trace non-increasing LOCC channel. By Proposition 1.1.9, we can assume that $\Lambda = \mathrm{Tr}_{reg}\, \mathcal{E}^n(J_n, f_n) \cdots \mathcal{E}^1(J_1, f_1)$ is a composition of trace non-increasing conditional quantum instruments followed by a coarse-graining, which must necessarily be the partial trace of the register, since the output $p\sigma$ which is independent of the register. Since $\Lambda(\rho) = p\sigma$ we must have

$$\mathcal{E}^n(J_n, f_n) \cdots \mathcal{E}^1(J_1, f_1)\rho = \sum_{j \in J_n} P(j) \rho_j \otimes |j\rangle\langle j| \,, \tag{1.22}$$

where $\sum_{j \in J_n} P(j)\rho_j = p\sigma$. Starting from the right, we iteratively replace each $\mathcal{E}^l(J_l, f_l)$ with the corresponding $\Lambda_l = \Pi_{g_l} \tilde{\mathcal{E}}^l(\tilde{J}_l, \tilde{f}_l)$ from (1.17), such that

$$\Lambda_n \cdots \Lambda_1 \rho = \sum_{j \in J_n} P(j) \rho_j + (1-p)\rho_0 \otimes |\bot\rangle\langle \bot| \,. \tag{1.23}$$

Now apply the coarse-graining g which contracts J_n to a single register point 0 and

$$\Pi_g \Lambda_n \cdots \Lambda_1 \rho = p\sigma \otimes |0\rangle\langle 0| + (1-p)\rho_0 \otimes |\bot\rangle\langle \bot| \,. \tag{1.24}$$

$\square$

Definition 1.1.13. *When $\rho \xrightarrow{\text{LOCC}} p\sigma$ for some $p > 0$ we say that ρ can be transformed to σ via* SLOCC *and write $\rho \xrightarrow{\text{SLOCC}} \sigma$.*

In particular when dealing with pure states, it is often convenient to break down the completely positive maps into their Kraus decompositions.

Theorem 1.1.14 (Kraus representation). *A map $\Lambda : \mathrm{End}(\mathcal{H}) \to \mathrm{End}(\mathcal{H}')$ is completely positive if and only if there exists a finite family of operators $(K_j)_{j\in J} \subset \mathrm{End}(\mathcal{H}, \mathcal{H}')$ such that for all states $\rho \in \mathcal{S}(\mathcal{H})$*

$$\Lambda(\rho) = \sum_{j\in J} K_j \rho K_j^*. \tag{1.25}$$

Furthermore, Λ is trace preserving if and only if $\sum_j K_j^* K_j = I_{\mathcal{H}}$ and trace non-increasing if and only if $\sum_j K_j^* K_j \le I_{\mathcal{H}}$.

The minimal number of Kraus operators needed to represent a completely positive map is called the Kraus rank. This coincides with the rank of the Choi-state associated to Λ. Now for any Kraus decomposition $\Lambda : \rho \mapsto \sum_j K_j \rho K_j^*$ of a completely positive map, we may associate the instrument $\mathcal{E}$ with components $\mathcal{E}_j : \rho \mapsto K_j \rho K_j^*$. Λ is trace preserving, respectively trace non-increasing, if and only if the associated instrument is trace preserving, respectively trace non-increasing. In fact, any conditional quantum instrument $\mathcal{E}(J, f)$ may be written as such a "Kraus instrument" followed by a coarse-graining. Simply replace each $\mathcal{E}_j$ with the instrument defined by a collection of Kraus operators $(K_l^j)_{l\in L_j}$ representing $\mathcal{E}_j$. That is, consider the instrument $\tilde{\mathcal{E}}(\bigsqcup_j L_j, \tilde{f})$ consisting of all the maps $\mathcal{E}_{j,l} : \rho \mapsto K_l^j \rho (K_l^j)^*$ for all $j \in J$ and $l \in L_j$ and let $\tilde{f} : (j, l) \mapsto f(j)$. Post-composing with the coarse-graining $L_j \to j$ reproduces the channel Λ. In light of this construction and Proposition 1.1.9 we get the following:

Proposition 1.1.15. *Λ is an LOCC channel (respectively trace non-increasing LOCC channel) if and only if*

$$\Lambda = \Pi_g \Lambda_n \cdots \Lambda_1 \tag{1.26}$$

for some $n \in \mathbb{N}$ where each Λ_l is a conditional application of local quantum instruments, with the components of each condtitional instrument having Kraus rank 1. I.e. Λ_l is of the form

$$\Lambda_l : \rho \mapsto \sum_j \Big(K_j \otimes |j\rangle\langle f(j)|\Big)\, \rho\, \Big(K_j^* \otimes |f(j)\rangle\langle j|\Big) \tag{1.27}$$

where $\rho \in \mathcal{S}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k) \otimes \mathrm{Diag}(\mathbb{C}^{\mathcal{X}})$, $f : J \to \mathcal{X}$, $K_j : \mathcal{H}_i \to \mathcal{H}_i'$ for some $i \in [k]$ and $\sum_{j\in J} K_j^* K_j = I_{\mathcal{H}_i}$ (respectively $\sum_{j\in J} K_j^* K_j \le I_{\mathcal{H}_i}$). Furthermore if the output of Λ has trivial register, $\Pi_g = \mathrm{Tr}_{reg}$.

Proposition 1.1.16. *Given two pure, k-partite states $|\psi\rangle\langle\psi|\, \mathcal{S}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k)$ and $|\phi\rangle\langle\phi| \in \mathcal{S}(\mathcal{H}_1' \otimes \cdots \otimes \mathcal{H}_k')$, the following are equivalent.*

1. *$|\psi\rangle\langle\psi| \xrightarrow{\mathrm{SLOCC}} |\phi\rangle\langle\phi|$*

2. *There exist linear maps $A_i : \mathcal{H}_i \to \mathcal{H}_i'$ such that*

$$(A_1 \otimes \cdots \otimes A_k)\, |\psi\rangle = |\phi\rangle \tag{1.28}$$

Proof. First we show 2 $\implies$ 1. Assume that $(A_1 \otimes \cdots \otimes A_k) |\psi\rangle = |\phi\rangle$. Let

$$\Lambda_l : \rho \mapsto \frac{1}{\operatorname{Tr}(A_l A_l^*)} A_l \rho A_l^*. \tag{1.29}$$

Then $\Lambda_n \cdots \Lambda_1 |\psi\rangle\langle\psi| = \frac{1}{\operatorname{Tr} A_n \cdots \operatorname{Tr} A_1} |\phi\rangle\langle\phi|$, as claimed.

Then we show 1 $\implies$ 2. Assume that $|\psi\rangle\langle\psi| \xrightarrow{\text{SLOCC}} |\phi\rangle\langle\phi|$. Then $\Lambda |\psi\rangle\langle\psi| = p |\phi\rangle\langle\phi|$ for some $p > 0$ and

$$\Lambda = \operatorname{Tr}_{reg} \Lambda_n \cdots \Lambda_1, \tag{1.30}$$

where each Λ_l is of the form (1.27). Let $J_l, f_l, (K_j^l)_{j \in J_l}, i_l$ be the objects definining Λ_l from Proposition 1.1.15. Now

$$\Lambda_n \cdots \Lambda_1 |\psi\rangle\langle\psi| = \sum_{(j_1,\ldots,j_n) \in \mathcal{J}} \left(K_{j_n}^n \cdots K_{j_1}^1\right) |\psi\rangle\langle\psi| \left(K_{j_n}^n \cdots K_{j_1}^1\right)^* \otimes |j_n\rangle\langle j_n|, \tag{1.31}$$

where $\mathcal{J} = \{(j_1, \ldots, j_n) \in J_1 \times \cdots \times J_n | f_l(j_l) = j_{l-1} \text{ for } l = 2, \ldots, n\}$. Since tracing out the register yields $p |\phi\rangle\langle\phi|$, we conclude that $\left(K_{j_n}^n \cdots K_{j_1}^1\right) |\psi\rangle\langle\psi| \left(K_{j_n}^n \cdots K_{j_1}^1\right)^*$ must be a multiple of $|\phi\rangle\langle\phi|$ for all $(j_1, \ldots, j_n) \in \mathcal{J}$. Since each K_j is local, $K_{j_n}^n \cdots K_{j_1}^1$ is of the form $A_1 \otimes \cdots \otimes A_k$, so by rescaling we obtain 2. □

Given a k-partite quantum state $\rho \in \mathcal{S}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k)$, each party can apply the local quantum channel $\operatorname{Tr}_i : \mathcal{S}(\mathcal{H}_i) \to \mathbb{C}$, resulting in the trivial state $\rho_0 = |0\ldots0\rangle\langle 0\ldots0|$. So $\rho \xrightarrow{\text{LOCC}} \rho_0$ for all k-partite states ρ, showing that ρ_0 is a smallest element in the resource theory of k-partite entanglement and any smallest element must necessarily be in the LOCC orbit of ρ_0.. The LOCC orbit of ρ is the class of states that can be reached from ρ via LOCC and for ρ_0 the LOCC orbit is the class of *separable* states. I.e. states of the form $\sum_j p_j \rho_{1,j} \otimes \cdots \otimes \rho_{k,j}$, where $\rho_{i,j} \in \mathcal{S}(\mathcal{H}_i)$ and $\sum_j p_j = 1$. These states are precisely the ones corresponding to the k parties only having shared classical randomness: The shared state is $\rho_{1,j} \otimes \cdots \otimes \rho_{k,j}$ with probability p_j. Therefore, entangled states are by definition the states which are not separable. States that are pure and separable are necessarily product states, i.e. states of the form $|\psi\rangle = |\psi_1\rangle \otimes \cdots \otimes |\psi_k\rangle$.

1.1.1. Multiple copies

When our imagined k parties share two states with density operators $\rho_1 \in \mathcal{S}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k)$ and $\rho_2 \in \mathcal{S}(\mathcal{K}_1 \otimes \cdots \otimes \mathcal{K}_k)$, then the combined state is described by the density operator $\rho_1 \otimes \rho_2 \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{K}_1 \otimes \cdots \otimes \mathcal{H}_k \otimes \mathcal{K}_k)$ and for LOCC protocols the i'th party may perform quantum operations that are local to the subsystem $(\mathcal{H}_i \otimes \mathcal{K}_i)$, as we imagine that both systems are in their possession. When we interpret $\rho_1 \otimes \rho_2$ as a k-partite state in this manner we shall write $\rho_1 \odot \rho_2$. But we might also interpret the density operator $\rho_1 \otimes \rho_2$ as a $2k$-partite state, where all quantum instruments in an LOCC protocol have to be local with respect to $\mathcal{H}_i$ or $\mathcal{K}_i$, but not across these systems. When we interpret $\rho_1 \otimes \rho_2$ as a $2k$-partite state we shall stick with the notation $\rho_1 \otimes \rho_2$. It is non-standard to use the notation $\rho_1 \odot \rho_2$ for the so-called Kronecker product or "flattened" tensor product, and in most of the literature the interpretation of $\otimes$ is inferred from context with little risk of confusion. The notation is mainly introduced here, because of the distinction being necesarry in the study of the multiplicativity of tensor rank in Chapter 2. In the paper [11], in which the non-multiplicativity of tensor rank under $\otimes$ was first described, the symbol $\boxtimes$ was used for the Kronecker product. The choice of using $\odot$ instead of $\boxtimes$ is mainly aesthetic and because of the lack of available Latex packages with a neat $\boxtimes$ version of the large symbols $\bigotimes$ and $\bigodot$ used for writing the application of the operation to a family of objects.

For some quantum states $\rho \overset{\text{LOCC}}{\not\longrightarrow} \sigma$, yet $\rho \odot \rho \xrightarrow{\text{LOCC}} \sigma \odot \sigma$. So if our k parties have a pool of multiple copies of a resource state, ρ, and wish to convert these to as many copies of a target state, σ, as possible, then they can often benefit from manipulating the entire resource pool jointly, rather than each resource state individually.

This thesis is mostly concerned with the asymptotic behavior of entanglement manipulation. Given many copies, $\rho^{\odot n} = \rho \odot \cdots \odot \rho$, of a resource state, ρ, we ask how many copies of a certain target state, σ, can be obtained per copy of ρ via LOCC conversion. We call this number of copies of σ obtained per copy of ρ the conversion rate. This rate has many variations, depending on which demands we make on the asymptotic precision of the LOCC conversion. We might demand the conversion to be exact or allow for some loss in fidelity and make certain demands on the asymptotic behavior of this fidelity loss. We could also allow for a probability of failed conversion and just like with fidelity make demands on the asymptotic behavior of success probability. Most of these rates are only well understood for pure, bipartite states, if even for these.

The exact, deterministic rate is

$$E_{exact}(\rho, \sigma) = \sup \left\{ \tau \in \mathbb{R}^+ \middle| \rho^{\odot n} \xrightarrow{\text{LOCC}} \sigma^{\odot \lfloor n\tau \rfloor} \text{ for } n \gg 1 \right\}, \tag{1.32}$$

which will be addressed for pure bipartite states in Chapter 5. The famous result on entropy of entanglement [1][19, ch. 19.4] mentioned in the introduction is that when $\rho = |\psi\rangle\langle\psi|$ and

$\sigma = |\phi\rangle\langle\phi|$ are pure bipartite states,

$$E_{\mathcal{F}>1-\epsilon}(\rho,\sigma) = \lim_{n\to\infty} \sup\left\{\tau \in \mathbb{R}^+ \middle| \exists \sigma_n : \rho^{\odot n} \xrightarrow{\text{LOCC}} \sigma_n \text{ and } F\left(\sigma_n, \sigma^{\odot \lfloor n\tau \rfloor}\right) > 1-\epsilon\right\} = \frac{H(P)}{H(Q)}, \tag{1.33}$$

where P and Q are the squared Schmidt coefficients of $|\psi\rangle$ and $|\phi\rangle$ respectively and $\epsilon > 0$ is some fixed arbitrarily small allowed error. $H(P)$ is defined in Definition 1.2.1.

1.2. Type classes and Shannon entropy

In this section we fix a set $\mathcal{X}$ of size $|\mathcal{X}| = d$ and consider the space of probability distributions $\mathcal{P} = \mathcal{P}(\mathcal{X}) = \left\{P : \mathcal{X} \to [0,1] \middle| \sum_{x\in\mathcal{X}} P(x) = 1\right\}$. In the following we shall sometimes interpret $0\log 0$ as 0, with the justification that $\lim_{x\to 0} x\log x = 0$. Now might also be a good time to mention that $\log = \log_2$ will always denote the 2-logarithm, rather than the natural logarithm, for the entirety of this thesis.

Definition 1.2.1. *For $P \in \mathcal{P}$, the Shannon entropy of P is defined as*

$$H(P) = -\sum_{x\in\operatorname{supp}(P)} P(x)\log P(x) \overset{\text{"}0\log 0=0\text{"}}{=} -\sum_{x\in\mathcal{X}} P(x)\log P(x). \tag{1.34}$$

For $\alpha \in (0,\infty)\backslash\{1\}$, the α-Rényi entropy is defined as

$$H_\alpha(P) = \frac{1}{1-\alpha}\log\sum_{x\in\mathcal{X}} P(x)^\alpha. \tag{1.35}$$

For $\alpha \in \{0,1,\infty\}$, we define $H_\alpha(P)$ by taking the limit $\lim\limits_{\beta\to\alpha} H_\beta(P)$. That is,

$$H_0(P) = \log|\operatorname{supp}(P)|, \tag{1.36}$$

$$H_1(P) = H(P), \tag{1.37}$$

and

$$H_\infty(P) = -\max_{x\in\mathcal{X}}\log P(x). \tag{1.38}$$

Definition 1.2.2. *Given two probability distributions $P, Q \in \mathcal{P}$ with $\operatorname{supp}(Q) \subseteq \operatorname{supp}(P)$, we define the relative entropy, also called the Kullback-Leibler divergence as*

$$D(Q||P) = \sum_{x\in\operatorname{supp}(Q)} Q(x)\log\left(\frac{Q(x)}{P(x)}\right) \overset{\text{"}0\log 0=0\text{"}}{=} \sum_{x\in\mathcal{X}} Q(x)\log\left(\frac{Q(x)}{P(x)}\right). \tag{1.39}$$

When $\operatorname{supp}(Q) \not\subseteq \operatorname{supp}(P)$, we set $D(Q||P) = \infty$.

We denote by $\mathcal{P}_n \subset \mathcal{P}$, the set of n-types on $\mathcal{X}$, i.e. the set of probability distributions on $\mathcal{X}$ such that all point probabilities are a multiple of $\frac{1}{n}$. For $Q \in \mathcal{P}_n$, and a finite sequence $I \in \mathcal{X}^n$, viewed as a function $I : [n] \to \mathcal{X}$, we say the I is of type Q, if $|I^{-1}(x)| = nQ(x)$ for all $x \in \mathcal{X}$. We denote by $T_Q^n \subset \mathcal{X}^n$ the sequences of type Q. T_Q^n will be called the type class of the type Q. In other words; T_Q^n is the set of sequences of length n, whose relative frequency of symbols are described by Q. Since choosing an element of $Q \in \mathcal{P}_n$ means choosing $nQ(x) \in \{0, \dots, n\}$ for each $x \in \mathcal{X}$, with the restriction that these numbers sum to n, we get a coarse but adequate upper bound on the number of type classes

$$|\mathcal{P}_n| \leq (n+1)^d , \tag{1.40}$$

showing that the number of type classes only grows polynomially in n. Therefore the size of the individual type classes must grow exponentially in n, and this exponential growth is captured by the following lemma:

Lemma 1.2.3. *[20, Lemma 2.3] Let $Q \in \mathcal{P}_n$, then*

$$\frac{1}{(n+1)^d} 2^{nH(Q)} \leq \left| T_Q^n \right| \leq 2^{nH(Q)} \tag{1.41}$$

Applying $\lim_{n\to\infty} \frac{1}{n} \log$ to Eq. (1.41), one obtains the following:

Proposition 1.2.4. *Let $Q \in \mathcal{P}_n$, then*

$$\lim_{n\to\infty} \frac{1}{n} \log \left| T_Q^n \right| = H(Q). \tag{1.42}$$

For a probability distribution $P \in \mathcal{P}(\mathcal{X})$, $P^{\otimes n} \in \mathcal{P}(\mathcal{X}^n)$ is the probability distribution given by $P^{\otimes n}(I) = \prod_{i\in[n]} P\left(I(i)\right)$, corresponding to a sequence of n independent and identically distributed stochastic variables with distribution P. Letting $Q \in \mathcal{P}_n$, $P \in \mathcal{P}$, and $I \in T_Q^n$ we have

$$P^{\otimes n}(I) = \prod_{x\in\mathcal{X}} P(x)^{nQ(x)} = 2^{n\sum_x Q(x)\log P(x)} = 2^{n\left(-H(Q)-D(Q||P)\right)}. \tag{1.43}$$

If we consider a stochastic variable $X = (X_1, \dots, X_n)$, where $X_1, \dots, X_n$ are independent and identically distributed, with distribution $P \in \mathcal{P}(X)$, then the probability that X is of type Q is

$$Pr\left(X \in T_Q^n\right) = \sum_{I\in T_Q^n} P^{\otimes n}(I) = \left|T_Q^n\right| 2^{n\left(-H(Q)-D(Q||P)\right)}. \tag{1.44}$$

By Lemma 1.2.3 and Eq. (1.44), we get:

Lemma 1.2.5. *Let $Q \in \mathcal{P}_n$, then*

$$\frac{1}{(n+1)^d} 2^{-nD(Q||P)} \leq \sum_{I\in T_Q^n} P^{\otimes n}(I) \leq 2^{-nD(Q||P)}. \tag{1.45}$$

Again, we may take the limit, to obtain

Proposition 1.2.6. *Let $Q \in \mathcal{P}_n$, then*

$$\lim_{n\to\infty} \frac{1}{n} \log \sum_{I \in T_Q^n} P^{\otimes n}(I) = -D(Q||P). \tag{1.46}$$

Chapter 2

Tensors

In some scientific communities a vector is simply a finite family of elements from some set, which is usually, but not always, a field, while in mathematics, a vector is an element of a vector space. When the underlying set is a field, then the two definitions of course translate to each other after choosing a basis for the vector space. Likewise, a *tensor* is defined somewhat differently depending on who is talking about them. We might think of a 3-tensor as a family of elements indexed over a multiindex $(a_{i,j,k})_{(i,j,k)\in I\times J\times K}$ or as a vector in the tensor product of vector spaces $\psi \in A\otimes B\otimes C$. If $(a_i)_{i\in I}, (b_j)_{j\in J}, (c_k)_{k\in K}$ are bases of A, B, C then $\psi \leftrightarrow \left(\langle a_i \otimes b_j \otimes c_k|\psi\rangle\right)_{i,j,k\in I\times J\times K}$ yields a correspondence between the two notions. In what follows, a tensor will simply mean a vector in a tensor product of vector spaces. Given a 2-tensor $\psi \in A \otimes B$, we may consider $|\psi\rangle$ as a linear map $B^* \ni \langle\phi| \mapsto \langle\phi|\psi\rangle \in A$, and as such it has a rank, which coincides with the Schmidt rank of $|\psi\rangle$ or equivalently the rank of the matrix $[\langle a_i \otimes b_j|\psi\rangle]_{i,j}$. This notion of rank has a natural extension to k-tensors.

2.1. Tensor Rank

Definition 2.1.1. *Let $\psi \in V_1 \otimes \cdots \otimes V_k$. The* tensor rank *of ψ is*

$$R(\psi) = \min\left\{ r \in \mathbb{N} \middle| \exists (\psi_{i,j})_{j=1}^r \subset V_i : \psi = \sum_{j=1}^r \psi_{1,j} \otimes \dots \psi_{k,j} \right\}. \tag{2.1}$$

The integer k will be called the order *of ψ and ψ will be called a k-tensor.*

Tensors of the form $\psi = \psi_1 \otimes \cdots \otimes \psi_k \in V_1 \otimes \cdots \otimes V_k$ are called simple tensors, or tensors of rank 1, and correspond to unnormalized, separable, pure, k-partite quantum states $|\psi\rangle\langle\psi| \in \mathcal{S}(V_1 \otimes \cdots \otimes V_k)$. The rank of ψ is then the smallest number of simple tensors needed to write ψ. While the rank of a tensor is simple enough to define, it is almost always hard to compute. Tensor rank has been studied much in the context of algebraic complexity theory, because of its relation to the multiplicative complexity of the corresponding forms: A tensor $\psi \in A \otimes B \otimes C$ may for instance be viewed as a trilinear form

$$A^* \times B^* \times C^* \ni \langle\psi_A| \times \langle\psi_B| \times \langle\psi_C| \mapsto \langle\psi_A \otimes \psi_B \otimes \psi_C|\psi\rangle \in \mathbb{C}$$

or as a bilinear map

$$A^* \times B^* \ni \langle\psi_A| \times \langle\psi_B| \times \mapsto \langle\psi_A \otimes \psi_B|\psi\rangle \in C.$$

A rank decomposition of ψ for $k = 3$ encodes a method for computing these operations, which is what makes the notion relevant in algebraic complexity theory. In particular, one might consider the 2×2 matrix multiplication tensor

$$|\mathrm{MaMu}_2\rangle = \sum_{i,j,k=1}^{2} |ij\rangle_A \otimes |jk\rangle_B \otimes |ik\rangle_C \in (\mathbb{C}^2 \odot \mathbb{C}^2)_A \otimes (\mathbb{C}^2 \odot \mathbb{C}^2)_B \otimes (\mathbb{C}^2 \odot \mathbb{C}^2)_C. \tag{2.2}$$

MaMu_2 is related to the complexity of performing matrix multiplication (see e.g. [21] [22] [23]), in the following sense: When $\langle\psi_A| = \sum_{ij} a_{ij} \langle ij|_A$ and $\langle\psi_B| = \sum_{jk} b_{jk} \langle jk|_B$ the map

$$\langle\psi_A| \otimes \langle\psi_B| \mapsto \langle\psi_A \otimes \psi_B|\mathrm{MaMu}_2\rangle \in C, \tag{2.3}$$

outputs $\sum_{k,i} c_{ik} |ik\rangle_C$ where the matrix $[c_{ik}]_{ik}$ is the product of the matrices $[a_{ij}]_{ij}$ and $[b_{jk}]_{jk}$.

One very simple and important tensor is the so-called unit k-tensor of rank r:

Definition 2.1.2. *We denote by*

$$u_r^k = \sum_{j=0}^{r-1} |\underbrace{j \cdots j}_{k \text{ times}}\rangle \in \mathbb{C}^r \otimes \cdots \otimes \mathbb{C}^r \tag{2.4}$$

the r-level unit k-tensor. When k is clearly implied from context it will be suppressed; $u_r = u_r^k$ and when $r = 2$ we shall sometimes omit it in notation. When $k = 2$ we use the notation $e_r = u_r^2$. More generally, when $I \subset [k]$, we denote by u_r^I, the unit tensor living on the subsystem $\bigotimes_{i \in I} V_i$:

$$u_r^I = \left[\sum_{j=0}^{r-1} \left(\bigotimes_{i \in I} |j\rangle_{V_i}\right)\right] \otimes \left(\bigotimes_{i \notin I} |0\rangle_{V_i}\right) \in V_1 \otimes \cdots \otimes V_k \tag{2.5}$$

where $V_i = \mathbb{C}^r$ for $i \in I$ and $V_i = \mathbb{C}$ otherwise. Specially, when $i, j \in [k]$, we write $e_r^{i,j} = u_r^{\{i,j\}}$.

Notice that $|u_2^k\rangle = \sqrt{2}\,|\mathrm{GHZ}^k\rangle$ is the unnormalized k-partite Greenberger–Horne–Zeilinger state and generally $|u_r^k\rangle = \sqrt{r}\,|\mathrm{GHZ}_r^k\rangle$ is what is sometimes called the unnormalized r-level GHZ-state. In the special case of $k = 2$, $|e_2\rangle = \sqrt{2}\,|\mathrm{EPR}\rangle$ is the unnormalized Einstein–Podolsky–Rosen pair (or a Bell state, if one pleases). Generally $\frac{1}{\sqrt{2}} |e_2^{i,j}\rangle$ corresponds to an EPR-pair being shared by parties i and j. A notion closely related to the concept of tensor rank is that of restrictions:

Definition 2.1.3. *Let $\psi \in V_1 \otimes \cdots \otimes V_k$ and $\phi \in W_1 \otimes \cdots \otimes W_k$. We say that ψ* restricts *to ϕ if and only if $(X_1 \otimes \cdots \otimes X_k)\psi = \phi$ for some $X_i \in \mathrm{End}(V_i, W_i)$ and we write*

$$\psi \geq \phi.$$

If $\psi \geq \phi$ and $\phi \geq \psi$, we say that the tensors are equivalent and write

$$\psi \sim \phi.$$

Let $\mathcal{T}_k$ denote the set of equivalence classes of k-tensors under this restriction equivalence.

It is immediately clear that $u_r \sim u_s$ if and only if $s = r$ and when $k = 2$ it follows from the singular value decomposition of matrices (or equivalently, the Schmidt decomposition) that any $\psi \in V_1 \otimes V_2$ is equivalent to a unit tensor e_r, where r is the rank of ψ.

Example 2.1.4. *[24] The first tensor one comes upon which is not restriction-equivalent to a unit tensor is the W tensor:*

$$|W\rangle = |001\rangle_{ABC} + |010\rangle_{ABC} + |100\rangle_{ABC} \in \mathbb{C}^2_A \otimes \mathbb{C}^2_B \otimes \mathbb{C}^2_C. \tag{2.6}$$

Here we have labelled each copy of $\mathbb{C}^2$ to distinguish between the spaces. By the way $|W\rangle$ is written above, it has rank at most 3. In fact the rank is exactly 3. To see this, we assume, for the sake of reaching a contradiction, that

$$|W\rangle = |\psi_1\rangle \otimes |\phi_1\rangle \otimes |\eta_1\rangle + |\psi_2\rangle \otimes |\phi_2\rangle \otimes |\eta_2\rangle \in \mathbb{C}^2_A \otimes \mathbb{C}^2_B \otimes \mathbb{C}^2_C. \tag{2.7}$$

Then

$$\langle 0|_A |W\rangle = \langle 0|\psi_1\rangle |\phi_1\rangle \otimes |\eta_1\rangle + \langle 0|\psi_2\rangle |\phi_2\rangle \otimes |\eta_2\rangle = |01\rangle_{BC} + |10\rangle_{BC}. \tag{2.8}$$

Since the latter has rank 2, $|\phi_1\rangle$ and $|\phi_2\rangle$ must be linearly independent, as must $|\eta_1\rangle$ and $|\eta_2\rangle$. But

$$\langle 1|_A |W\rangle = \langle 1|\psi_1\rangle |\phi_1\rangle \otimes |\eta_1\rangle + \langle 1|\psi_2\rangle |\phi_2\rangle \otimes |\eta_2\rangle = |00\rangle_{BC} \tag{2.9}$$

is rank 1. This implies that either $\langle 1|\psi_1\rangle$ or $\langle 1|\psi_2\rangle$ is 0. Without loss of generality we assume that $|\psi_1\rangle = |0\rangle$ and so $|\phi_2\rangle \otimes |\eta_2\rangle$ must be a multiple of $|00\rangle_{BC}$, but this makes (2.8) impossible as $|01\rangle_{BC} + |10\rangle_{BC} - z\,|00\rangle_{BC}$ has rank 2 for any $z \in \mathbb{C}$ and $|\phi_1\rangle \otimes |\eta_1\rangle$ has rank 1. So

$$R(W) = 3. \tag{2.10}$$

Arguably, this is the easiest non-trivial rank computation for a tensor. The technique used is a case of what is called the substitution method (see e.g. [25, Proposition 3.1]), which is a crude way to gain lower bounds on the rank of a tensor. To see that W is not equivalent to a unit *tensor, note that since rank is stable under equivalence, if W is equivalent to any unit tensor, it must be u_3. But $W \not\geq u_3$ since as a linear map from A^* to $B \otimes C$, $|W\rangle$ has rank 2 and $|u_3\rangle$ has rank 3, and a restriction will necessarily lower the rank of a linear map.*

Note that by Proposition 1.1.16, $\psi \geq \phi \iff |\psi\rangle\langle\psi| \xrightarrow{\text{SLOCC}} |\phi\rangle\langle\phi|$. In the resource theory of SLOCC, normalizing the states is of no matter, since $|\psi\rangle\langle\psi| \xrightarrow{\text{SLOCC}} |\phi\rangle\langle\phi|$ if and only if the same is true when multiplying by any positive real on either side. In other words $\mathcal{T}_k$ is invariant under rescaling, i.e. $\mathbb{R}^+\psi \subset [\psi]_\sim \in \mathcal{T}_k$. Now that it is clear how the resource theory of pure quantum states under SLOCC is the same as the resource theory of tensors under restriction, let us see how they both relate to tensor rank

Proposition 2.1.5. *Let $\psi \in V_1 \otimes \cdots \otimes V_k$ be a k-tensor. Then*

$$R(\psi) = \min\{r \in \mathbb{N} | u_r \geq \psi\}. \tag{2.11}$$

Proof. Let

$$\psi = \sum_{j=1}^{r} \psi_{1,j} \otimes \cdots \otimes \psi_{k,j}, \tag{2.12}$$

be a rank decomposition of ψ. Let $X_i = \sum_{j=1}^{r} |\psi_{i,j}\rangle\langle j|$, then

$$(X_1 \otimes \cdots \otimes X_k)|u_r\rangle = |\psi\rangle, \tag{2.13}$$

showing that $R(\psi) \geq \min\{r \in \mathbb{N} | u_r \geq \psi\}$.

Conversely, suppose that $(X_1 \otimes \ldots \otimes X_k)|u_r\rangle = |\psi\rangle$, for some $X_i \in \mathrm{End}(\mathbb{C}^r, V_i)$. Then

$$|\psi\rangle = \sum_{j=1}^{r} X_1|j\rangle \otimes \cdots \otimes X_k|j\rangle, \tag{2.14}$$

is a rank decomposition of $|\psi\rangle$, showing that $R(\psi) \leq \min\{r \in \mathbb{N} | u_r \geq \psi\}$. ☐

Inspired by Proposition 2.1.5 we define the sub-rank

$$R_{\text{sub}}(\psi) = \max\{r \in \mathbb{N} | \psi \geq u_r\}. \tag{2.15}$$

By Proposition 2.1.5, the tensor rank of a tensor ψ is simply the minimal size of GHZ-state needed to extract $|\psi\rangle$ in the resource theory of k-partite quantum states under SLOCC, while the sub-rank by definition is the maximal GHZ-state extractable from $|\psi\rangle$. Given two tensors $\psi \in V_1 \otimes \cdots \otimes V_k$, $\phi \in W_1 \otimes \cdots \otimes W_k$, we might consider the tensor product $\psi \otimes \phi$ and ask what the rank is or which tensors it restricts to or from. However, here we run into the same ambiguity as discussed in Section 1.1.1. Are we considering $|\psi\rangle \otimes |\phi\rangle$ as a k-tensor, or are we thinking of it as a $2k$-tensor? That is, does it live in

$$(V_1 \otimes W_1) \otimes \ldots \otimes (V_k \otimes W_k)\,. \tag{2.16}$$

or in

$$V_1 \otimes \ldots \otimes V_k \otimes W_1 \otimes \ldots \otimes W_k \tag{2.17}$$

The two are of course isomorphic as vector spaces, through a canonical isomorphism, but the notion of "simple tensor" is different, as is the notion of restriction and the notion of rank. From the perspective of k-partite quantum states, the distinction is whether we think of $|\psi\rangle \otimes |\phi\rangle$ as k parties sharing two entangled states, or as k parties sharing a state $|\psi\rangle$ and k other parties sharing a state $|\phi\rangle$. To deal with this ambiguity we shall use $\odot$ and $\otimes$ in the same manner as in Section 1.1.1.

$$\psi \odot \phi \in (V_1 \otimes W_1) \otimes \cdots \otimes (V_k \otimes W_k)\,. \tag{2.18}$$

$$\psi \otimes \phi \in V_1 \otimes \cdots \otimes V_k \otimes W_1 \otimes \cdots \otimes W_k. \tag{2.19}$$

Since there are more simple tensors in (2.16) than in (2.17), $R(\psi \otimes \phi) \geq R(\psi \odot \phi)$. When $\psi = \sum_{j=1}^{R(\psi)} \psi_j$ and $\phi = \sum_{j=1}^{R(\phi)} \phi_j$ are k-tensors, written as a sum of simple tensors, ψ_j and ϕ_j, then $\psi \otimes \phi = \sum_{j_1=1}^{R(\psi)} \sum_{j_2=1}^{R(\phi)} \psi_{j_1} \otimes \phi_{j_2}$ is a sum of simple tensors, so

$$R\left(\psi\right)R(\phi) \geq R\left(\psi \otimes \phi\right) \geq R\left(\psi \odot \phi\right). \tag{2.20}$$

And when both systems are binary ($k = 2$), the above inequalities are in fact equalities, since it follows from multiplicativity of matrix rank under Kronecker product, that

$$R(\psi)R(\phi) \geq R(\psi \otimes \phi) \geq R(\psi \odot \phi) = R(\psi)R(\phi). \tag{2.21}$$

It is well-known that for $k \geq 3$ there is often strict inequality between $R\left(\psi\right)R(\phi)$ and $R\left(\psi \odot \phi\right)$, as we shall also see in examples below. In fact, as was shown in [11] and will be shown again below, the first inequality in Eq. (2.20) can also be strict. The fact that rank is sometimes strictly sub-multiplicative with respect to $\odot$ and $\otimes$ makes it worth considering the asymptotic rank:

$$R^{\odot}(\phi) \stackrel{def}{=} \lim_{n\to\infty} R\left(\phi^{\odot n}\right)^{1/n}, \tag{2.22}$$

and the less studied non-flattening asymptotic rank

$$R^{\otimes}(\phi) \stackrel{def}{=} \lim_{n\to\infty} R\left(\phi^{\otimes n}\right)^{1/n}. \tag{2.23}$$

Similarly we define the asymptotic sub-rank

$$R_{\mathrm{sub}}^{\odot}(\phi) \stackrel{def}{=} \lim_{n\to\infty} R_{\mathrm{sub}}\left(\phi^{\odot n}\right)^{1/n}. \tag{2.24}$$

We shall now see that the asymptotic sub-rank of a tensor is 0 if and only if it is separable over some bipartition, which was also shown in [26]. Given a bipartition $[k] = I_1 \sqcup I_2$, we say that $\psi \in V_1 \otimes \cdots \otimes V_k$ is separable over the I_1-I_2 bipartition if $\psi = \psi_1 \otimes \psi_2$ for some $\psi_1 \in \bigotimes_{i\in I_1} V_i$ and $\psi_2 \in \bigotimes_{i\in I_2} V_i$. Clearly being separable over some bipartition means that the asymptotic subrank is 0, since separability is preserved by both Kronecker product and restriction. To see the converse, we first show that we can extract EPR-pairs between each pair of parties.

Lemma 2.1.6 (slighty stronger version of [26, Lemma 4])**.** *Let $i_1, i_2 \in [k]$ and let $\psi \in V_1\otimes\cdots\otimes V_k$ be such that ψ is not separable over any partition $[k] = I_1 \sqcup I_2$ with $i_1 \in I_1$ and $i_2 \in I_2$. Then $\psi \geq e^{i_1,i_2}$*

Proof. Let i_0 be any $i_0 \in [k]$ different from i_1 and i_2. Now let $I_1 \sqcup I_2 = [k]\backslash\{i_0\}$ with $i_1 \in I_1$ and $i_2 \in I_2$. Let $\mathrm{Sep}_{I_1,I_2} \subset \bigotimes_{i\neq i_0} V_i$ be the tensors which are separable across the I_1-I_2 partition and let $A_{I_1,I_2} \subset V_{i_0}^*$ be the subset characterized by

$$A_{I_1,I_2} = \left\{ \langle\psi_{i_0}| \in V_{i_0}^* \middle| \langle\psi_{i_0}|\psi\rangle \in \mathrm{Sep}_{I_1,I_2} \right\}. \tag{2.25}$$

A_{I_1,I_2} cannot be all of $V_{i_0}^*$, since this would imply separability of ψ. Since Sep_{I_1,I_2} is an algebraic set, so is A_{I_1,I_2}. Now $\bigcup A_{I_1,I_2}$, where the union is taken over all partitions $I_1 \sqcup I_2 = [k]\backslash\{i_0\}$ with $i_1 \in I_1$ and $i_2 \in I_2$, is a finite union of proper algebraic subsets and therefore not all of $V_{i_0}^*$. So for some $\langle\psi_{i_0}| \in V_{i_0}^*$, $\langle\psi_{i_0}|\psi\rangle$ is non-separable over all partitions $I_1 \sqcup I_2 = [k]\backslash\{i_0\}$ with $i_1 \in I_1$ and $i_2 \in I_2$. Iterating this process we end up with a restriction $X = \bigotimes_{i\neq i_1,i_2} \langle\psi_i|$ such that $X|\psi\rangle \in V_{i_1} \otimes V_{i_2}$ is non-separable and therefore equivalent to $e_r^{i_1,i_2}$ for some $r \geq 2$ which restricts to e^{i_1,i_2}. □

We say that a tensor ψ is globally entangled if ψ is non-separable over any bipartition. We can now show that it is possible to extract a unit tensor, by applying what is basically a simplified stochastic version of quantum teleportation.

Proposition 2.1.7. *Let $\psi \in V_1 \otimes \cdots \otimes V_k$ be globally entangled. Then $\psi^{\odot k-1} \geq u_2$.*

Proof. By the previous lemma $\psi^{\odot k-1} \geq \bigodot_{i=2}^k e^{1,i}$.

$$\bigodot_{i=2}^{k} |e^{1,i}\rangle = \sum_{j_2\ldots j_k=0}^{2} |j_2 \ldots j_k\rangle_1 \otimes |j_2\rangle_2 \otimes \ldots \otimes |j_k\rangle_k. \tag{2.26}$$

Applying the restriction $|0\rangle\langle 0\ldots 0|_1 + |1\rangle\langle 1\ldots 1|_1$ to the first system results in a tensor equivalent to u_2. □

Definition 2.1.8. *Given two k-tensors ψ and ϕ, we define the asymptotic restriction cost of ϕ with respect to ψ as*

$$\omega(\psi, \phi) = \inf \left\{ \tau \in \mathbb{R}^+ \middle| \psi^{\odot \lfloor \tau n \rfloor} \geq \phi^{\odot n} \textit{ for } n \gg 1 \right\}. \tag{2.27}$$

Since restriction of tensors correspond to SLOCC conversion of pure quantum states, we may define the corresponding asymptotic notion for general quantum states. That is, the asymptotic extraction rate of one resource, σ, from another, ρ:

$$E_{\text{SLOCC}}(\rho, \sigma) = \sup \left\{ \tau \in \mathbb{R}^+ \middle| \rho^{\odot n} \xrightarrow{\text{SLOCC}} \sigma^{\odot \lfloor \tau n \rfloor} \text{ for } n \gg 1 \right\}. \tag{2.28}$$

Or we may conversely consider the cost of σ in terms of ρ

$$C_{\text{SLOCC}}(\rho, \sigma) = \frac{1}{E_{\text{SLOCC}}(\rho, \sigma)} = \inf \left\{ \tau \in \mathbb{R}^+ \middle| \rho^{\odot \lfloor \tau n \rfloor} \xrightarrow{\text{SLOCC}} \sigma^{\odot n} \text{ for } n \gg 1 \right\}. \tag{2.29}$$

In light of Proposition 2.1.5,

$$\omega(\psi, \phi) = C_{\text{SLOCC}}\left(|\psi\rangle\langle\psi| , |\phi\rangle\langle\phi| \right). \tag{2.30}$$

When the resource is $\psi = u_2$, the entanglement cost of a pure state $|\psi\rangle$ is closely related to the concept of asymptotic rank;

$$\omega(\phi) \overset{def}{=} \omega(u_2, \phi) = \log R^{\odot}(\phi). \tag{2.31}$$

And like sub-rank, we also consider

$$\omega_{\text{sub}}(\phi) \overset{def}{=} \frac{1}{\omega(\phi, u_2)} = \log R^{\odot}_{\text{sub}}(\phi). \tag{2.32}$$

As previously mentioned, one of the most studied tensors is MaMu_2. Especially determining the value of $\omega = \omega(\text{MaMu}_2)$ has been the subject of much research focus. Currently the best known bounds are $2 \leq \omega \leq 2.3729$ [27], or equivalently, that $4 \leq R^{\odot}(\text{MaMu}_2) \leq 2^{2.3729} \approx 5.18$. Since it is known that $R(\text{MaMu}_2) = 7$ [28], we can conclude that $R(\text{MaMu}^{\odot n}) < R(\text{MaMu})^n$ for sufficiently large n, giving an example that the left-hand-side and right-hand-side of Eq. (2.20) are not always equal, and indeed that $R^{\odot}$ and R are not the same thing.

Another notion of rank that has been used to study $R^{\odot}$ is the border rank, $\underline{R}$, which shall now be introduced.

Border rank and degenerations

A notion of rank for tensors, which is popular in the field of algebraic geometry, is the border rank. Border rank is in a sense the closure of tensor rank. The border rank is upper bounded by tensor rank and often coincides with tensor rank.

Definition 2.1.9. *Given $\psi \in V_1 \otimes \cdots \otimes V_k$, the border rank of ψ is defined as*

$$\underline{R}(\psi) = \min \left\{ r \in \mathbb{N} \middle| \exists \psi_n \subset V_1 \otimes \cdots \otimes V_k \text{ such that } \lim_n \psi_n = \psi \text{ and } \forall n : R(\psi_n) \leq r \right\}. \quad (2.33)$$

When $k = 2$, tensor rank coincides with the rank of the corresponding linear map and since the matrices of rank at most r form a closed set, tensor rank and border rank coincide for 2-tensors.

Example 2.1.10. *The W tensor is the simplest example of a tensor where the rank and border rank differ [29, Sec. 2]. Let*

$$|\psi_\varepsilon\rangle = (|0\rangle + \varepsilon |1\rangle)^{\otimes 3} - |000\rangle = \varepsilon |W\rangle + \varepsilon^2 (|011\rangle + |101\rangle + |110\rangle) + \varepsilon^3 |111\rangle. \quad (2.34)$$

Then $R(\psi_\varepsilon) = 2$ and $\frac{1}{\varepsilon}\psi_\varepsilon \to W$ as $\varepsilon \to 0$, showing that $\underline{R}(W) \leq 2$. In fact, it is then not hard to see that $\underline{R}(W) = 2 < 3 = R(W)$, by a flattening argument (see Section 2.3).

Just like tensor rank and restriction of unit tensors relate, so does border rank relate to the concept of degeneration:

Definition 2.1.11. *Let $\psi \in V_1 \otimes \cdots \otimes V_k$ and $\phi \in W_1 \otimes \cdots \otimes W_k$. We write $\psi \trianglerighteq \phi$ if*

$$\phi \in \overline{\{\phi' \in W_1 \otimes \cdots \otimes W_k | \psi \geq \phi'\}}. \quad (2.35)$$

When this is the case we say that ψ degenerates to ϕ.

Note that by the above definitions and Proposition 2.1.5, we have

$$\underline{R}(\psi) = \min \{r \in \mathbb{N} | u_r \trianglerighteq \psi\}. \quad (2.36)$$

A neat thing about degenerations is that the set of tensors that degenerate from ψ form an algebraic variety, which leads to us only having to consider polynomial approximations (see Theorem 2.1.13 below).

Definition 2.1.12. *Let $\psi \in V_1 \otimes \cdots \otimes V_k$ and $\phi \in W_1 \otimes \cdots \otimes W_k$. We say that ψ degenerates to ϕ with* error degree *e and* approximation degree *d, denoted $\psi \trianglerighteq_d^e \phi$, if there exists $f_i(\varepsilon) \in \mathrm{Hom}(V_i, W_i)[\varepsilon]$ such that*

$$(f_1(\varepsilon) \otimes \cdots \otimes f_k(\varepsilon))\, \psi = \varepsilon^d \phi + \varepsilon^{d+1}\phi_1 + \cdots + \varepsilon^{d+e}\phi_e, \quad (2.37)$$

for some $\phi_1, \ldots, \phi_e$. Here $\mathrm{Hom}(V_i, W_i)[\varepsilon]$ is the polynomial ring over $\mathrm{Hom}(V_i, W_i)$, which, given choice of bases, may be thought of as matrices with entries that are polynomial in ε. The map $(f_1(\varepsilon) \otimes \cdots \otimes f_k(\varepsilon))$ is called a degeneration. We write $\psi \trianglerighteq^e \phi$ and $\psi \trianglerighteq_d \phi$, if $\psi \trianglerighteq_d^e \phi$ for any $e, d \in \mathbb{N}$.

In the example with the W tensor above, the error degree, e, was 2 and the approximation degree, d, was 1. The following important theorem is due to Strassen. A proof will not be presented here.

Theorem 2.1.13. *[21, Theorem 5.8] $\psi \trianglerighteq \phi$ if and only if $\psi \trianglerighteq_d^e \phi$ for some $e, d \in \mathbb{N}$.*

Applying the fact that

$$\begin{aligned}&\Big[\left(f_1(\varepsilon) \otimes \cdots \otimes f_k(\varepsilon)\right) \otimes \left(g_1(\varepsilon) \otimes \cdots \otimes g_k(\varepsilon)\right)\Big]\, \psi_1 \otimes \psi_2 \\ = &\Big[\left(f_1(\varepsilon) \otimes \cdots \otimes f_k(\varepsilon)\right) \psi_1\Big] \otimes \Big[\left(g_1(\varepsilon) \otimes \cdots \otimes g_k(\varepsilon)\right) \psi_2\Big],\end{aligned} \tag{2.38}$$

to (2.37) we immediately get:

Proposition 2.1.14. *If $\psi_1 \trianglerighteq_{d_1}^{e_1} \phi_1$ and $\psi_2 \trianglerighteq_{d_2}^{e_2}$, then*

$$\psi_1 \odot \psi_2 \trianglerighteq_{d_1+d_2}^{e_1+e_2} \phi_1 \odot \phi_2 \tag{2.39}$$

and

$$\psi_1 \otimes \psi_2 \trianglerighteq_{d_1+d_2}^{e_1+e_2} \phi_1 \otimes \phi_2. \tag{2.40}$$

The error degree is upper bounded by the approximation degree in the following sense: If $\psi \trianglerighteq_d \phi$ via a degeneration

$$\left(f_1(\varepsilon) \otimes \cdots \otimes f_k(\varepsilon)\right) \psi = \varepsilon^d \phi + o(\varepsilon^{d+1}), \tag{2.41}$$

then by removing all terms in each $f_i(\varepsilon)$ of degree strictly higher than d, (2.41) remains true, and then we have no terms of degree higher than kd on the right-hand-side. Therefore

$$\psi \trianglerighteq_d \phi \implies \psi \trianglerighteq_d^{kd-d}. \tag{2.42}$$

Proposition 2.1.15. *Given k-tensors ψ and ϕ such that $\psi \trianglerighteq^e \phi$, then $\psi \odot u_{e+1} \geq \phi$.*

Proof. By assumption there are $f_i(\varepsilon)$ such that

$$\left(\varepsilon^{-d} f_1(\varepsilon) \otimes \cdots \otimes f_k(\varepsilon)\right) \psi = \phi + \varepsilon\phi_1 + \cdots + \varepsilon^e \phi_e. \tag{2.43}$$

Let $p(\varepsilon) = \phi + \varepsilon\phi_1 + \cdots + \varepsilon^e\phi_e$. Given distinct non-zero $z_0, \ldots, z_e \in \mathbb{C}$, we have by Lagrange interpolation [30] (or e.g. [31])

$$\phi = p(0) = \sum_{j=0}^{e} \lambda_j p(z_j), \tag{2.44}$$

where $\lambda_j = \prod_{i \neq j} \frac{z_i}{z_i - z_j}$. Let $g_1 = \sum_{j=0}^{e} \lambda_j z_j^{-d} f_1(z_j) \odot \langle j|$ and $g_i = \sum_{j=0}^{e} \lambda_j f_i(z_j) \odot \langle j|$. Then

$$\begin{aligned}\left(g_1 \otimes \cdots \otimes g_k\right)\left(|\psi\rangle \odot |u\rangle\right) &= \sum_{j=0}^{n} \lambda_j \left(f_1(z_j) \otimes \cdots \otimes f_k(z_j)\right) |\psi\rangle \\ &= \sum_{j=0}^{e} \lambda_j p(z_j) = p(0) = |\phi\rangle.\end{aligned} \tag{2.45}$$

□

By Eq. (2.36) and Theorem 2.1.13 we always have $u_{\underline{R}(\psi)} \trianglerighteq^e \psi$ for some error degree e.

Corollary 2.1.16.

$$R(\psi^{\otimes n}) \leq \underline{R}(\psi)^n (ne+1), \tag{2.46}$$

where e is the error degree of a degeneration $u_{\underline{R}(\psi)} \trianglerighteq^e \psi$

Proof. By Proposition 2.1.14 $u_{\underline{R}(\psi)}^{\otimes n} \trianglerighteq^{ne} \psi^{\otimes n}$. By Proposition 2.1.15,

$$u_{\underline{R}(\psi)^n (ne+1)} \sim u_{\underline{R}(\psi)}^{\otimes n} \odot u_{ne+1} \geq \psi^{\otimes n}, \tag{2.47}$$

showing that $R(\psi^{\otimes n}) \leq \underline{R}(\psi)^n (ne+1)$. □

By Corollary 2.1.16, the border rank versions of asymptotic rank $\underline{R}^{\odot}(\phi) \overset{def}{=} \lim_{n\to\infty} \underline{R}\left(\phi^{\odot n}\right)^{1/n}$ and $\underline{R}^{\otimes}(\phi) \overset{def}{=} \lim_{n\to\infty} \underline{R}\left(\phi^{\otimes n}\right)^{1/n}$ coincide with the same notions for tensor rank; $\underline{R}^{\odot} = R^{\odot}$, $\underline{R}^{\otimes} = R^{\otimes}$. In other words, if we wish to upper bound asymptotic rank of ψ, it is just as good to find degenerations from u to ψ as to find restrictions. This gives some extra freedom.

Another consequence of Corollary 2.1.16 is the first example of how the first inequality in Eq. (2.20) can be strict:

Corollary 2.1.17. *Given a tensor ψ such that $\underline{R}(\psi) < R(\psi)$, then for large enough n*

$$R(\psi^{\otimes n}) < R(\psi)^n. \tag{2.48}$$

In the concrete case of the W-tensor, there is a degeneration with error degree $e = 2$, so $R(W^{\otimes n}) \leq \underline{R}(W)^n (2n+1) = (2n+1)2^n$. This is smaller than $R(W)^n = 3^n$ for $n \geq 7$, so we get an example of rank non-multiplicativity with $\psi = W^{\otimes j}$ and $\phi = W$ for some $j \leq 6$. In fact, the rank drop happens already at $j = 1$:

Proposition 2.1.18. $R(W^{\otimes 2}) \leq 8 < 9 = R(W)^2$

Proof. For any $z \in \mathbb{C} \neq 0$, let $\sqrt{z}$ be some square root of z. Then

$$|W\rangle + z\,|111\rangle = \frac{1}{2\sqrt{z}}\left(\left(|0\rangle + \sqrt{z}\,|1\rangle\right)^{\otimes 3} - \left(|0\rangle - \sqrt{z}\,|1\rangle\right)^{\otimes 3}\right), \tag{2.49}$$

showing that $R(|W\rangle + z\,|111\rangle) = 2$ for all $z \neq 0$. Now

$$|W\rangle \otimes |W\rangle = (W + |111\rangle)^{\otimes 2} - \left(W + \frac{1}{2}\,|111\rangle\right) \otimes |111\rangle - |111\rangle \otimes \left(W + \frac{1}{2}\,|111\rangle\right). \tag{2.50}$$

The ranks of the three terms on the right-hand-side are, from left to right, at most $2 \cdot 2 = 4$, $2 \cdot 1 = 2$ and $1 \cdot 2 = 2$, respectively, so the entire thing has rank at most $4 + 2 + 2 = 8$. □

As Proposition 2.1.18 shows, we didn't have to look far for an example of non-multiplicativity of tensor rank. In Section 2.2 we shall see that we pretty much couldn't have found a simpler example.

2.2. Multiplicativity for matrix pencils and 2-tensors

The goal of this section is to prove the following proposition.

Proposition 2.2.1. *Let $\psi \in \mathbb{C} \otimes \mathbb{C}^d \otimes \mathbb{C}^d$ and $\phi \in \mathbb{C}^2 \otimes \mathbb{C}^n \otimes \mathbb{C}^m$. Then*

$$R(\phi \odot \psi) = R(\phi \otimes \psi) = R(\phi)R(\psi). \tag{2.51}$$

Remark 2.2.2. *Proposition 2.2.1 shows that the non-multiplicativity example in Proposition 2.1.18 is essentially minimal. Namely, any example of non-multiplicativity of tensor rank under $\otimes$ must either be with a 5-tensor in $(\mathbb{C}^d \otimes \mathbb{C}^d) \otimes (\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \otimes \mathbb{C}^{d_3})$ with $d_1, d_2, d_3 \geq 3$, $d \geq 2$ or in a tensor space of order 6 or more. Whether counterexamples with 5-tensors exist at all is still an open question. Moreover, one can show using Proposition 2.2.1 and the well-known classification of the $\mathrm{GL}_2^{\times 3}$-orbits in $\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ that if $\psi, \phi \in \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ and $\mathrm{R}(\psi \otimes \phi) < \mathrm{R}(\psi)\,\mathrm{R}(\phi)$, then ψ and ϕ are both equivalent to W.*

The elements of $\mathbb{C}^2 \otimes \mathbb{C}^n \otimes \mathbb{C}^m$ are often called matrix pencils. The tensor rank of matrix pencils is completely understood, in the sense that every matrix pencil is equivalent under local isomorphisms to a pencil in canonical form (Theorem 2.2.4), for which the rank is given by a simple formula (Theorem 2.2.6). This formula allows a short proof of Proposition 2.2.1.

We begin with introducing the canonical form for matrix pencils (Theorem 2.2.4). For a proof of Theorem 2.2.4 see [32, Chapter XII].

Definition 2.2.3. *Given $\psi_i \in U \otimes V_i \otimes W_i$, define $\mathrm{Diag}_U(\psi_1, \ldots, \psi_n)$ as the image of $\bigoplus_{i=1}^n \psi_i$ under the natural inclusion $U \otimes \bigoplus_i (V_i \otimes W_i) \to U \otimes (\bigoplus_i V_i) \otimes (\bigoplus_i W_i)$. For $\zeta \in \mathbb{N}$ define the tensor $L_\zeta \in \mathbb{C}^2 \otimes \mathbb{C}^\zeta \otimes \mathbb{C}^{\zeta+1}$ by*

$$\begin{aligned} |L_\zeta\rangle &:= |0\rangle \otimes \sum_{i=0}^{\zeta-1} |ii\rangle \;+\; |1\rangle \otimes \sum_{i=1}^{\zeta-1} |i\rangle\,|i+1\rangle \\ &= |0\rangle \otimes \begin{pmatrix} 1 & & & 0 \\ & 1 & & 0 \\ & & \ddots & \vdots \\ & & 1 & 0 \end{pmatrix} + |1\rangle \otimes \begin{pmatrix} 0 & 1 & & \\ 0 & & 1 & \\ \vdots & & & \ddots \\ 0 & & & & 1 \end{pmatrix} \end{aligned}$$

and for $\eta \in \mathbb{N}$ define the tensor $N_\eta \in \mathbb{C}^2 \otimes \mathbb{C}^{\eta+1} \otimes \mathbb{C}^\eta$ by

$$\begin{aligned} |N_\eta\rangle &:= |0\rangle \otimes \sum_{i=0}^{\eta-1} |ii\rangle \;+\; |1\rangle \otimes \sum_{i=0}^{\eta-1} |i+1\rangle\,|i\rangle \\ &= |0\rangle \otimes \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ 0 & 0 & \cdots & 0 \end{pmatrix} + |1\rangle \otimes \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}. \end{aligned}$$

The matrix notation in above equations is via the association $|ij\rangle \leftrightarrow |i\rangle\langle j|$.

Theorem 2.2.4 (Canonical form)**.** *Let $\psi \in \mathbb{C}^2 \otimes \mathbb{C}^n \otimes \mathbb{C}^m$. There exist invertible linear maps $A \in \mathrm{GL}_2$, $B \in \mathrm{GL}_n$ and $C \in \mathrm{GL}_m$ and natural numbers $\zeta_1, \ldots, \zeta_p, \eta_1, \ldots, \eta_q \in \mathbb{N}$ and an $\ell \times \ell$ Jordan matrix F such that, with $M = |0\rangle \otimes I_\ell + |1\rangle \otimes F$, we have*

$$(A \otimes B \otimes C)\psi = \mathrm{Diag}_{\mathbb{C}^2}(0, L_{\zeta_1}, \ldots, L_{\zeta_p}, N_{\eta_1}, \ldots, N_{\eta_q}, M), \tag{2.52}$$

where the 0 stands for some 0-tensor of appropriate dimensions. The right-hand side of (2.52) is called the canonical form of ψ.

We now have the necessary notation to present the formula for the tensor rank of matrix pencils in canonical form (Theorem 2.2.6). Theorem 2.2.6 is due to Grigoriev [33], JáJá [34] and Teichert [35], see also [36, Theorem 19.4] or [22, Theorem 3.11.1.1].

Definition 2.2.5. *Let F be a Jordan matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$. Let $d(\lambda_i)$ be the number of Jordan blocks in F of size at least two with eigenvalue λ_i. Define $m(F) \coloneqq \max_i d(\lambda_i)$.*

Theorem 2.2.6. *Let $\psi = \mathrm{Diag}_{\mathbb{C}^2}(0, L_{\zeta_1}, \ldots, L_{\zeta_p}, N_{\eta_1}, \ldots, N_{\eta_q}, |0\rangle \otimes I_\ell + |1\rangle \otimes F)$ be a tensor in canonical form as in (2.52). The tensor rank of ψ equals*

$$\mathrm{R}(\psi) = \sum_{i=1}^{p}(\zeta_i + 1) + \sum_{i=1}^{q}(\eta_i + 1) + \ell + m(F).$$

We are now ready to give the short proof of Proposition 2.2.1.

Proof of Proposition 2.2.1. Let $\psi \in \mathbb{C}\otimes\mathbb{C}^d\otimes\mathbb{C}^d$, $\phi \in \mathbb{C}^2\otimes\mathbb{C}^n\otimes\mathbb{C}^m$. By Eq. (2.20) it suffices to show that $R(\psi)R(\phi) = R(\psi\odot\phi)$. We may assume that $|\psi\rangle = 1\otimes\sum_{i=0}^{r-1}|ii\rangle$ with $r = \mathrm{R}(\psi)$. By Theorem 2.2.4 we may assume that ϕ is in canonical form, $\phi = \mathrm{Diag}_{\mathbb{C}^2}(0, L_{\zeta_1}, \ldots, L_{\zeta_\ell}, N_{\varepsilon_1}, \ldots, N_{\varepsilon_k}, M)$. The Kronecker product $\phi \odot \psi$ is isomorphic to

$$\phi \odot \psi \cong \mathrm{Diag}_{\mathbb{C}^2}(\underbrace{\phi, \ldots, \phi}_{r}).$$

By an appropriate local basis transformation we put this in canonical form

$$\psi \odot \phi \cong \mathrm{Diag}_{\mathbb{C}^2}(L_{\zeta_1}^{\oplus r}, \ldots, L_{\zeta_\ell}^{\oplus r}, N_{\varepsilon_1}^{\oplus r}, \ldots, N_{\varepsilon_k}^{\oplus r}, M^{\oplus r}),$$

which by Theorem 2.2.6 has rank $r \cdot \mathrm{R}(\phi) = \mathrm{R}(\psi)\,\mathrm{R}(\phi)$. $\square$

2.3. Non-multiplicativity of border rank

As we saw in Corollary 2.1.17 and Proposition 2.1.18, tensor rank is sometimes strictly sub-multiplicative. A natural question to ask is then, what about border rank? In Corollary 2.1.17 the discrepancy between border rank and tensor rank was the witness of strict sub-multiplicativity, so there does not immediately seem to be a good way to generalize this technique. Furthermore, the W tensor cannot be a counterexample to border rank multiplicativity, as a flattening argument will show.

Flattenings

Given a k-tensor, one might group some of the tensor legs together and view it as a tensor of order $k' < k$. The fact that this can never increase the tensor rank or border rank of a tensor, gives a way of providing lower bounds on tensor and border rank. In particular if $\psi \in V_1 \otimes \cdots \otimes V_k$, we may consider $|\psi\rangle$ as a linear map $V_1^* \otimes \cdots \otimes V_j^* \to V_{j+1} \otimes \cdots \otimes V_k$, corresponding to $k' = 2$. The rank of this linear map then lower bounds both the rank and border rank of ψ.

Example 2.3.1. *Consider the tensor* $W \in \mathbb{C}_A^{\otimes 2} \otimes \mathbb{C}_B^{\otimes 2} \otimes \mathbb{C}_C^{\otimes 2}$ *of Example 2.1.4. As a linear map* $\mathbb{C}_A^{\otimes 2*} \otimes \mathbb{C}_B^{\otimes 2*} \to \mathbb{C}_C^{\otimes 2}$*, this has rank 2, so* $\underline{R}(W) \geq 2$*.*

If $V = V_1 \otimes \cdots \otimes V_k$ a generalized flattening is a linear map $F : V \to \mathrm{Hom}(A, B)$, where A and B are vector spaces. This generalizes the usual flattenings with $k' = 2$.

Proposition 2.3.2. *Let* $F : V_1 \otimes \cdots \otimes V_k \to \mathrm{Hom}(A, B)$ *and*

$$r_0 = \max \left\{ R\left(F(\psi)\right) \middle| R\left(\psi\right) = 1 \right\}. \tag{2.53}$$

Then for all $T \in V_1 \otimes \cdots \otimes V_k$*,*

$$\underline{R}\left(T\right) \geq \frac{R\left(F(T)\right)}{r_0} \quad \forall T \in V. \tag{2.54}$$

Proof. Let $T \in V_1 \otimes \cdots \otimes V_k$ with $\underline{R}(T) = r$ and let $(T_i)_{i\in\mathbb{N}}$ be a sequence of tensors with $R(T_i) \leq r$ converging to T. Since $F\left(\psi\right) \leq r_0$ for all simple tensors, ψ, we have $R\left(F(T_i)\right) \leq r_0 R(T_i) \leq r_0 r$. And since $F(T_i) \overset{i\to\infty}{\to} F(T)$, it follows that $R\left(F(T)\right) = \underline{R}\left(F(T)\right) \leq \max_{i\in\mathbb{N}} R\left(F(T_i)\right) \leq r_0 r$, from which Eq. (2.54) follows. □

One might also consider flattenings, F, that map into higher order tensor spaces, but this is rarely done, since the merit of the flattening technique is precisely that ranks are easy to compute for linear maps, which is why one wants to reduce a problem to that of linear maps.

If $F_1 : V_1 \otimes \cdots \otimes V_k \to \mathrm{Hom}(A_1, B_1)$ and $F_2 : W_1 \otimes \cdots \otimes W_k \to \mathrm{Hom}(A_2, B_2)$, then $F_1 \otimes F_2$ maps into $\mathrm{Hom}(A_1, B_1) \otimes \mathrm{Hom}(A_2, B_2)$. Let r_1 and r_2 be the maximum values of Eq. (2.53) for F_1 and F_2 respectively. It follows from Eq. (2.21) that $R\left((F_1 \otimes F_2)(T_1 \otimes T_2)\right) = R\left(F_1(T_1)\right) R\left(F_2(T_2)\right)$ and that

$$\max \left\{ R\left((F_1 \otimes F_2)(\psi)\right) \middle| R\left(\psi\right) = 1 \right\} = r_1 r_2, \tag{2.55}$$

such that Proposition 2.3.2 applied to $F_1 \otimes F_2$ gives

$$\underline{R}\left(T_1 \otimes T_2\right) \geq \frac{R\left(F_1(T_1)\right) R\left(F_2(T_2)\right)}{r_1 r_2}. \tag{2.56}$$

In this sense lower bounds provided by generalized flattenings are multiplicative. As a consequence, if a flattening provides a lower bound r to the border rank of T, then $R^{\otimes}(T)$ is also lower bounded by r.

Example 2.3.3. *We saw in Example 2.3.1 how a simple flattening yielded a lower bound $\underline{R}(W) \geq 2$. Since flattening lower bounds are multiplicative $\underline{R}\left(W^{\otimes n}\right) \geq 2^n$. Since we saw in Example 2.1.4 that $\underline{R}(W) \leq 2$, we conclude that $\underline{R}\left(W^{\otimes n}\right) = \underline{R}\left(W^{\odot n}\right) = 2^n$.*

Counter example for border rank

Even if the W tensor cannot produce a counterexample to border rank multiplicativity, Proposition 2.1.18 still provides a technique that can work for proving non-multiplicativity of border rank for other tensors. Note that Eq. (2.50) generalizes as

$$T^{\otimes 2} = (T+\psi)^{\otimes 2} - \left(T + \frac{1}{2}\psi\right) \otimes \psi - \psi \otimes \left(T + \frac{1}{2}\psi\right), \tag{2.57}$$

for any tensors T and ψ. If we can find T and ψ such that ψ is a simple tensor and $\underline{R}\,(T+\psi) = \underline{R}\left(T + \frac{1}{2}\psi\right) = \underline{R}\,(T) - 1$, then the border rank of Eq. (2.57) is upper bounded by

$$\left(\underline{R}\,(T) - 1\right) + \left(\underline{R}\,(T) - 1\right) + \left(\underline{R}\,(T) - 1\right) = \underline{R}\,(T)^2 - 1, \tag{2.58}$$

meaning that $\underline{R}\left(T^{\otimes 2}\right) \leq \underline{R}\,(T)^2 - 1 < \underline{R}\,(T)^2$, providing a counterexample to border rank multiplicativity. In fact we may look for something even more general. Consider the line

$$L_{T,\psi} = \{T + z\psi | z \in \mathbb{C}\}. \tag{2.59}$$

Since the tensors of border rank at most $n-1$, commonly denoted σ_{n-1}, forms an algebraic variety [22], $L_{T,\psi}$ is either entirely contained in σ_{n-1} or intersects with σ_{n-1} at finitely many points. If the intersection is finite and consists of at least two points, then we get an example as in Eq. (2.57). This can be done by rescaling ψ and leting $T' = T + z\psi$ be of border rank n with $T' + \psi$ and $T' + \frac{1}{2}\psi$ of border rank $n-1$. Proposition 2.3.4 presents such an example.

Proposition 2.3.4. *Consider the tensors $T, \psi \in A \otimes B \otimes C$, with $A = B = C = \mathbb{C}^3$, given by*

$$\begin{aligned} |T\rangle = &|000\rangle_{ABC} + |111\rangle_{ABC} + |222\rangle_{ABC} + \left(|0\rangle + |1\rangle + |2\rangle\right)^{\otimes 3} \\ &+ 2\left(|0\rangle_A + |1\rangle_A\right)\left(|0\rangle_B + |2\rangle_B\right)\left(|1\rangle_C + |2\rangle_C\right), \end{aligned} \tag{2.60}$$

$$|\psi\rangle = \left(|0\rangle_A + |1\rangle_A\right)\left(|0\rangle_B + |2\rangle_B\right)\left(|1\rangle_C + |2\rangle_C\right). \tag{2.61}$$

Then $\underline{R}\,(T) = 5$ and $\underline{R}\,(T-\psi) = \underline{R}\,(T-2\psi) = 4$, implying that

$$\underline{R}\left(T^{\otimes 2}\right) \leq 24 < 25 = \underline{R}\,(T)^2. \tag{2.62}$$

Proof. $\underline{R}(T) \leq 5$ and $\underline{R}(T - 2\psi) \leq 4$ by the decomposition Eq. (2.60). There are two things to show. First that $\underline{R}(T) \geq 5$, which is done by a flattening, and that $\underline{R}(T - \psi) \leq 4$, which is done by simply providing the decomposition:

$$\begin{aligned} |T\rangle - |\psi\rangle =& 2\left(|0\rangle + |1\rangle + \frac{1}{2}|2\rangle\right)\left(|0\rangle + \frac{1}{2}|1\rangle + |2\rangle\right)\left(\frac{1}{2}|0\rangle + |1\rangle + |2\rangle\right) \\ &+ \left(|1\rangle + \frac{1}{2}|2\rangle\right)|1\rangle\left(\frac{1}{2}|0\rangle + |1\rangle\right) \\ &+ \left(|0\rangle + \frac{1}{2}|2\rangle\right)\left(|0\rangle + \frac{1}{2}|1\rangle\right)|0\rangle \\ &+ |2\rangle\left(\frac{1}{2}|1\rangle + |2\rangle\right)\left(\frac{1}{2}|0\rangle + |2\rangle\right). \end{aligned} \tag{2.63}$$

Now consider the flattening $F : A \otimes B \otimes C \to \mathrm{Hom}\left(A \otimes B^*, \left(\wedge^2 A\right) \otimes C\right)$ given by

$$F(\phi) : |\psi_1\rangle_A \otimes \langle\psi_2|_B \mapsto |\psi_1\rangle_A \wedge \langle\psi_2|_B |\phi\rangle . \tag{2.64}$$

Written differently, $F(\phi)$ is a composition of the maps $A \otimes B^* \overset{\mathrm{Id}_A \otimes |\psi\rangle}{\to} A \otimes A \otimes C \overset{\pi_{\wedge^2 A} \otimes \mathrm{Id}_C}{\to} \left(\wedge^2 A\right) \otimes C$. If $|\phi\rangle = |\phi_1\rangle_A \otimes |\phi_2\rangle_B \otimes |\phi_3\rangle_C$, then

$$F(\phi) : |\psi_1\rangle_A \otimes \langle\psi_2|_B \mapsto \langle\psi_2|\phi_2\rangle_B |\phi_1\rangle_A \wedge |\psi_1\rangle_A \otimes |\phi_3\rangle_C , \tag{2.65}$$

has rank at most 2, since the image of $F(\phi)$ is contained in the 2-dimensional space $\{|\phi_1\rangle_A \wedge |\psi\rangle_A \otimes |\phi_3\rangle_C | \psi \in A\}$. So the r_0 of Eq. (2.54) is 2. Now inserting Eq. (2.60) into Eq. (2.64) yields

$$\begin{aligned} F(T)|0\rangle_A \otimes \langle 0|_B = |0\rangle_A \wedge \Big[& |00\rangle_{AC} + (|0\rangle + |1\rangle + |2\rangle)_A (|0\rangle + |1\rangle + |2\rangle)_C \\ & + 2(|0\rangle + |1\rangle)_A (|1\rangle + |2\rangle)_C \Big] \\ = & |0 \wedge 1\rangle_A (|0\rangle + 3|1\rangle + 3|2\rangle)_C + |0 \wedge 2\rangle_A (|0\rangle + |1\rangle + |2\rangle)_C . \end{aligned} \tag{2.66}$$

Similarly we calculate $F(T)|i\rangle_A \otimes \langle j|_B$ and express it in the basis $(|i \wedge j\rangle_A \otimes |k\rangle_C)_{i<j,k}$. This results in the following matrix, where the first column was calculated in Eq. (2.66) above:

$$\begin{bmatrix} 1 & 1 & 1 & -2 & -1 & -1 & 0 & 0 & 0 \\ 3 & 2 & 3 & -3 & -1 & -3 & 0 & 0 & 0 \\ 3 & 1 & 3 & -3 & -1 & -3 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & -2 & -1 & -1 \\ 1 & 1 & 1 & 0 & 0 & 0 & -3 & -1 & -3 \\ 1 & 1 & 2 & 0 & 0 & 0 & -3 & -1 & -3 \\ 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 & 1 & -3 & -2 & -3 \\ 0 & 0 & 0 & 1 & 1 & 1 & -3 & -1 & -3 \end{bmatrix}, \tag{2.67}$$

which has full rank, 9. By Eq. (2.54) $\underline{R}(T) \geq \frac{9}{2} > 4$ as wanted. By Eq. (2.58)

$$\underline{R}\left(T^{\otimes 2}\right) \leq 24 < 25 = \underline{R}(T)^2 . \tag{2.68}$$

$\square$

By now we have seen examples that both tensor rank and border rank can be strictly sub-multiplicative under the tensor product. The fact that tensor rank can be strictly sub-multiplicative under the non-flattened tensor product implies that $R^{\otimes}$ and R are not the same thing. The fact that border rank can also be strictly sub-multiplicative, implies that the second inequality below, can be strict.

$$R(\psi) \geq \underline{R}(\psi) \geq R^{\otimes}(\psi) \geq R^{\odot}(\psi). \tag{2.69}$$

So some of the difference between R and $R^{\odot}$ does not come the flattening of tensor legs. A natural question to ask now might then be if the last inequality above can be strict, or if in fact $R^{\otimes}$ and $R^{\odot}$ are the same. This is also not the case. As previously mentioned, the best known upper bound on $R^{\odot}(\mathrm{MaMu}_2)$ is $2^{2.3729} \approx 5.18$. Yet $R^{\otimes}(\mathrm{MaMu}_2) \geq 6$:

Proposition 2.3.5. $R^{\otimes}(\mathrm{MaMu}_2) \geq 6$

Proof. We prove this lower bound by a Young flattening, as in Proposition 2.3.4. By applying the map $I_A + |00\rangle\langle 11|_A - |11\rangle\langle 11|_A$, to the first tensor leg, we obtain a restriction.

$$\begin{aligned}
|\mathrm{MaMu}_2\rangle &= \sum_{i,j,k=0}^{1} |ij\rangle_A |jk\rangle_B |ki\rangle_C \\
&\geq |00\rangle_A \left(|00\rangle_B |00\rangle_C + |10\rangle_B |01\rangle_C + |11\rangle_B |11\rangle_C + |01\rangle_B |10\rangle_C\right) \\
&\quad + |10\rangle_A \left(|01\rangle_B |11\rangle_C + |00\rangle_B |01\rangle_C\right) \\
&\quad + |01\rangle_A \left(|10\rangle_B |00\rangle_C + |11\rangle_B |10\rangle_C\right) \in \mathbb{C}^3 \otimes \mathbb{C}^4 \otimes \mathbb{C}^4.
\end{aligned} \tag{2.70}$$

Let us call the tensor on the right-hand-side above T. Like in the proof of Proposition 2.3.4, we consider the flattening $F : A \otimes B \otimes C \to \mathrm{Hom}\left(A \otimes B^*, \bigwedge^2 A \otimes C\right)$ given by

$$F(\phi) : |\psi_1\rangle_A \otimes \langle\psi_2|_B \mapsto |\psi_1\rangle_A \wedge \langle\psi_2|_B |\phi\rangle . \tag{2.71}$$

The image of a simple tensor under the flattening F, has at rank at most 2, so

$$\underline{R}(T) \geq \frac{R(F(T))}{2}. \tag{2.72}$$

Expressing the linear map $F(T)$ in the basis $\left(|ij\rangle_A \langle kl|_B\right)_{\{i,j,k,l \in \{0,1\} | (i,j) \neq (1,1)\}}$, for the domain and $\left(|ij \wedge kl\rangle_A |nm\rangle_C\right)_{\{i,j,k,l,n,m \in \{0,1\} | (i,j) \neq (1,1), (k,l) \neq (1,1)\}}$ for the co-domain, both ordered lexicographically, we get the matrix:

$$\begin{bmatrix} 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{2.73}$$

which has rank 12. So $\underline{R}(T) \geq 6$. As discussed previously, and expressed in Eq. (2.56), flattening lower bounds are multiplicative under tensor product, so $R^{\otimes}(\mathrm{MaMu}_2) \geq 6$. $\square$

The exact value of $R^{\otimes}(\mathrm{MaMu}_2)$ is unknown, but it must necessarily be in the interval $[6, 7]$.

Chapter 3

The asymptotic spectrum of tensors

Tensor rank, border rank and their asymptotic versions, introduced in Chapter 2, constitute what could reasonably be called entanglement monotones for pure states, in the sense that they decrease under application of any SLOCC channel, and therefore specially any LOCC channel. However there are certain properties that are often considered desirable for entanglement monotones. One is a sense of continuity, which shall not be touched upon in this thesis. Another is additivity under tensor product. That is, for states ρ and σ we wish for a monotone to satisfy $f(\rho \odot \sigma) = f(\rho) + f(\sigma)$. In other words, the combined entanglement of two resources should be the sum of the entanglement of the parts. These kinds of entanglement monotones can necessarily not distinguish between asymptotic and single-shot conversions of resources. In the case of pure states under SLOCC (equivalently tensors under restriction), it was shown by V. Strassen that in fact the asymptotics are entirely determined by a subset of the additive entanglement monotones, namely by what Strassen called the asymptotic spectrum of tensors, which shall now be introduced. Note that in what follows we are considering monotones which are *multiplicative* under tensor product: $f(\rho \odot \sigma) = f(\rho)f(\sigma)$, but this is of course equivalent to $\log f$ being additive. Considering $\log f(\psi)$ rather than $f(\psi)$ is the same difference as considering $\omega(\psi)$ rather than $R^{\odot}(\psi)$.

3.1. The asymptotic spectrum of a preordered semiring

In [15], Strassen considers the semiring $(\mathcal{T}_k, \odot, \oplus, \geq)$ of equivalence classes of k-tensors under mutual restriction, see Definition 2.1.3. This is a semiring with respect to direct sum and Kronecker product and the partial order, given by restriction, respects the algebraic structure of this semiring. By applying the spectral theorem [15, Theorem 2.3] (here Theorem 3.1.4 below), one gets the asymptotic spectrum $\Delta(\mathcal{T}_k)$ of tensors.

Definition 3.1.1. *A commutative semiring* $(\mathcal{S}, +, \cdot)$ *is a set* $\mathcal{S}$ *with two binary, commutative, and associative operations* $(+, \cdot)$ *containing distinct additive and multiplicative identity elements* $0, 1 \in \mathcal{S}$*, satisfying the distributive law:*

$$a(b + c) = ab + ac. \tag{3.1}$$

Note that what distinguishes a semiring from a ring, is that there is no guarantee of an additive inverse. In fact the semiring we will consider in this chapter has no additive inverses, except for 0, which is always its own additive inverse. In this thesis all semirings are commutative and semiring shall therefore be understood to implicitly mean commutative semiring.

Definition 3.1.2. *A preorder* $\leq$ *on* $\mathcal{S}$ *is a binary relation which is transitive and reflexive (but not necessarily antisymmetric). We say that* $(\mathcal{S}, +, \cdot, \leq)$ *is a preordered semiring, if* $\leq$ *respects the algebraic structure on* $\mathcal{S}$*. That is, when* $a \leq b$ *and* $c \leq d$*:*

$$a + c \leq b + d \tag{3.2}$$

$$ac \leq bd. \tag{3.3}$$

Remark 3.1.3. *Note that in order to show conditions (3.2) and (3.3) it suffices to show* $a + c \leq b + c$ *and* $ac \leq bc$ *whenever* $a \leq b$*, since this implies* $a + c \leq b + c \leq b + d$ *whenever* $a \leq b$ *and* $c \leq d$*, and similarly for the product.*

One can always turn a semiring into a preordered semiring by defining $\leq$ to be either the equality preorder ($x \leq y \iff x = y$) or the other extreme preorder ($\forall x, y \in \mathcal{S} : x \leq y$). We shall only be interested in certain non-trivial preorders, namely semirings where $\mathbb{N} \subset \mathcal{S}$ and the preorder restricted to $\mathbb{N}$ is the usual ordering of $\mathbb{N}$.

Theorem 3.1.4 (Strassen, [15], see also [37, Theorem 2.2])**.** *Let* $(\mathcal{S}, \leq)$ *be a preordered semiring with* $\mathbb{N} \subseteq \mathcal{S}$ *satisfying the following:*

1. $\leq$ *restricted to* $\mathbb{N}$ *is the usual ordering of* $\mathbb{N}$*.*

2. *For any* $a, b \in \mathcal{S} \backslash \{0\}$ *there is an* $r \in \mathbb{N}$ *such that* $a \leq rb$*.*

Define the asymptotic preorder $\lesssim$ *on* $\mathcal{S}$ *by;* $a \lesssim b$ *if and only if* $a^N \leq 2^{x_N} b^N$ *for some integer-valued sequence* $x_N \in o(N)$*. Then* $(\mathcal{S}, \lesssim)$ *is also a preordered semiring. Let*

$$\Delta(\mathcal{S}) = \left\{ f \in \mathrm{Hom}(\mathcal{S}, \mathbb{R}^+) | \forall a, b \in \mathcal{S} : a \leq b \implies f(a) \leq f(b) \right\}.$$

That is $\Delta(\mathcal{S})$ *is the set of order-preserving semiring homomorphisms from* $\mathcal{S}$ *to* $\mathbb{R}^+$*. Then*

$$a \lesssim b \iff \forall f \in \Delta(\mathcal{S}) : f(a) \leq f(b). \tag{3.4}$$

Let $\Delta(\mathcal{S})$ be equipped with the topology generated by the maps $\hat{a} : \Delta(\mathcal{S}) \to \mathbb{R}$, given by $\hat{a} : f \mapsto f(a)$. That is, $\Delta(\mathcal{S})$ is equipped with the coarsest topology making these maps continuous. Then $\Delta(\mathcal{S})$ is a compact Hausdorff space and $a \mapsto \hat{a}$ is a semiring homomorphism $\mathcal{S} \to C(\Delta(\mathcal{S}))$, which, by Eq. (3.4), respects both $\lesssim$ and $\leq$ on $\mathcal{S}$. $\Delta(\mathcal{S})$ will be called the asymptotic spectrum of $\mathcal{S}$.

$\mathcal{T}_k$ of Definition 2.1.3 naturally comes equipped with the structure of a partially ordered semiring. When $\psi \in V_1 \otimes \cdots \otimes V_k$ and $\phi \in W_1 \otimes \cdots \otimes W_k$, the operations

$$\psi \odot \phi \in (V_1 \odot W_1) \otimes \ldots \otimes (V_k \odot W_k) \tag{3.5}$$

and

$$\psi \oplus \phi \in (V_1 \oplus W_1) \otimes \ldots \otimes (V_k \oplus W_k) \tag{3.6}$$

both respect restriction and therefore also equivalence, making $(\mathcal{T}_k, \odot, \oplus, \geq)$ a partially ordered semiring. Furthermore, the conditions of Theorem 3.1.4 are met (for details see [21] and [15]). The unit element of the semiring $\mathcal{T}_k$ is the equivalence class of simple tensors, which we might represent by $[u_1] = [|0 \ldots 0\rangle]$ and since $u_i \oplus u_j \sim u_{i+j}$ and $u_i \odot u_j \sim u_{ij}$, the embedding of the natural numbers in $\mathcal{T}_k$ is $r \mapsto [u_r]$. So the term 2^{x_N} of Theorem 3.1.4, in the case of $\mathcal{S} = \mathcal{T}_k$, corresponds to the element $u_2^{\odot x_N} \sim u_{2^{x_N}}$, and per definition

$$[\psi] \gtrsim [\phi] \iff \psi^{\odot N} \odot u_2^{\odot o(N)} \geq \phi^{\odot N}. \tag{3.7}$$

By [15, eq (2.7)-(2.10)], also

$$\psi \gtrsim \phi \iff \forall \theta > 1 \; \forall N \gg 1 : \psi^{\odot N} \odot u_{\lceil \theta^N \rceil} \geq \phi^{\odot N}. \tag{3.8}$$

Since for the rest of this chapter we shall be dealing with $\mathcal{T}_k$, we shall sometimes omit the brackets and simply write ψ in place of $[\psi]$, as this is unlikely to cause confusion.

Lemma 3.1.5. *Let $\psi, \phi \in \mathcal{T}_k$ with ψ globally entangled. Then*

$$\psi \gtrsim \phi \iff \omega(\psi, \phi) \leq 1. \tag{3.9}$$

Proof. Assume first that $\omega(\psi, \phi) \leq 1$ and let $\theta > 1$ be given. Let $r = \log R(\psi)$, such that $u_2^{\odot n} \geq \psi^{\odot \lfloor \frac{n}{r} \rfloor}$ for all n. Then

$$u_{\lceil \theta^N \rceil} \geq \psi^{\odot \left\lfloor \frac{N \log \theta}{r} \right\rfloor}, \tag{3.10}$$

which implies

$$\psi^{\odot N} \odot u_{\lceil \theta^N \rceil} \geq \psi^{\odot N} \odot \psi^{\odot \left\lfloor \frac{N \log \theta}{r} \right\rfloor} = \psi^{\odot \left\lfloor \left(1 + \frac{\log \theta}{r}\right) N \right\rfloor} \geq \phi^{\odot N}. \tag{3.11}$$

Here the last relation holds for $N \gg 1$, since $1 + \frac{\log \theta}{r} > 1 \geq \omega(\psi, \phi)$.

Conversely, assume that $[\psi] \gtrsim [\phi]$, such that $\psi^{\odot N} \odot u_2^{\odot x_N} \geq \phi^{\odot N}$ for some $x_N \in o(N)$. Let $\tau > 1$ be given. We need to show that $\psi^{\odot \lfloor \tau N \rfloor} \geq \phi^{\odot N}$ for large N. By Proposition 2.1.7 $\psi^{\odot \lfloor (\tau - 1) n \rfloor} \geq u_2$ for some n. So for large N:

$$\psi^{\odot \lfloor \tau N \rfloor} = \psi^{\odot N} \odot \psi^{\odot \lfloor (\tau-1) N \rfloor} \geq \psi^{\odot N} \odot u_2^{\odot \lfloor N/n \rfloor} \geq \psi^{\odot N} \odot u_2^{\odot x_N} \geq \phi^{\odot N}. \tag{3.12}$$

Here the second to last relation holds because $\lfloor N/n \rfloor$ dominates any function in $o(N)$ for large N. $\square$

Proposition 3.1.6. *For $(\mathcal{T}_k, \geq)$, the asymptotic preorder $\lesssim$ of Theorem 3.1.4, is precisely the preorder*

$$\psi \gtrsim \phi \iff \omega(\psi, \phi) \leq 1. \tag{3.13}$$

Proof. The proof of the implications $\Longleftarrow$ in Lemma 3.1.5 did not use the fact that $\psi^{\odot N}$ was globally entangled, so we just have to prove the reverse implication for general ψ. If $\psi \in V_1 \otimes \cdots \otimes V_k$ is separable across some bipartition where ϕ is not, then $\psi \not\gtrsim \phi$ and $\omega(\psi, \phi) = \infty$. So we can assume that $\psi = \psi_1 \otimes \cdots \otimes \psi_l$ and $\phi = \phi_1 \otimes \cdots \otimes \phi_l$ for some partition $I_1 \sqcup \cdots \sqcup I_l = [k]$ and $\psi_j, \phi_j \in \bigotimes_{i \in I_j} V_i$ with each ψ_j globally entangled in $\bigotimes_{i \in I_j} V_i$. It is not hard to see that $\omega(\psi, \phi) \leq 1$ if and only if $\omega(\psi_j, \phi_j) \leq 1$ for all j. Showing that also

$$\psi \gtrsim \phi \iff \forall j : \psi_j \gtrsim \phi_j, \tag{3.14}$$

finishes the proof, as Lemma 3.1.5 implies $\psi_j \gtrsim \phi_j \iff \omega(\psi_j, \phi_j) \leq 1$. So we prove (3.14). The implication $\Longleftarrow$ follows from the fact that $u_2^{[k]} \geq u_2^{I_j}$. The implication $\Longrightarrow$ is slightly harder. For the quantum information reader, who is comfortable with arguments by protocol description, we have to show that the sublinear number of global GHZ states can be replaced by local GHZ states within each entanglement cluster. This is done by having one party in each cluster locally create the states of the other entanglement clusters and then mimic the global protocol. In mathematical terms: Assume that $(f_1 \otimes \cdots \otimes f_k)(\psi^{\odot n} \odot u_2^{\odot x_n}) = \phi^{\odot n}$ and let a cluster $j_0 \in [l]$ be given. It suffices to prove that $\psi_{j_0}^{\odot n} \odot \left(u_2^{I_{j_0}}\right)^{\odot x_n} \geq \phi_{j_0}^{\odot n}$. Fix any party $i_0 \in I_{j_0}$. Consider the flattening of $\psi^{\odot n} \odot u_2^{\odot x_n}$, to a $|I_{j_0}|$-tensor:

$$Flat\left(\psi^{\odot n} \odot u_2^{\odot x_n}\right) \in \left(V_{i_0} \odot \bigodot_{i \notin I_j} V_i\right) \otimes \bigotimes_{i \in I_{j_0} \setminus \{i_0\}} V_i. \tag{3.15}$$

Notice that by applying a suitable map to V_{i_0}, we have $\psi_{j_0}^{\odot n} \odot \left(u_2^{I_{j_0}}\right)^{\odot x_n} \geq Flat\left(\psi^{\odot n} \odot u_2^{\odot x_n}\right)$ and $Flat\left(\psi^{\odot n} \odot u_2^{\odot x_n}\right) \geq Flat(\phi^{\odot n}) \geq \phi_{j_0}$, where the first restriction is the flattening of $(f_1 \otimes \cdots \otimes f_k)$, while the second is taking the partial trace on the systems $\left(V_{i_0} \odot \bigodot_{i \notin I_j} V_i\right)$. $\square$

Combining Proposition 3.1.6 and Theorem 3.1.4 we see that $\omega(\psi,\phi) \leq 1$ if and only if $f(\psi) \geq f(\phi)$ for all $f \in \Delta(\mathcal{T}_k)$. Since for any $a, b \in \mathbb{N}$: $\omega(\psi^{\odot a}, \phi^{\odot b}) = \frac{b}{a}\omega(\psi,\phi)$ we obtain the following formula:

Corollary 3.1.7.

$$\omega(\psi,\phi) = \max_{f\in\Delta(\mathcal{T}_k)} \frac{\log f(\phi)}{\log f(\psi)}, \tag{3.16}$$

where we ignore all f that are 1 on both ϕ and ψ, and interpret $\frac{x}{0}$ as ∞

Since $\log f(u_2) = 1$ for all $f \in \Delta$, we infer the following corollary from Corollary 3.1.7 applied to Eq. (2.31) and Eq. (2.32).

Corollary 3.1.8.

$$R^{\odot}(\psi) = \max_{f\in\Delta(\mathcal{T}_3)} f(\psi). \tag{3.17}$$

$$R^{\odot}_{sub}(\psi) = \min_{f\in\Delta(\mathcal{T}_3)} f(\psi). \tag{3.18}$$

Since we are often interested in expressing costs in terms of exponents, i.e. computing $\omega(\psi,\phi)$, we shall often consider the logarithmic spectrum $\log \Delta(\mathcal{T}_k)$, which is simply $\{\log f\}_{f\in\Delta(\mathcal{T}_k)}$. By Corollary 3.1.7, knowing the entire spectrum $\Delta(\mathcal{T}_k)$ means seeing the complete picture of asymptotic restrictions of k-tensor, or equivalently asymptotic SLOCC transformation rates between pure k-partite quantum states. It is then not surprising that finding members of $\Delta(\mathcal{T}_k)$ is not an easy task.

Example 3.1.9. *The first points in Δ_k which one finds are the local ranks. Given $\psi \in V_1\otimes\cdots\otimes V_k$ and any bipartition $I_1 \sqcup I_2 = [k]$, we may consider the flattening rank $R_{I_1,I_2}(\psi)$ of ψ viewed as a map $\bigotimes_{i\in I_1} V_i^* \to \bigotimes_{i\in I_2} V_i$. That is, the rank of the map $\langle\phi|_{I_1} \mapsto \langle\phi|_{I_1}|\psi\rangle_{I_1\sqcup I_2}$. It is easy to see that R_{I_1,I_2} is additive under $\oplus$ and multiplicative under $\odot$ and respects the restriction order, so $R_{I_1,I_2} \in \Delta(\mathcal{T}_k)$.*

In the following section, some non-trivial spectral points found in [13] and [8] will be presented.

3.2. Support and quantum functionals

In [13] V. Strassen found spectral points in $\Delta(\mathcal{O}_3)$, called support functionals, for a certain sub-semiring $\mathcal{O}_3 \subset \mathcal{T}_3$ of *oblique* tensors (see Definition 3.2.2 below). These spectral points were recently extended in [8, see Cor. 3.31] to monotones in $\Delta(\mathcal{T}_k)$. In the following sections we shall only work with $k = 3$, so here is a quick recap of the relevant results from [13] and [8] for $k = 3$.

Definition 3.2.1. *Let $\psi \in \mathbb{C}^X \otimes \mathbb{C}^Y \otimes \mathbb{C}^Z$ for some finite sets X, Y, Z, then* $\operatorname{supp}(\psi) \subset X \times Y \times Z$ *are the tuples (x, y, z), such that $\langle xyz|\psi\rangle \neq 0$.*

Definition 3.2.2. *Let $\psi \in \mathbb{C}^X \otimes \mathbb{C}^Y \otimes \mathbb{C}^Z$, then ψ is said to be oblique if for some total orderings of X, Y and Z, $\operatorname{supp}(\psi) \subset X \times Y \times Z$ forms as antichain in the product order. I.e. for $(x_1, y_1, z_1), (x_2, y_2, z_2) \in \operatorname{supp}(\psi)$;*

$$x_1 \geq x_2 \text{ and } y_1 \geq y_2 \text{ and } z_1 \geq z_2 \implies (x_1, y_1, z_1) = (x_2, y_2, z_2). \tag{3.19}$$

$\mathcal{O}_3$ is the subset of $[\psi] \in \mathcal{T}_3$ such that ψ is equivalent to an oblique tensor.

Definition 3.2.3. *Given $\Phi \subset X \times Y \times Z$ and probability measure $\theta : \{1, 2, 3\} \to [0, 1]$, define $h_\theta(\Phi)$ as*

$$h_\theta(\Phi) = \max_{P \in \mathcal{P}(\Phi)} \sum_{i=1}^{3} \theta(i) H(P_i), \tag{3.20}$$

where P_i is the marginal probability distribution on system i, e.g. $P_1(x') = \sum_{(x,y,z) \in \Phi : x = x'} P(x, y, z)$.

Theorem 3.2.4. *[13, Thm. 4.4.] For $[\psi] \in \mathcal{O}_3$, let $\phi \sim \psi$ be an oblique representative of the equivalence class, then the map*

$$\rho^\theta : [\psi] \to h_\theta(\operatorname{supp}(\phi)), \tag{3.21}$$

is well-defined and $\zeta^\theta = 2^{\rho^\theta} \in \Delta(\mathcal{O}_3)$. Or equivalently, $\rho^\theta \in \log \Delta(\mathcal{O}_3)$.

The subset $\{\zeta^\theta\}_\theta \subset \Delta(\mathcal{O}_3)$ is what Strassen calls the *support simplex*, as it is the image of the simplex $\mathcal{P}([3])$, under the continuous map $\theta \mapsto \zeta^\theta$. Note that if $\left(\theta(i)\right)_{i=1}^3 = (1, 0, 0)$, then ζ^θ is simply the local rank $R_{\{1\},\{2,3\}}$ of Example 3.1.9.

Theorem 3.2.5. *[8, Cor. 3.31] For $\psi \in \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ and probability measure $\theta : \{1, 2, 3\} \to [0, 1]$, let $H_\theta(\psi) = \sum_{i=1}^3 \theta(i) H(\rho_i)$, where $\rho_i = \operatorname{Tr}_{\{1,2,3\}\setminus\{i\}} \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle}$ is the normalized marginal state of $|\psi\rangle$ on system i and $H(\rho) = -\operatorname{Tr}(\rho \log \rho)$ is the von Neumann entropy. For $[\psi] \in \mathcal{T}_3$, let*

$$E_\theta(\psi) = \sup_{\phi \leq \psi} H_\theta(\phi). \tag{3.22}$$

Then $F_\theta = 2^{E_\theta} \in \Delta(\mathcal{T}_3)$. Furthermore $F^\theta|_{\mathcal{O}_3} = \zeta^\theta$ (see [8, Sec. 3.4]).

Whether these so-called *quantum functionals* or Theorem 3.2.5 form the entirety of $\Delta(\mathcal{T}_3)$ is unclear. If this was the case, it would imply that the maximal local dimension is an upper bound to the asymptotic rank, and in particular that $\omega(\mathrm{MaMu}_2) = 2$. Though no such strong results are known for $R^\odot$, a weaker version, for the sub-semiring of tight tensors (defined below) is known for $R^\odot_{sub}$, as shall now be explained.

Definition 3.2.6. *A tensor $\psi \in \mathbb{C}^X \otimes \mathbb{C}^Y \otimes \mathbb{C}^Z$ is said to be* tight *if there exist injective maps $\alpha : X \to \mathbb{Z}$, $\beta : Y \to \mathbb{Z}$, $\gamma : Z \to \mathbb{Z}$ such that*

$$(x, y, z) \in \mathrm{supp}(\psi) \implies \alpha(x) + \beta(y) + \gamma(z) = 0. \tag{3.23}$$

Tight tensors are necessarily oblique and form a sub-semiring of $\mathcal{T}_3$ (see [13, sec. 5]). We denote by $\mathcal{O}_3^{\text{tight}} \subset \mathcal{T}_3$ the sub-semiring of equivalence classes of tight tensors.

Theorem 3.2.7. *[13, Theorem 5.5] For any sub-semiring $\mathcal{S} \subset \mathcal{O}_3^{tight}$ we order $\Delta(\mathcal{S})$ by the usual ordering of functions by point-wise majorization. The minimal points of $\{\zeta^\theta|_{\mathcal{S}}\}_\theta \subset \Delta(\mathcal{S})$ coincide with the minimal points of $\Delta(\mathcal{S})$.*

When $\mathcal{S} = \langle \psi_1, \dots, \psi_n \rangle \subset \mathcal{T}$ is the sub-semiring generated by $\psi_1, \dots, \psi_n$, let us write $\Delta(\psi_1, \dots, \psi_n) = \Delta(\mathcal{S})$. We shall always think of sub-semirings as unital. For $[\psi] \in \mathcal{O}_3^{\text{tight}}$, we may consider the sub-semiring $\langle \psi \rangle \subset \mathcal{T}_3$ generated by ψ. Since any $f \in \Delta(\psi)$ is determined by its action on ψ, it follows from Theorem 3.2.7 that

$$R_{\text{sub}}^{\odot}(\psi) = \min_{f \in \Delta(\psi)} f(\psi) = \min_{\theta \in \mathcal{P}(\{1,2,3\})} \zeta^\theta(\psi). \tag{3.24}$$

So while the asymptotic rank is not known to be determined by the support simplex, for tight tensors, the asymptotic sub-rank is. Whether this generalizes to the sub-rank of general tensors and the quantum functionals, is currently unknown.

Example 3.2.8. *The tensor $W \in \mathbb{C}^X \otimes \mathbb{C}^Y \otimes \mathbb{C}^Z$, with $X = Y = Z = \{0, 1\}$ has*

$$\mathrm{supp}(W) = \{x_1 = (1, 0, 0), x_2 = (0, 1, 0), x_3 = (0, 0, 1)\} \subset \{0, 1\}^{\times 3}, \tag{3.25}$$

which is tight as witnessed by the injective maps $\alpha, \beta, \gamma : \{0, 1\} \to \mathbb{Z}$, given by $\alpha(x) = x$, $\beta(y) = y$, $\gamma(z) = z - 1$. When $P \in \mathcal{P}(\{x, y, z\})$, the entropy of the marginals are $H(P_1) = h(P(x_1))$, $H(P_2) = h(P(x_2))$, $H(P_3) = h(P(x_3))$, where $h(p) = -p \log p - (1 - p) \log(1 - p)$ is the binary entropy function. Let Q be the uniform distribution on $\{x_1, x_2, x_3\}$, then

$$\begin{aligned} \rho^\theta(W) &= \max_{P \in \mathcal{P}(\mathrm{supp} W)} \theta(1)h(P(x_1)) + \theta(2)h(P(x_2)) + \theta(3)h(P(x_3)) \\ &\geq \theta(1)h(Q(x_1)) + \theta(2)h(Q(x_2)) + \theta(3)h(Q(x_3)) = h\left(\frac{1}{3}\right) = \log(3) - \frac{2}{3}. \end{aligned} \tag{3.26}$$

When θ_0 is the uniform distribution, the fact that h is concave, implies, by Jensen's inequality, that Eq. (3.26) becomes an equality, so

$$\rho^{\theta_0}(W) = \log(3) - \frac{2}{3}. \tag{3.27}$$

In conclusion

$$\log R_{sub}^{\odot}(W) = \min_{\theta \in \mathcal{P}([3])} \rho^\theta(W) = \log(3) - \frac{2}{3} \approx 0.9183, \tag{3.28}$$

as was also shown in [26]. Furthermore, since $\log R^{\odot}(W) = \max_{\theta \in \mathcal{P}([3])} \rho^\theta(W) = 1$ and $\rho^{(1,0,0)}(W) = 1$ and $\theta \mapsto \rho^\theta$ is continuous, we conclude that $\log \Delta(W) \cong [\log(3) - \frac{2}{3}, 1]$.

3.3. The support simplex of W and one EPR-pair

Let us add another tensor into the mix and see if we can determine the spectrum. Let us consider the unnormalized EPR-pair $|e\rangle = |e^{1,2}\rangle = |000\rangle + |110\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}$, which is also tight, as witnessed by the maps $\alpha, \beta, \gamma : \{0,1\} \to \mathbb{Z}$, given by $\alpha(x) = -x$, $\beta(y) = y$, $\gamma(z) = z$. We shall determine $\log \Delta(e, W)$.

Since $\mathrm{supp}(e)$ is a two-point set, any $P \in \mathcal{P}\left(\mathrm{supp}(e)\right)$ is determined by the value $p = P\left((0,0,0)\right)$. Note also that the marginal P_3 is supported on a one-point set, so that $H(P_3) = 0$. So

$$\begin{aligned} \rho^\theta(e) &= \max_{P \in \mathcal{P}(\mathrm{supp}(e))} \theta(1)H(P_1) + \theta(2)H(P_2) \\ &= \max_{p \in [0,1]} \left(\theta(1) + \theta(2)\right) h(p) \\ &= \theta(1) + \theta(2). \end{aligned} \tag{3.29}$$

Since $f \in \log \Delta(e, W)$ is determined by its value at e and W, we may visualize $\log \Delta(e, W)$ as a subset of $\mathbb{R}^2_+$ by the inclusion $\log \Delta(e, W) \ni f \mapsto \left(f(e), f(W)\right) \in \mathbb{R}^2_+$. The points $\rho^\theta \in \log \Delta(e, W)$, for $\theta \in \mathcal{P}\left(\{1,2,3\}\right)$, then correspond to the points

$$\Big(\left[\theta(1) + \theta(2)\right], \rho^\theta(W)\Big).$$

We describe what this set looks like. Let $\mathrm{supp}(W) = \{x_1, x_2, x_3\}$ as in Eq. (3.25) and for $P \in \mathcal{P}\left(\mathrm{supp}(W)\right)$, let

$$F(\theta, P) = \sum_{i \in \{1,2,3\}} \theta(i) h\left(P(x_i)\right), \tag{3.30}$$

such that

$$\rho^\theta(W) = \max_{P \in \mathcal{P}\left(\{x_1, x_2, x_3\}\right)} F(\theta, P). \tag{3.31}$$

For any $\theta \in \mathcal{P}\left(\{1,2,3\}\right)$, let $\tilde{\theta}$ be given by $\tilde{\theta}(1) = \tilde{\theta}(2) = \frac{\theta(1)+\theta(2)}{2}$, $\tilde{\theta}(3) = \theta(3)$. Then

$$\rho^{\tilde{\theta}}(W) = \max_{P \in \mathcal{P}} F(\tilde{\theta}, P) = \max_{\substack{P \in \mathcal{P} \\ P(x_1) = P(x_2)}} F(\tilde{\theta}, P) = \max_{\substack{P \in \mathcal{P} \\ P(x_1) = P(x_2)}} F(\theta, P) \leq \rho^\theta(W), \tag{3.32}$$

where the second equality holds by concavity of h and the third is true since $F(\tilde{\theta}, P) = F(\theta, P)$ whenever $P(x_1) = P(x_2)$. From Eq. (3.32) it follows that for fixed $t = \rho^\theta(e) = \theta(1) + \theta(2)$, the minimal value of $\rho^\theta(W)$ is attained at $(\theta_t(1), \theta_t(2), \theta_t(3)) = (t/2, t/2, 1-t)$, with value

$$\rho^{\theta_t}(W) = \max_p \frac{t}{2} h(p) + \frac{t}{2} h(p) + (1-t) h(1-2p) = \max_p \left[t h(p) + (1-t) h(2p)\right]. \tag{3.33}$$

Let L_p be the line parametrized by $t \mapsto \left(t, t h(p) + (1-t) h(2p)\right) \in \mathbb{R}^2$. Then the lines L_p cut out the support simplex from below. See Fig. 3.1 for illustration.

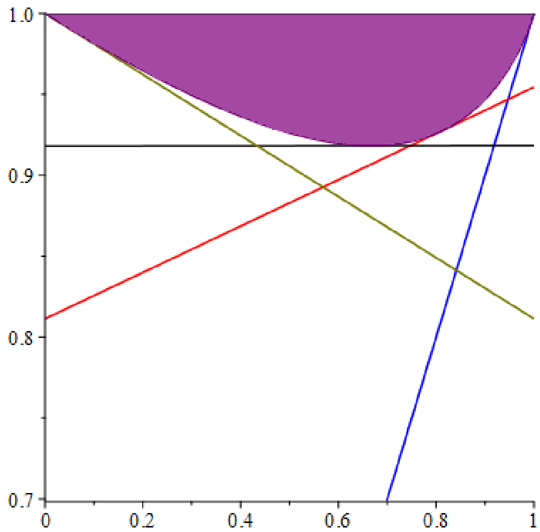

Figure 3.1: Support simplex $\{\rho^\theta\}_{\theta\in\mathcal{P}(\{1,2,3\})}$ shown in purple.
$\rho^\theta(e)$ along the x-axis and $\rho^\theta(W)$ along the y-axis.
Tangent lines $L_{1/4}, L_{1/3}, L_{3/8}, L_{1/2}$ shown in respectively olive, black, red and blue.

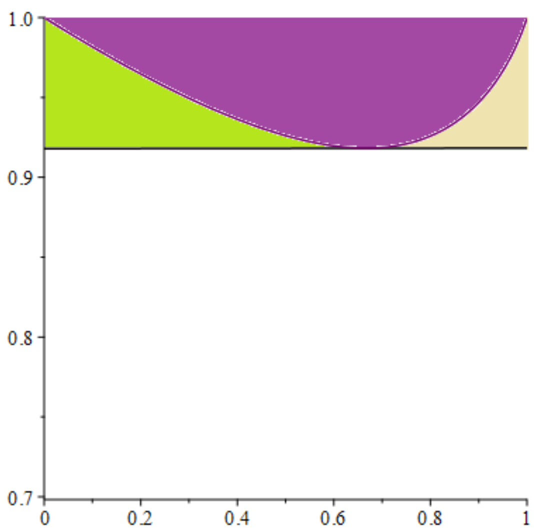

Figure 3.2: The green region is not in $\log\Delta(e, W)$ by Theorem 3.2.7
The yellow region needs to be determined

Now by Theorem 3.2.7, we can rule out the green region in Fig. 3.2 as points in $\log\Delta(e, W)$. In other words, all points below the lines L_p for $p \in \left[\frac{1}{4}, \frac{1}{3}\right]$ are not in $\log\Delta(e, W)$. To deal with the yellow region will take some work, deferred to Section 3.4, but let us first see what our current knowledge of the spectrum allows us to say about ω and SLOCC. Let $p \in \left[\frac{1}{4}, \frac{1}{3}\right]$. Since L_p cuts out $\log\Delta(e, W)$ from below, we know that for any $f \in \log\Delta(e, W)$, we have $f(W) \geq f(e)h(p) + (1 - f(e))\, h(2p)$. So for any $f \in \log\Delta(e, W)$

$$\begin{aligned} f(W) + (h(2p) - h(p))\, f(e) &\geq f(e)h(p) + (1 - f(e))\, h(2p) + (h(2p) - h(p))\, f(e) \\ &\geq h(2p), \end{aligned} \tag{3.34}$$

from which it follows, by Corollary 3.1.8, that for any $n \in \mathbb{N}$ and $p \in \left[\frac{1}{4}, \frac{1}{3}\right]$

$$W^{\odot n} \odot e^{\odot\lceil (h(2p)-h(p))n\rceil} \gtrsim u_2^{\odot\lfloor nh(2p)\rfloor}. \tag{3.35}$$

In terms of SLOCC, this means that given a large number of W-states and $(h(2p) - h(p))$ EPR-pairs shared between two of the parties per W-state, we can extract $h(2p)$ GHZ-states per W. Putting $p = \frac{1}{3}$, we recover the asymptotic sub-rank of W from Example 3.2.8. For $p = \frac{1}{4}$ we see that we can asymptotically convert W states to GHZ states at rate 1-to-1 given the aid of $(h(1/2) - h(1/4)) \approx 0.189$ EPR-pairs per W-state.

Theorem 3.2.7 can only tell us that the green region of Fig. 3.2 is not in $\log\Delta(e, W)$. In the following section we will show that Eq. (3.34) also holds for $p \in \left[\frac{1}{3}, \frac{1}{2}\right]$. Note, that for such p, $h(2p) - h(p)$ is negative. Therefore the expression Eq. (3.35) does not really make sense, as we have not defined negative Kronecker powers of tensors. The correct interpretation is then that negative powers should be moved to the other side of the inequality. In the proof of Theorem 3.4.1 below, we will essentially prove the negative power version of Eq. (3.35) in the more general setting where, W is replaced with any tight tensor.

3.4. Extracting EPR-pairs and GHZ-states from a tight tensor

In order to determine whether the yellow section of Fig. 3.2 is in $\log\Delta(e, W)$, we need to get a bit technical. The proof of the following theorem uses a variation of the technique used for proving Theorem 3.2.7 in [13]. The technique in [13], which is an integral part of what is called the laser method, is a way of extracting large unit tensors from a tight tensor. In what follows, the same technique is used for extracting direct sums of high-level maximally entangled states between a fixed pair of parties.

Theorem 3.4.1. *Let $\psi \in \mathbb{C}^X \otimes \mathbb{C}^Y \otimes \mathbb{C}^Z$ be a tight tensor with $\Psi = \mathrm{supp}(\psi) \subset X \times Y \times Z$ and $|e\rangle = |000\rangle + |110\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}$. Let $f \in \log \Delta(e, \psi)$ and $P \in \mathcal{P}(\Psi)$ such that $H(P_3) \leq \min\{H(P_1), H(P_2)\}$. Let $s = f(\psi)$ and $t = f(e)$. Then*

$$s \geq t \left(\min_{i \in \{1,2\}} H(P_i) - H(P_3) \right) + H(P_3). \tag{3.36}$$

Proof. The overall idea is to first take a large tensor power $\psi^{\odot n} \in \mathbb{C}^{X^n} \otimes \mathbb{C}^{Y^n} \otimes \mathbb{C}^{Z^n}$, and then choose subsets $X_n \subset X^n, Y_n \subset Y^n, Z_n \subset Z^n$, such that projecting onto these subsets creates a tensor $\phi_n = (\pi_{X_n} \otimes \pi_{Y_n} \otimes \pi_{Z_n}) \psi^{\odot n}$, equivalent to $u_2^{\odot \alpha} \odot e^{\odot \beta}$, where $\alpha \approx nH(P_3)$ and $\beta \approx n \left(\min_{i \in \{1,2\}} H(P_i) - H(P_3) \right)$.

We may assume that the values of P are rational and obtain the general result by taking limits. Let $\delta > 0$ be given. Let $n \in \mathbb{N}$ be large with nP integral, such that we may consider the type classes $T^n_{P_1}, T^n_{P_2}, T^n_{P_3}$ (defined in Section 1.2). Let $\Psi^n_P = \Psi^n \cap \left[T^n_{P_1} \times T^n_{P_2} \times T^n_{P_3} \right]$. Note that for each $z \in T^n_{P_3}$ the sets $\Psi^n_P \cap (X^n \times Y^n \times \{z\})$ are of the same size, so $\left| \Psi^n_P \cap (X^n \times Y^n \times \{z\}) \right| = \frac{|\Psi^n_P|}{\left| T^n_{P_3} \right|}$, and similarly for $x \in T^n_{P_1}$ and $y \in T^n_{P_2}$. Let $M = M(n)$ be a prime depending on n with

$$\frac{M}{2} \leq \frac{|\Psi_{P^n}|}{\min\left(|T^n_{P_1}|, |T^n_{P_2}| \right)} 2^{n\delta} \leq M. \tag{3.37}$$

By [13, Proof of Lemma 5.1 pp 151-153] (for completeness the statement is proven in Lemma 3.4.2 below), we may choose n and therefore M sufficiently large that the following holds: There exist random maps $a : X^n \to \mathbb{Z}_M$, $b : Y^n \to \mathbb{Z}_M$, $c : Z^n \to \mathbb{Z}_M$ and a set $S \subset \mathbb{Z}_M$ of size $|S| > M^{1-\delta}$, such that for $(x, y, z) \in \Psi^n_P$

$$\Pr\left[a(x), b(y), c(z) \in S \right] = \frac{|S|}{M^2}, \tag{3.38}$$

and for all pairs of distinct $(x, y, z), (x', y', z') \in \Psi^n_P$ with $x = x'$ or $y = y'$ or $z = z'$

$$\Pr\left[a(x'), a(x'), b(y), b(y'), c(z), c(z') \in S \right] = \frac{|S|}{M^3}. \tag{3.39}$$

Note that a, b, c are not random maps in the sense that their domains X^n, Y^n, Z^n are measure spaces, but rather, in the sense that the maps a, b, c themselves are randomly chosen. That is, for some probability space Ω, we have $\Omega \ni \omega \mapsto a_\omega \in \mathrm{Hom}_{\mathrm{set}}(X^n, \mathbb{Z}_M)$, where we have suppressed the ω in the notation, as one usually does with stochastic variables.

Now let $V = \Psi_P^n \cap \left(a^{-1}(S) \times b^{-1}(S) \times c^{-1}(S)\right)$ and

$$E = \left\{ \left\{(x,y,z),(x',y',z')\right\} \in \binom{V}{2} \middle| x = x' \text{ or } y = y' \right\}. \tag{3.40}$$

Let $\Gamma \subset V$ be the set of isolated vertices in the graph (V, E). Furthermore, let $V_z = V \cap (X^n \times Y^n \times \{z\})$ be the vertices with fixed z coordinate, and E_z be the edges with at least one end-node in V_z. For each z let $\Gamma_z = \Gamma \cap V_z$. Let $X_n = \pi_{X^n}(\Gamma), Y_n = \pi_{Y^n}(\Gamma), Z_n = \pi_{Z^n}(\Gamma)$, such that the support of $\phi_n = (\pi_{X_n} \otimes \pi_{Y_n} \otimes \pi_{Z_n})\, \psi^{\odot n}$ is $\operatorname{supp}(\phi_n) = \Gamma = \bigcup_{z \in T_{P_3}^n} \Gamma_z$. By the fact that Γ is the set of isolated vertices in (V, E), all x and y coordinates of points in Γ are unique. So

$$\Gamma = \bigcup_{z \in T_{P_3}^n} \bigcup_{(x,y,z) \in \Gamma_z} (x, y, z), \tag{3.41}$$

with each x and y unique in the union above. So by a rescaling $f_X : \mathbb{C}^{X_n} \mapsto \mathbb{C}^{X_n}$ given by $f_X : |x\rangle \mapsto \frac{1}{\langle xyz|\phi_n\rangle} |x\rangle$, we get

$$|\phi_n\rangle \overset{f_X \otimes I \otimes I}{\sim} \sum_{z \in T_{P_3}^n} \sum_{(x,y,z) \in \Gamma_z} |x\rangle\, |y\rangle\, |z\rangle \sim \sum_{z \in T_{P_3}^n} \sum_{x=1}^{|\Gamma_z|} |x, z\rangle\, |x, z\rangle\, |z\rangle \sim \bigoplus_z |e_{|\Gamma_z|}\rangle\,. \tag{3.42}$$

Here $|e_m\rangle = \sum_{i=0}^{m-1} |ii0\rangle$ denotes the m-level bipartite unit tensor shared between the first two systems. For reasons which will become clear, we are interested in lower bounding the t'th moment $\mathbb{E}|\Gamma_z|^t$ for $z \in T_{P_3}^n$. To do this, we lower bound $\mathbb{E}|\Gamma_z|$ and then upper bound $\mathbb{E}|\Gamma_z|^2$, since a lower bound of $\mathbb{E}|\Gamma_z|^t$ then follows from Hölder's inequality. First, the expectation value: Since a subgraph with no edges consists of isolated points, and adding an edge will de-isolate only the vertices it connects we get the coarse lower bound $|\Gamma_z| \geq |V_z| - 2|E_z|$.

$$\begin{aligned} \mathbb{E}|V_z| &= \sum_{(x,y,z) \in \Psi_P^n \cap \left(X^n \times Y^n \times \{z\}\right)} \Pr\left[a(x) \in S, b(y) \in S, c(z) \in S\right] \\ &= \left|\Psi_P^n \cap (X^n \times Y^n \times \{z\})\right| \frac{|S|}{M^2} \\ &= \frac{\left|\Psi_P^n\right| |S|}{\left|T_{P_3}^n\right| M^2}. \end{aligned} \tag{3.43}$$

$$
\begin{aligned}
\mathbb{E}|E_z| \leq &\sum_{(x,y,z)\in\Psi_P^n\cap\left(X^n\times Y^n\times\{z\}\right)} \left(\sum_{\substack{(x',y,z')\in\Psi_P^n\cap(X^n\times\{y\}\times Z^n)\\(x',y,z')\neq(x,y,z)}} \Pr\left[a(x),a(x'),b(y),c(z),c(z')\in S\right] \right.\\
&\left. + \sum_{\substack{(x,y',z')\in\Psi_P^n\cap(\{x\}\times Y^n\times Z^n)\\(x',y,z')\neq(x,y,z)}} \Pr\left[a(x),b(y),b(y'),c(z),c(z')\in S\right] \right)\\
&< \left|\Psi_P^n\cap\left(X^n\times Y^n\times\{z\}\right)\right| \left[2\max\left\{\left|\Psi_P^n\cap\left(X^n\times\{y\}\times Z^n\right)\right|,\left|\Psi_P^n\cap\left(\{x\}\times Y^n\times Z^n\right)\right|\right\}\frac{|S|}{M^3}\right]\\
&= \frac{2\left|\Psi_P^n\right|^2|S|}{\left|T_{P_3}^n\right|\min\left\{\left|T_{P_1}^n\right|,\left|T_{P_2}^n\right|\right\}M^3}.
\end{aligned}
\tag{3.44}
$$

Combining Eq. (3.43) and Eq. (3.44) with both inequalities in Eq. (3.37) yields

$$
\begin{aligned}
\mathbb{E}|\Gamma_z| \geq \mathbb{E}|V_z| - 2\mathbb{E}|E_z| &\overset{3.43,3.44}{\geq} \frac{\left|\Psi_P^n\right||S|}{\left|T_{P_3}^n\right|M^2}\left(1-\frac{4\left|\Psi_P^n\right|}{M\min\left\{\left|T_{P_1}^n\right|,\left|T_{P_2}^n\right|\right\}}\right)\\
&\overset{3.37(second)}{\geq} \frac{\left|\Psi_P^n\right||S|}{\left|T_{P_3}^n\right|M^2}\left(1-\frac{4}{2^{n\delta}}\right)\\
&\overset{3.37(first)}{\geq} \frac{\min\left\{\left|T_{P_1}^n\right|,\left|T_{P_2}^n\right|\right\}}{\left|T_{P_3}^n\right|}\frac{1}{M^\delta 2^{n\delta+1}}\left(1-\frac{4}{2^{n\delta}}\right).
\end{aligned}
\tag{3.45}
$$

Now for the second moment. First off $|\Gamma_z|^2 \leq |V_z|^2$, and the latter can be estimated from above as follows:

$$
\begin{aligned}
\mathbb{E}\left|V_z\right|^2 = &\sum_{(x,y,z)\in\Psi_P^n\cap\left(X^n\times Y^n\times\{z\}\right)}\ \sum_{(x',y',z)\in\Psi_P^n\cap\left(X^n\times Y^n\times\{z\}\right)} \Pr\left[a(x),a(x'),b(y),b(y'),c(z)\in S\right]\\
&< \sum_{(x,y,z)\in\Psi_P^n\cap\left(X^n\times Y^n\times\{z\}\right)} \frac{|S|}{M^2} + \sum_{(x,y,z)\in\Psi_P^n\cap\left(X^n\times Y^n\times\{z\}\right)}\ \sum_{(x',y',z)\in\Psi_P^n\cap\left(X^n\times Y^n\times\{z\}\right)} \frac{|S|}{M^3}\\
&= \frac{\left|\Psi_P^n\right||S|}{\left|T_{P_3}^n\right|M^2} + \frac{\left|\Psi_P^n\right|^2|S|}{\left|T_{P_3}^n\right|^2M^3} = \frac{\left|\Psi_P^n\right||S|}{\left|T_{P_3}^n\right|M^2}\left(1+\frac{\left|\Psi_P^n\right|}{\left|T_{P_3}^n\right|M}\right)\\
&\overset{3.37}{\leq} \frac{\min\left\{\left|T_{P_1}^n\right|,\left|T_{P_2}^n\right|\right\}}{\left|T_{P_3}^n\right|}\left(1+\frac{\min\left\{\left|T_{P_1}^n\right|,\left|T_{P_2}^n\right|\right\}}{\left|T_{P_3}^n\right|}\right) \leq 2\left(\frac{\min\left\{\left|T_{P_1}^n\right|,\left|T_{P_2}^n\right|\right\}}{\left|T_{P_3}^n\right|}\right)^2.
\end{aligned}
\tag{3.46}
$$

Here the last inequality follows from the fact that $\min\left\{\left|T_{P_1}^n\right|,\left|T_{P_2}^n\right|\right\} \geq \left|T_{P_3}^n\right|$. We can now use the estimates of the first and second moment to lower bound $\mathbb{E}\left|\Gamma_z\right|^t$. Let $\rho = \frac{t}{2-t}$, $p = 2-t$, $q = \frac{2-t}{1-t}$.

Since $\frac{1}{p}+\frac{1}{q}=\frac{1}{2-t}+\frac{1-t}{2-t}=1$, it follows from Hölder's inequality applied to $|\Gamma_z|=|\Gamma_z|^{\rho}|\Gamma_z|^{1-\rho}$ that

$$\mathbb{E}|\Gamma_z| \leq \left(\mathbb{E}|\Gamma_z|^{\rho p}\right)^{\frac{1}{p}}\left(\mathbb{E}|\Gamma_z|^{(1-\rho)q}\right)^{\frac{1}{q}}, \tag{3.47}$$

which implies, by taking the p'th power on both sides, that

$$\left(\mathbb{E}|\Gamma_z|\right)^{2-t} \leq \left(\mathbb{E}|\Gamma_z|^{t}\right)\left(\mathbb{E}|\Gamma_z|^{2}\right)^{1-t}. \tag{3.48}$$

Therefore

$$\begin{aligned}\mathbb{E}|\Gamma_z|^t &\geq \frac{\left(\mathbb{E}|\Gamma_z|\right)^{2-t}}{\left(\mathbb{E}|\Gamma_z|^2\right)^{1-t}} > \frac{\left(\frac{\min\left\{\left|T^n_{P_1}\right|,\left|T^n_{P_2}\right|\right\}}{\left|T^n_{P_3}\right|}\frac{1}{M^{\delta}2^{n\delta+1}}\left(1-\frac{4}{2^{n\delta}}\right)\right)^{2-t}}{2^{1-t}\left(\frac{\min\left\{\left|T^n_{P_1}\right|,\left|T^n_{P_2}\right|\right\}}{\left|T^n_{P_3}\right|}\right)^{2-2t}} \\ &= \frac{\min\left\{\left|T^n_{P_1}\right|,\left|T^n_{P_2}\right|\right\}^{t}}{\left|T^n_{P_3}\right|}2^{t-1}\left(\frac{1}{M^{\delta}2^{n\delta+1}}\left(1-\frac{4}{2^{n\delta}}\right)\right)^{2-t},\end{aligned} \tag{3.49}$$

such that

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}|\Gamma_z|^t &\geq t\lim_{n\to\infty}\frac{1}{n}\log\frac{\min\left\{\left|T^n_{P_1}\right|,\left|T^n_{P_2}\right|\right\}}{\left|T^n_{P_3}\right|} \\ &\quad+(2-t)\left(-\delta-\delta\lim_{n\to\infty}\frac{1}{n}\log M+\lim_{n\to\infty}\frac{1}{n}\log\left(1-\frac{4}{2^{n\delta}}\right)\right) \\ &\geq t\left(\min\left\{H(P_1),H(P_2)\right\}-H(P_3)\right)+(2-t)(-\delta)\left(1+\log|\Psi|+\delta\right).\end{aligned} \tag{3.50}$$

Here the last inequality comes from the upper bound on M in Eq. (3.37):

$$\lim_{n\to\infty}\frac{1}{n}\log M(n) \leq \lim_{n\to\infty}\frac{1}{n}\log|\Psi|^n 2^{n\delta} = \log|\Psi|+\delta. \tag{3.51}$$

If $f\in\log\Delta(e,W)$, then

$$2^{f\left(\bigoplus_{z\in T^n_{P_3}} e_{|\Gamma_z|}\right)} = \sum_{z\in T^n_{P_3}} 2^{f\left(e_{|\Gamma_z|}\right)} = \sum_{z\in T^n_{P_3}} 2^{f(e)\log|\Gamma_z|} = \sum_{z\in T^n_{P_3}}|\Gamma_z|^t. \tag{3.52}$$

So

$$ns = nf(\psi) = f\left(\psi^{\odot n}\right) \geq f(\phi_n) = f\left(\bigoplus_{z\in T^n_{P_3}} e_{|\Gamma_z|}\right) = \log\sum_{z\in T^n_{P_3}}|\Gamma_z|^t. \tag{3.53}$$

For some realization of the stochastic variable $\sum_{z\in T^n_{P_3}}|\Gamma_z|^t$, we have $\sum_{z\in T^n_{P_3}}|\Gamma_z|^t \geq \mathbb{E}\sum_{z\in T^n_{P_3}}|\Gamma_z|^t$, and for this realization

$$s \geq \frac{1}{n}\log\sum_{z\in T^n_{P_3}}|\Gamma_z|^t \geq \frac{1}{n}\log\mathbb{E}\sum_{z\in T^n_{P_3}}|\Gamma_z|^t = \frac{1}{n}\log\sum_{z\in T^n_{P_3}}\mathbb{E}|\Gamma_z|^t \geq \frac{1}{n}\log\left|T^n_{P_3}\right| + \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}|\Gamma_z|^t, \tag{3.54}$$

which in the limit $n \to \infty$, gives

$$s \geq H(P_3) + t\Big(\min\{H(P_1), H(P_2)\} - H(P_3)\Big) + (2-t)\left(-\delta(1 + \log|\Psi| + \delta)\right). \tag{3.55}$$

Taking $\delta \to 0$ removes the last term and concludes the proof. $\square$

Applying Theorem 3.4.1 to $\psi = W$, we see that $\log \Delta(W, e)$ is exactly the purple region in Fig. 3.2, so that the support simplex (or equivalently the quantum functionals of [8]) is the entire spectrum. By Theorem 3.4.1, for any $p \in \left[\frac{1}{3}, \frac{1}{2}\right]$,

$$W^{\odot n} \gtrsim e^{\odot \left\lfloor \left(h(p) - h(2p)\right)n \right\rfloor} \odot u_2^{\odot \lfloor nh(2p) \rfloor}. \tag{3.56}$$

When $p = \frac{1}{3}$ we recover the subrank of W. For $p = \frac{1}{2}$ we get $W^{\odot n} \gtrsim e^{\odot n}$, which follows trivially from $W \geq e$. But for p strictly between $\frac{1}{3}$ and $\frac{1}{2}$, Eq. (3.56) seems about as non-trivial as Theorem 3.4.1.

In terms of asymptotic restrictions for tight tensors, ψ, in general, Theorem 3.4.1 implies

$$\psi^{\odot n} \gtrsim e^{\odot \left\lfloor \left(\min_{i \in \{1,2\}} H(P_i) - H(P_3)\right)n \right\rfloor} \odot u_2^{\odot \lfloor nH(P_3) \rfloor} \tag{3.57}$$

for any probability distribution $P \in \mathcal{P}(\mathrm{supp}(\psi))$ with $H(P_3) < \min_{i \in \{1,2\}} H(P_i)$.

Before ending this chapter, a proof of the technical lemma, used in the proof of Theorem 3.4.1 is presented below.

Lemma 3.4.2 (Technical lemma following the proof in [13]). *Let $\Psi \subset X \times Y \times Z$ be tight and $\delta > 0$ be given. For sufficiently large n and prime $M = M(n) \xrightarrow{n} \infty$ there exist random maps $a : X^n \to \mathbb{Z}_M$, $b : Y^n \to \mathbb{Z}_M$, $c : Z^n \to \mathbb{Z}_M$ and a set $S \subset \mathbb{Z}_M$ of size $|S| \geq M^{1-\delta}$ such that Eq. (3.38) and Eq. (3.39) hold.*

Proof. Let $\delta > 0$ be given and let $\alpha, \beta, \gamma : X, Y, Z \to \mathbb{Z}$ be the injective maps witnessing tightness of Ψ. By adding a large constant K, to both α and β, while subtracting $2K$ from γ, we can assume that α and β only take positive values, while γ only takes negative values. This will be important later and resolves an issue with the original proof in [13], which seems to have gone unnoticed. Let M be a prime larger then $2|\alpha(x)|, 2|\beta(y)|$ and $2|\gamma(z)|$ for all x, y, z. Let $\omega_1, \ldots, \omega_{n+3}$ be independent uniformly distributed stochastic variables with values in $\mathbb{Z}_M$. Now define $a : X^n \to \mathbb{Z}_M$, $b : Y^n \to \mathbb{Z}_M$ and $c : Z^n \to \mathbb{Z}_M$ by

$$a(x) = \sum_{i=1}^{n} \alpha(x_i)\omega_i + \omega_{n+1} - \omega_{n+2}, \tag{3.58}$$

$$b(x) = \sum_{i=1}^{n} \beta(y_i)\omega_i + \omega_{n+2} - \omega_{n+3}, \tag{3.59}$$

$$c(x) = -\frac{1}{2}\left(\sum_{i=1}^{n} \gamma(z_i)\omega_i + \omega_{n+3} - \omega_{n+1}\right). \tag{3.60}$$

By [38] we may find $S \subset \left\{0, \dots, \frac{M-1}{2}\right\}$ of size $|S| \geq M^{1-\delta}$ such that for $s_1, s_2, s_3 \in S$,

$$s_1 + s_3 = 2s_2 \implies s_1 = s_2 = s_3. \tag{3.61}$$

In other words, S has no three elements in arithmetic progression. Clearly $a(x)$, $b(y)$ and $c(z)$ are uniformly distributed and pairwise independent for each x, y, z. Furthermore for $(x, y, z) \in \Psi^n$, tightness implies

$$a(x) + b(y) - 2c(z) = \sum_{i=1}^{n} \left(\alpha(x_i) + \beta(y_i) + \gamma(z_i)\right)\omega_i = 0. \tag{3.62}$$

Since S has no three-term arithmetic progressions,

$$a(x), b(y), c(z) \in S \iff a(x) = b(y) \in S \iff b(y) = c(z) \in S \iff a(x) = c(z) \in S, \tag{3.63}$$

which by pairwise independence and uniformity implies Eq. (3.38), as wanted.

We now need to establish Eq. (3.39). Let $(x, y, z), (x', y', z') \in \Psi_P^n$ be distinct points which agree on either the x, y or z coordinate. We consider the case where $z \neq z'$ and $x = x'$. The other cases are shown in the exact same way and the particular case is chosen for the simple fact that the typographical distance between c and γ is larger than for the other letters in use. By Eq. (3.63), $a(x), b(y), b(y'), c(z), c(z') \in S$ if and only if $a(x) = c(z) = c(z') \in S$. So all we need to show is that $a(x), c(z), c(z')$ are independent. $a(x)$ is independent of the pair $(c(z), c(z'))$, by the fact that this pair does not depend on ω_{n+2}, so it suffices to show that $c(z)$ and $c(z')$ are independent, which is equivalent to showing that $\sum_{i=1}^{n} \gamma(z_i)\omega_i$ and $\sum_{i=1}^{n} \gamma(z_i')\omega_i$ are independent. This will follow as we show that the γ part is linearly independent:

Since z and z' both belong to the same type class, and γ is injective, the vectors $\gamma(z) := (\gamma(z_i))_{i=1}^{n} \in \mathbb{Q}^n$ and $\gamma(x') := (\gamma(z_i'))_{i=1}^{n} \in \mathbb{Q}^n$ also have the same type (i.e. they take values in $\mathbb{Q}$ equally many times), and by injectivity, $\gamma(z) \neq \gamma(z')$. Since $\gamma(z)$ and $\gamma(z')$ have all entries of the same sign (negative, by the assumption at the beginning of the proof), are different, yet of the same type, they must necessarily be *linearly* independent. When M is sufficiently large this implies that $\gamma(z)$ and $\gamma(z')$ are also linearly independent in $\mathbb{Z}_M^n$.

We now return to showing that $\sum_{i=1}^n \gamma(z_i)\omega_i = \langle \gamma(z)|\omega\rangle$ and $\sum_{i=1}^n \gamma(z_i')\omega_i = \langle \gamma(z')|\omega\rangle$ are uniformly and independently distributed. Since $\gamma(z), \gamma(z') \in \mathbb{Z}_M^n$ are linearly independent we may find $A \in \mathrm{GL}_2(\mathbb{Z}_M)$ such that the matrix

$$\begin{bmatrix} v \\ v' \end{bmatrix} = A \begin{bmatrix} \gamma(z) \\ \gamma(z') \end{bmatrix} \in M_{2\times n}(\mathbb{Z}_M), \tag{3.64}$$

has j'th column $\begin{bmatrix} v_j \\ v_j' \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ for some j. It follows that the stochastic variables $\sum_{i=1}^n v_i\omega_i = \langle v|\omega\rangle$ and $\sum_{i=1}^n v_i'\omega_i = \langle v'|\omega\rangle$ expressible as

$$\begin{bmatrix} \langle v|\omega\rangle \\ \langle v'|\omega\rangle \end{bmatrix} = A \begin{bmatrix} \langle \gamma(z)|\omega\rangle \\ \langle \gamma(z')|\omega\rangle \end{bmatrix}, \tag{3.65}$$

are stochastically independent and uniformly distributed, since only $\langle v|\omega\rangle$ depends on ω_j. Since A acts as a permutation on $\mathbb{Z}_M^2$, it follows that since $\begin{bmatrix} \langle v|\omega\rangle \\ \langle v'|\omega\rangle \end{bmatrix}$ is uniformly distributed on $\mathbb{Z}_M^2$, so is $\begin{bmatrix} \langle \gamma(z)|\omega\rangle \\ \langle \gamma(z')|\omega\rangle \end{bmatrix} = A^{-1} \begin{bmatrix} \langle v|\omega\rangle \\ \langle v'|\omega\rangle \end{bmatrix}$, which is equivalent to the variables $\langle \gamma(z)|\omega\rangle$ and $\langle \gamma(z')|\omega\rangle$ being independently and uniformly distributed in $\mathbb{Z}_M$ as wanted. $\square$

Chapter 4

The asymptotic spectrum of LOCC transformations

In the previous chapter, we considered the spectrum $\Delta(\mathcal{T}_k)$ of the ordered semiring, $(\mathcal{T}_k, \odot, \oplus, \geq)$, of tensors under restriction, corresponding to SLOCC channels and pure states. In SLOCC, there is no control on the probability by which the conversion succeeds. If we wish to retain information on this probability when passing to monotones, then for $p \in \mathbb{R}\backslash\{0\}$ we must distinguish between ψ and $p\psi$, which in $\mathcal{T}_k$ belong to the same equivalence class. In this chapter we shall refine the notion of equivalence of tensors, creating a finer division of the class of states. Two tensors ψ and ϕ will be equivalent if and only if their corresponding operators $|\psi\rangle\langle\psi|$, $|\phi\rangle\langle\phi|$ are equivalent under LOCC. In particular they must have the same norm. It turns out that the set of equivalence classes, ordered by LOCC, which will be denoted $\mathcal{S}_k$, is also an ordered semiring, to which Theorem 3.1.4 applies. The asymptotic ordering of Theorem 3.1.4 then gives asymptotic LOCC conversion given a converse error exponent and Theorem 3.1.4 provides a relation between this rate and the spectrum $\Delta(\mathcal{S}_k)$.

In this chapter we consider asymptotic, exact, probabilistic LOCC conversion of pure k-partite states. That is, given many copies of a resource state $|\psi\rangle$, we ask how many copies of a target state $|\phi\rangle$ we can probabilistically obtain through exact LOCC, as a function of the asymptotic behavior of success. In [39] they considered the special bipartite case when the target $|\phi\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right)$ was the maximally entangled state. The optimal exact extraction rate, was described as a function of the asymptotic exponential behavior of the probability of successful transformation: If probability of failure behaves like 2^{-nr}, where n is the number of copies of the resource, r is called the *direct error exponent*. If probability of success behaves like 2^{-nr}, r is called a *converse error exponent*. Given converse error exponent r, the following formula for the concentration rate was given in [39]:

$$E^*(r) = \inf_{\alpha\in[0,1)} \frac{r\alpha + \log\sum_i p_i^\alpha}{1-\alpha}. \tag{4.1}$$

Here $(\sqrt{p_i})_i$ are the Schmidt coefficients of the resource state $|\psi\rangle$. In [14] this formula was extended to multi-partite states and other target states than the maximally entangled state. This chapter largely follows the content of this paper.

For two multipartite pure states, $|\psi\rangle$ and $|\phi\rangle$, we let $E^*_{\mathcal{P}}(r,\psi,\phi)$ be the number of copies of $|\phi\rangle$ that can be asymptotically extracted per copy of $|\psi\rangle$ with success probability behaving like $2^{-nr+o(n)}$. The "P" stands for probabilistic and the $*$ represents the fact that we are considering the converse error exponent, rather than the direct error exponent. We say that the optimal extraction rate from $|\psi\rangle$ to $|\phi\rangle$ with converse error exponent r is $E^*_{\mathcal{P}}(r,\psi,\phi)$. In the light of Proposition 1.1.12, $E^*_{\mathcal{P}}(r,\psi,\phi)$ is formally defined, also for mixed states, as

$$E^*_{\mathcal{P}}(r,\rho,\sigma) = \sup\left\{\tau\in\mathbb{R}^+ \middle| \rho^{\odot n} \xrightarrow{\text{LOCC}} 2^{-nr+o(n)}\sigma^{\odot\lfloor\tau n\rfloor} \text{ for } n\gg 1\right\}. \tag{4.2}$$

As a consequence of Theorem 4.1.4 below we show in Eq. (4.51) that for k parties and globally entangled resource state $|\psi\rangle$, the optimal extraction rate between pure states can be expressed as

$$E^*_{\mathcal{P}}(r,\psi,\phi) = \inf_{f\in\Delta(\mathcal{S}_k)} \frac{r\alpha(f) + \log f(|\psi\rangle)}{\log f(|\phi\rangle)}. \tag{4.3}$$

Here $\Delta(\mathcal{S}_k)$ is the spectrum of a certain partially ordered semiring, which is to be constructed, from Theorem 3.1.4. Concretely $\Delta(\mathcal{S}_k)$ is the set of real, LOCC-monotone functions on the set of unnormalized states, that are additive under direct sum, multiplicative under tensor product and normalized, and $\alpha(f) = \log f\left(\sqrt{2}\,|0\ldots 0\rangle\right)$. $\Delta(\mathcal{S}_k)$ will be called the asymptotic LOCC spectrum. Due to Eq. (4.3), an explicit description of the $\Delta(\mathcal{S}_k)$ would imply a complete understanding of the asymptotic extraction rates given a converse error exponent. In Theorem 4.2.1 a characterization of the functions in $\Delta(\mathcal{S}_k)$ is presented. Note that in the case where the target $|\phi\rangle$ is the normalized GHZ-state $\frac{|0\ldots0\rangle+|1\ldots1\rangle}{\sqrt{2}}$, we have $\log f(|\phi\rangle) = 1-\alpha(f)$, showing the resemblance between Eq. (4.1) and Eq. (4.3) and once the bipartite spectrum is fully described in Section 4.3, we shall see that Eq. (4.1) is the special case of Eq. (4.3) with $k=2$ and $|\phi\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle+|11\rangle\right)$.

4.1. The semiring of unnormalized pure states

We wish to apply Strassen's Theorem 3.1.4 to a semiring of unnormalized pure quantum states with preorder determined by LOCC conversion via trace non-increasing LOCC channels cf. Definition 1.1.10. Since the class of all k-partite states acting on some tensor product of finite dimensional Hilbert spaces is not a set we shall deal with equivalence classes of pure states, just like in Chapter 3. These equivalence classes will be the elements of the semiring.

Definition 4.1.1. *Given $k \in \mathbb{N}$ and finite dimensional Hilbert spaces $\mathcal{H}_1, \mathcal{H}_1', \ldots, \mathcal{H}_k, \mathcal{H}_k'$ we say that two unnormalized states $|\phi\rangle \in \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k$ and $|\psi\rangle \in \mathcal{H}_1' \otimes \cdots \otimes \mathcal{H}_k'$ are locally unitarily equivalent (or LU-equivalent), if there exist partial isometries $U_i : \mathcal{H}_i \to \mathcal{H}_i'$ such that*

$$|\psi\rangle = (U_1 \otimes \cdots \otimes U_k) |\phi\rangle$$

and

$$|\phi\rangle = (U_1^* \otimes \cdots \otimes U_k^*) |\psi\rangle \, .$$

Let $\mathcal{S}_k$ denote the set of local unitary equivalence classes, commonly abbreviated as the LU classes.

To call this local unitary equivalence could be somewhat confusing, since the witnessing operators are partial isometries rather than unitaries, and indeed in most of the literature the U_i's are to be unitaries. The reason the definition uses partial isometries instead is because this corresponds to disregarding the dimension of the ambient spaces. The definitions are morally the same, since if we choose any family of Hilbert spaces $(\mathcal{K}_i)_{i\in[k]}$ that are large enough that $\mathcal{H}_i$ and $\mathcal{H}_i'$ can both be embedded in $\mathcal{K}_i$ for each i, then the images of $|\psi\rangle$ and $|\phi\rangle$ in $\mathcal{K}_1 \otimes \cdots \otimes \mathcal{K}_k$ under the product of these embeddings are unitarily equivalent in the sense of unitaries U_i acting on $\mathcal{K}_i$ if and only if $|\psi\rangle$ and $|\phi\rangle$ are equivalent in the sense of Definition 4.1.1.

In the bipartite case $k = 2$, each LU equivalence class is uniquely represented by its ordered non-zero Schmidt coefficients. In the case $k \geq 3$ characterizing LU classes is a highly non-trivial task. For a characterization of LU classes in the k-qubit case, i.e. when each of the local systems are 2-dimensional, see [5].

Note that for any two representatives, $[|\psi\rangle] = [|\phi\rangle]$, of an element of $\mathcal{S}_k$, the partial isometries witnessing this equivalence define k-step LOCC channels mapping one to the other and back; $|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle \xrightarrow{\text{LOCC}} |\psi\rangle$. In other words, unnormalized states that are locally unitarily equivalent are also LOCC-equivalent. The following preorder on $\mathcal{S}_k$ is therefore well-defined:

$$[|\psi\rangle] \geq [|\phi\rangle] \text{ iff } |\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle \, .$$

By [40, Corollary 1], LOCC equivalence also implies local unitary equivalence. So the above preorder is in fact a partial order. This is not of importance for the theory to work, but still worth noting.

When $|\psi\rangle \in \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k$ and $|\phi\rangle \in \mathcal{H}'_1 \otimes \cdots \otimes \mathcal{H}'_k$ we may take their direct sum and tensor product to get new k-partite states:

$$|\psi\rangle \oplus |\phi\rangle \in (\mathcal{H}_1 \oplus \mathcal{H}'_1) \otimes \cdots \otimes (\mathcal{H}_k \oplus \mathcal{H}'_k),$$

$$|\psi\rangle \odot |\phi\rangle \in (\mathcal{H}_1 \odot \mathcal{H}'_1) \otimes \cdots \otimes (\mathcal{H}_k \odot \mathcal{H}'_k).$$

The direct sum, $|\psi\rangle \oplus |\phi\rangle$, is the sum of the images of $|\psi\rangle$ and $|\phi\rangle$ under the natural inclusions into $(\mathcal{H}_1 \oplus \mathcal{H}'_1) \otimes \cdots \otimes (\mathcal{H}_k \oplus \mathcal{H}'_k)$.

Both sum and product respect local unitary equivalence, turning $(\mathcal{S}_k, \oplus, \odot)$ into a semiring. As mentioned, we wish to apply Theorem 3.1.4 to $(\mathcal{S}_k, \oplus, \odot, \leq)$. For this purpose, what remains to be shown is that $(\mathcal{S}_k, \oplus, \odot, \leq)$ is a preordered semiring and that conditions 1 and 2 of Theorem 3.1.4 are satisfied. We start out by showing that it is a preordered semiring. Eq. (3.3) is immediate, so we proceed to proving Eq. (3.2), which is done in Proposition 4.1.2.

We say that a state $\rho \in \mathcal{B}(\mathcal{H}_1) \otimes \cdots \otimes \mathcal{B}(\mathcal{H}_k) \otimes \mathrm{Diag}(\mathbb{C}^{\mathcal{X}})$ is conditionally pure if it can be written in the form

$$\rho = \sum_{x \in \mathcal{X}} |\phi_x\rangle\langle\phi_x| \otimes |x\rangle\langle x| \,. \tag{4.4}$$

Proposition 4.1.2. *Let $|\psi\rangle, |\phi\rangle$ and $|\eta\rangle$ be unnormalized, k-partite pure states then*

$$|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle \quad \implies \quad |\psi\rangle \oplus |\eta\rangle \xrightarrow{\text{LOCC}} |\phi\rangle \oplus |\eta\rangle \,. \tag{4.5}$$

Proof. Assume that $|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle$ and $|\eta\rangle \in \mathcal{K}_1 \otimes \cdots \otimes \mathcal{K}_k$. By Proposition 1.1.15

$$|\phi\rangle\langle\phi| = \mathrm{Tr}_{reg}\, \Lambda_n \cdots \Lambda_1 \, |\psi\rangle\langle\psi| \,, \tag{4.6}$$

for some trace non-increasing channels $(\Lambda_l)_{l=1}^n$ of the form (1.27). Let J_l, i_l, $f_j : J_l \to J_{l-1}$ and $(K_j^l)_{j \in J_l}$ be the defining objects of Λ_l for $l = 1, \ldots, n$ as in Proposition 1.1.15. Then

$$\Lambda_n \cdots \Lambda_1 \, |\psi\rangle\langle\psi| = \sum_{j \in J_n} a_j \, |\phi\rangle\langle\phi| \otimes |j\rangle\langle j| \tag{4.7}$$

for some $a_j \geq 0$ with $\sum_j a_j = 1$. We wish to show that there exists an LOCC channel Φ such that

$$\Phi \, |\psi \oplus \eta\rangle\langle\psi \oplus \eta| = \sum_{j \in J_n} a_j \, |\phi \oplus \eta\rangle\langle\phi \oplus \eta| \otimes |j\rangle\langle j| \,. \tag{4.8}$$

This is shown by induction on n. The induction hypothesis is; if for any unnormalized states $|\psi\rangle$, $|\phi\rangle$ and $|\eta\rangle$ and any $\Lambda_n, \ldots, \Lambda_1$ of the form Eq. (1.27) such that Eq. (4.7) holds, then there exists Λ' such that Eq. (4.8) holds. Assume the hypothesis holds for $n-1$, we are to show that it holds for n. So assume Eq. (4.7) holds. Now

$$\Lambda_1 |\psi\rangle\langle\psi| = \sum_{j_1\in J_1} |\psi_{j_1}\rangle\langle\psi_{j_1}| \otimes |j_1\rangle\langle j_1| , \tag{4.9}$$

where $|\psi_{j_1}\rangle = K^1_{j_1}|\psi\rangle$ in the notation of Proposition 1.1.15. Now Λ_2 acts on $\sum_{j_1\in J_1} |\psi_{j_1}\rangle\langle\psi_{j_1}|\otimes|j_1\rangle\langle j_1|$, and we may break Λ_2 down to the sum of its conditional actions

$$\Lambda_2 = \sum_{j_1\in J_1} \Lambda_2^{j_1} \tag{4.10}$$

where $\Lambda_2^{j_1} : \rho\otimes|j\rangle\langle j| \mapsto \delta_{j=j_1}\sum_{j_2\in f_2^{-1}(j_1)} K_{j_2}\rho K^*_{j_2}\otimes|j_2\rangle\langle j_2|$. Now $\Lambda_n\ldots\Lambda_2 = \sum_{j_1\in J_1}\Lambda_n\ldots\Lambda_2^{j_1}$ and

$$\Lambda_n\ldots\Lambda_2^{j_1}|\psi_{j_1}\rangle\langle\psi_{j_1}|\otimes|j_1\rangle\langle j_1| = \sum_{j_n\in J_n^{j_1}} a_{j_n}|\phi\rangle\langle\phi|\otimes|j_n\rangle\langle j_n| = \sum_{j_n\in J_n^{j_1}}\frac{a_{j_n}}{c_{j_1}}\left|\sqrt{c_{j_1}}\phi\right\rangle\left\langle\sqrt{c_{j_1}}\phi\right|\otimes|j_n\rangle\langle j_n| , \tag{4.11}$$

where $J_n^{j_1}\subset J_n$ is the subset $J_n^{j_1} = (f_n\circ\ldots\circ f_2)^{-1}(j_1)$ and $c_{j_1} = \sum_{j_n\in J_n^{j_1}} a_{j_n}$. Since $\sum_{j_n\in J_n^{j_1}}\frac{a_{j_n}}{c_{j_1}} = 1$ and since $\Lambda_2^{j_1}$ only acts on a single part of the register, we may apply the induction hypothesis to the states $|\psi_{j_1}\rangle$, $|\phi_{j_1}\rangle = |\sqrt{c_{j_1}}\phi\rangle$ and $|\eta_{j_1}\rangle = |\sqrt{c_{j_1}}\eta\rangle$ with the $n-1$ channels $\Lambda_n,\ldots,\Lambda_2^{j_1}$. The induction hypothesis grants a channel Φ^{j_1} such that

$$\begin{aligned}\Phi^{j_1} : |\psi_{j_1}\oplus\eta_{j_1}\rangle\langle\psi_{j_1}\oplus\eta_{j_1}|\otimes|j_1\rangle\langle j_1| &\mapsto \sum_{j_n\in J_n^{j_1}}\frac{a_{j_n}}{c_{j_1}}|\phi_{j_1}\oplus\eta_{j_1}\rangle\langle\phi_{j_1}\oplus\eta_{j_1}|\otimes|j_n\rangle\langle j_n| \\ &= \sum_{j_n\in J_n^{j_1}} a_{j_n}|\phi\oplus\eta\rangle\langle\phi\oplus\eta|\otimes|j_n\rangle\langle j_n| .\end{aligned} \tag{4.12}$$

Finally, let $\Phi_1 = \tilde{\mathcal{E}}(J_1, f_1)$ be the map acting on $|\psi\oplus\eta\rangle\langle\psi\oplus\eta|$ by the Kraus operators $K^1_{j_1}\oplus\sqrt{c_{j_1}}I$, where $(K^1_{j_1})$ are the Kraus operators of $\Lambda_1 = \mathcal{E}(J_1, f_1)$ and I is the identity operator acting on $\mathcal{K}_{i_1}$. Here i_1 is the index of the Hilbert space on which Λ_1 acts. Φ_1 is trace non-increasing since

$$\sum_{j_1\in J_1}\left(K^1_{j_1}\oplus\sqrt{c_{j_1}}I\right)^*\left(K^1_{j_1}\oplus\sqrt{c_{j_1}}I\right) = \sum_{j_1\in J_1}(K^1_{j_1})^*K^1_{j_1}\oplus\sum_{j_1\in J_1}c_{j_1}I \le I_{\mathcal{H}_{i_1}}\oplus I_{\mathcal{K}_{i_1}} , \tag{4.13}$$

and

$$\Phi_1|\psi\oplus\eta\rangle\langle\psi\oplus\eta| = \sum_{j_1\in J_1}|\psi_{j_1}\oplus\eta_{j_1}\rangle\langle\psi_{j_1}\oplus\eta_{j_1}|\otimes|j_1\rangle\langle j_1| . \tag{4.14}$$

Now

$$\begin{aligned}\left(\sum_{j_1\in J_1}\Phi^{j_1}\right)\Phi_1|\psi\oplus\eta\rangle\langle\psi\oplus\eta| &= \left(\sum_{j_1\in J_1}\Phi^{j_1}\right)\sum_{j_1\in J_1}|\psi_{j_1}\oplus\eta_{j_1}\rangle\langle\psi_{j_1}\oplus\eta_{j_1}|\otimes|j_1\rangle\langle j_1| \\ &= \sum_{j_1\in J_1}\sum_{j_n\in J_n^{j_1}} a_{j_n}|\phi\oplus\eta\rangle\langle\phi\oplus\eta|\otimes|j_n\rangle\langle j_n| \\ &= \sum_{j\in J_n} a_j|\phi\oplus\eta\rangle\langle\phi\oplus\eta|\otimes|j\rangle\langle j| ,\end{aligned} \tag{4.15}$$

finishing the induction step. Tracing out the register results in $|\psi\rangle\oplus|\eta\rangle \xrightarrow{\text{LOCC}} |\phi\rangle\oplus|\eta\rangle$. $\square$

By Remark 3.1.3 it follows that Eq. (3.2) holds for $(\mathcal{S}_k, \oplus, \otimes, \leq)$, which is therefore a pre-ordered semiring.

It remains to be shown that conditions 1 and 2 in 3.1.4 are satisfied. The multiplicative unit in $\mathcal{S}_k$ is represented by the pure state $|0\dots 0\rangle \in \mathbb{C}^{\otimes k}$ and the additive unit is represented by the zero-vector $0 \in \mathbb{C}^{\otimes k}$. $\mathbb{N}$ embeds into $\mathcal{S}_k$ in the following sense: An integer $d \in \mathbb{N}$ is represented in $\mathcal{S}_k$ by the d-level, k-partite, unnormalized GHZ state, which we in Definition 2.1.2 identified with the unit tensor

$$\sqrt{d}\,|\mathrm{GHZ}_d\rangle = |u_d\rangle = \sum_{i=0}^{d-1} |i\dots i\rangle \in (\mathbb{C}^d)^{\otimes k}. \tag{4.16}$$

If $d_1 \geq d_2$, then the one-step trace reducing LOCC channel, defined by the single Kraus operator $K = \sum_{i=0}^{d_2-1} |i\rangle\langle i|$ acting on the first system, witnesses $|u_{d_1}\rangle \xrightarrow{\text{LOCC}} |u_{d_2}\rangle$. And since LOCC channels never increase the trace, we have $|u_{d_1}\rangle \xrightarrow{\text{LOCC}} |u_{d_2}\rangle$ iff $d_1 \geq d_2$, so condition 1 holds.

We proceed by proving that condition 2 holds:

Proposition 4.1.3. *For any non-zero pure states $|\psi\rangle$ and $|\phi\rangle$, there is a $d \in \mathbb{N}$ such that*

$$|u_d\rangle \odot |\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle\,. \tag{4.17}$$

Proof. By having one party locally construct the normalized $|\phi\rangle$, converting GHZ states to EPR pairs between parties [41] [42] and using quantum teleportation [43] [19, s. 6.5.3], one obtains a protocol that extracts the normalized version of $|\phi\rangle$ from $|u_d\rangle$ for large enough d. Furthermore $|\psi\rangle \xrightarrow{\text{LOCC}} \|\psi\|\,|u_1\rangle$. So for sufficiently large d

$$|u_d\rangle \odot |\psi\rangle \xrightarrow{\text{LOCC}} \frac{1}{\|\phi\|}\,|\phi\rangle \odot |\psi\rangle \xrightarrow{\text{LOCC}} \frac{\|\psi\|}{\|\phi\|}\,|\phi\rangle\,. \tag{4.18}$$

In order to obtain $|\phi\rangle$ one simply increases d to dn for large enough n and traces out the GHZ states not used for teleportation:

$$|u_{dn}\rangle \odot |\psi\rangle = |u_n\rangle \odot |u_d\rangle \odot |\psi\rangle \xrightarrow{\text{LOCC}} \frac{\|\psi\|}{\|\phi\|}\,|u_n\rangle \odot |\phi\rangle \xrightarrow{\text{LOCC}} \sqrt{n}\frac{\|\psi\|}{\|\phi\|}\,|\phi\rangle\,. \tag{4.19}$$

And for $n > \frac{\|\phi\|^2}{\|\psi\|^2}$

$$\sqrt{n}\frac{\|\psi\|}{\|\phi\|}\,|\phi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle\,. \tag{4.20}$$

□

Theorem 3.1.4 now applies to $\mathcal{S}_k$ and we get the following:

Theorem 4.1.4. *Let $\Delta(S_k)$ be the set of order preserving semiring homomorphisms $\mathcal{S}_k \to \mathbb{R}^+$. Then*

$$[|\psi\rangle] \gtrsim [|\phi\rangle] \iff \forall f \in \Delta(\mathcal{S}_k) : f\left(|\psi\rangle\right) \geq f\left(|\phi\rangle\right). \tag{4.21}$$

We call $\Delta(\mathcal{S}_k)$ the asymptotic LOCC spectrum.

Concretely $[|\psi\rangle] \gtrsim [|\phi\rangle]$ means that

$$|u_2\rangle^{\odot o(n)} \odot |\psi\rangle^{\odot n} \xrightarrow{\text{LOCC}} |\phi\rangle^{\odot n}, \tag{4.22}$$

where $|u_2\rangle = |0\dots 0\rangle + |1\dots 1\rangle$ is the unnormalized two-level GHZ state. In other words; to extract n copies of $|\phi\rangle$, we need n copies of $|\psi\rangle$, a proportionally vanishing number of GHZ states and the success probability decays as $2^{n(\log\|\psi\|^2 - \log\|\phi\|^2)+o(n)}$. Since we only need a proportionally vanishing number of GHZ states, we may, assuming that $|\psi\rangle$ is globally entangled, extract these GHZ states from $|\phi\rangle^{\odot n}$ without further cost in the asymptotic limit (see Proposition 2.1.7). Indeed, one can show that when $|\psi\rangle$ is globally entangled, $|\psi\rangle^{\odot k} \xrightarrow{\text{LOCC}} x\,|u_2\rangle$ for some $x > 0$, e.g. by extracting EPR-pairs (see [26, Lemma 4]) and using teleportation. That is, for any globally entangled $|\psi\rangle$,

$$\begin{aligned} E^*_{\mathcal{P}}(r,\psi,\phi) &= \sup\left\{\tau \in \mathbb{R}^+ \middle| 2^{nr+o(n)}\,|\psi\rangle\!\langle\psi|^{\odot n} \xrightarrow{\text{LOCC}} |\phi\rangle\!\langle\phi|^{\odot\lfloor\tau n\rfloor} \text{ for } n \gg 1\right\} \\ &= \sup\left\{\tau \in \mathbb{R}^+ \middle| \left[\left(2^{r/2}\,|\psi\rangle\right)^{\odot n}\right] \gtrsim \left[|\phi\rangle^{\odot\lfloor\tau n\rfloor}\right] \text{ for } n \gg 1\right\} \\ &= \sup\left\{\tau \in \mathbb{R}^+ \middle| \forall f \in \Delta(\mathcal{S}_k) : f\left(2^{r/2}\,|\psi\rangle\right) \geq f\left(|\phi\rangle\right)^{\tau}\right\}. \end{aligned} \tag{4.23}$$

It is unclear whether this is also true for states which are separable across some bipartition (i.e. not globally entangled). It is the intuition of the author that this is indeed the case, but a proof has not been found.

4.2. Characterizing the spectrum

The goal of this section is to prove Theorem 4.2.1 below, which establishes a condition for a semiring homomorphism $f : \mathcal{S}_k \to \mathbb{R}^+$ to be monotone (i.e. order preserving) and hence define a point in $\Delta(\mathcal{S}_k)$.

Theorem 4.2.1. *Let $f : \mathcal{S}_k \to \mathbb{R}^+$ be a semiring homomorphism. Then f is monotone if and only if there is an $\alpha \in [0,1]$ such that $f(\sqrt{p}\,|0\dots 0\rangle) = p^{\alpha}$ for all $p > 0$ and*

$$f(|\phi\rangle) \geq \left(f\left(\Pi\,|\phi\rangle\right)^{1/\alpha} + f\left((I-\Pi)\,|\phi\rangle\right)^{1/\alpha}\right)^{\alpha} \tag{4.24}$$

for any $|\phi\rangle \in \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k$, $i \in \{1,\dots,k\}$ and orthogonal projection $\Pi \in \mathcal{B}(\mathcal{H}_i)$.

In the above theorem, and for the duration of this section, the case $\alpha = 0$ should be interpreted in the sense of $\alpha \to 0$. That is, Eq. (4.24) becomes

$$f(|\phi\rangle) \geq \max\left\{f\left(\Pi\,|\phi\rangle\right), f\left(\Pi\,|\psi\rangle\right)\right\}. \tag{4.25}$$

We start by showing that the α of Theorem 4.2.1 necessarily exists.

Proposition 4.2.2. *Let $f : \mathcal{S}_k \to \mathbb{R}^+$ be a monotone semiring homomorphism. There is an $\alpha \geq 0$ such that*

$$f\left(\sqrt{p}\,|\phi\rangle\right) = p^\alpha f\left(|\phi\rangle\right) \tag{4.26}$$

for each $|\phi\rangle$ and each $p > 0$.

Proof. Since $p \mapsto f(\sqrt{p}\,|0\ldots0\rangle)$ is multiplicative, non-decreasing, sends 0 to 0 and 1 to 1, it follows from the solution to the Cauchy functional equation that

$$f\left(\sqrt{p}\,|0\ldots0\rangle\right) = p^\alpha$$

for all $p > 0$ and some $\alpha \geq 0$. Therefore

$$\begin{aligned} f\left(\sqrt{p}\,|\phi\rangle\right) &= f\left(\sqrt{p}\,|\phi\rangle \otimes |0\ldots0\rangle\right) = f\left(|\phi\rangle\right) f\left(\sqrt{p}\,|0\ldots0\rangle\right) = p^\alpha f\left(|\phi\rangle\right) f\left(|0\ldots0\rangle\right) \\ &= p^\alpha f\left(|\phi\rangle\right). \end{aligned}$$

□

For the proof of Theorem 4.2.1 the following extension of a monotone homomorphism $f : \mathcal{S}_k \to \mathbb{R}^+$ to conditionally pure states is introduced: Given f such that $f(\sqrt{p}\,|0\ldots0\rangle) = p^\alpha$ for some $\alpha > 0$, define

$$f\left(\sum_{x\in\mathcal{X}} |\phi_x\rangle\langle\phi_x| \otimes |x\rangle\langle x|\right) = \left(\sum_{x\in\mathcal{X}} f(|\phi_x\rangle)^{1/\alpha}\right)^\alpha,$$

and if $\alpha = 0$, let the extension be defined as

$$f\left(\sum_{x\in\mathcal{X}} |\phi_x\rangle\langle\phi_x| \otimes |x\rangle\langle x|\right) = \max_{x\in\mathcal{X}} f\left(|\phi_x\rangle\right).$$

Proposition 4.2.3. *The extension of f is multiplicative under tensor product.*

Proof. For $\alpha > 0$

$$
\begin{aligned}
&f\left(\left(\sum_{x\in\mathcal{X}} |\phi_x\rangle\langle\phi_x| \otimes |x\rangle\langle x|\right) \odot \left(\sum_{y\in\mathcal{Y}} |\psi_y\rangle\langle\psi_y| \otimes |y\rangle\langle y|\right)\right) \\
&= f\left(\sum_{\substack{x\in\mathcal{X}\\ y\in\mathcal{Y}}} \Big(|\phi_x\rangle\langle\phi_x| \odot |\psi_y\rangle\langle\psi_y|\Big) \otimes |xy\rangle\langle xy|\right) \\
&= \left(\sum_{\substack{x\in\mathcal{X}\\ y\in\mathcal{Y}}} f(|\phi_x\rangle \odot |\psi_y\rangle)^{1/\alpha}\right)^{\alpha} = \left(\sum_{\substack{x\in\mathcal{X}\\ y\in\mathcal{Y}}} f(|\phi_x\rangle)^{1/\alpha} f(|\psi_y\rangle)^{1/\alpha}\right)^{\alpha} \\
&= \left(\sum_{x\in\mathcal{X}} f(|\phi_x\rangle)^{1/\alpha} \sum_{y\in\mathcal{Y}} f(|\psi_y\rangle)^{1/\alpha}\right)^{\alpha} = \left(\sum_{x\in\mathcal{X}} f(|\phi_x\rangle)^{1/\alpha}\right)^{\alpha} \left(\sum_{y\in J} f(|\psi_y\rangle)^{1/\alpha}\right)^{\alpha} \\
&= f\left(\sum_{x\in\mathcal{X}} |\phi_x\rangle\langle\phi_x| \otimes |x\rangle\langle x|\right) f\left(\sum_{y\in\mathcal{Y}} |\psi_y\rangle\langle\psi_y| \otimes |y\rangle\langle y|\right).
\end{aligned}
\tag{4.27}
$$

If $\alpha = 0$, then

$$
\begin{aligned}
&f\left(\left(\sum_{x\in\mathcal{X}} |\phi_x\rangle\langle\phi_x| \otimes |x\rangle\langle x|\right) \odot \left(\sum_{y\in\mathcal{Y}} |\psi_y\rangle\langle\psi_y| \otimes |y\rangle\langle y|\right)\right) \\
=&\max_{\substack{x\in\mathcal{X}\\ y\in\mathcal{Y}}} f\left(|\phi_x\rangle\right) f\left(|\psi_y\rangle\right) = \max_{x\in\mathcal{X}} f\left(|\phi_x\rangle\right) \max_{y\in\mathcal{Y}} f\left(|\psi_y\rangle\right) \\
=&f\left(\sum_{x\in\mathcal{X}} |\phi_x\rangle\langle\phi_x| \otimes |x\rangle\langle x|\right) f\left(\sum_{y\in\mathcal{Y}} |\psi_y\rangle\langle\psi_y| \otimes |y\rangle\langle y|\right).
\end{aligned}
\tag{4.28}
$$

□

Proposition 4.2.4. *If f is monotone, then the extension of f to conditionally pure states is monotone under conditional application of local quantum instruments with components having Kraus rank 1. That is, f is monotone under maps of the form Eq. (1.27) acting on conditionally pure states.*

Proof. First assume $\alpha > 0$ and start with the case where the initial state is pure:

$$
|\psi\rangle\langle\psi| \xrightarrow{\text{LOCC}} \sum_{i\in I} P(i) \, |\phi_i\rangle\langle\phi_i| \otimes |i\rangle\langle i| \, . \tag{4.29}
$$

Here the $|\phi_i\rangle$'s are normalized and $P : I \to \mathbb{R}^+$ is a map.

Recall from Section 1.2, that for $n \in \mathbb{N}$ we say that a probability distribution $Q : I \to \mathbb{R}^+$ is an n-type, if $nQ(i) \in \mathbb{N}$ for each $i \in \mathbb{N}$. Given an n-type Q, we say that a sequence in I^n is of type Q, if i appears $nQ(i)$ times. The type class $T_Q^n \subset I^n$ is the set of sequences of type Q. Given any n-type Q:

$$\begin{aligned}
|\psi\rangle\langle\psi|^{\odot n} &\xrightarrow{\text{LOCC}} \left(\sum_{i\in I} P(i)\,|\phi_i\rangle\langle\phi_i|\otimes|i\rangle\langle i|\right)^{\odot n} \\
&= \sum_{a\in I^n}\prod_{j=1}^{n} P(a_j)\bigodot_{j=1}^{n}\left|\phi_{a_j}\middle\rangle\middle\langle\phi_{a_j}\right|\otimes|a\rangle\langle a| \\
&\xrightarrow{\text{LOCC}} |T_Q^n|2^{-n(H(Q)+D(Q||P))}\bigodot_{i\in I}|\phi_i\rangle\langle\phi_i|^{\odot nQ(i)}\,.
\end{aligned} \tag{4.30}$$

The last LOCC transformation is the projection onto the multi-indices of type Q followed by a unitary reshuffling of indices and a partial trace on the classical register. $H(Q)$ is the Shannon entropy of Q, defined in Definition 1.2.1, and $D(Q||P)$ is the relative entropy as defined in Definition 1.2.2. Since the last expression in Eq. (4.30) is a pure state we can apply monotonicity of f on pure states to get

$$f(|\psi\rangle)^n \geq \left(|T_Q^n|2^{-n(H(Q)+D(Q||P))}\right)^{\alpha}\prod_{i\in I} f(|\phi_i\rangle)^{nQ(i)}. \tag{4.31}$$

By Lemma 1.2.3 $|T_Q^n| \geq 2^{nH(Q)-|I|\log(n+1)}$ this implies, by taking the n-th root of the above expression:

$$f(|\psi\rangle) \geq \left(2^{-D(Q||P)+\sum_i Q(i)\log f(|\phi_i\rangle)^{1/\alpha}}\right)^{\alpha} 2^{-\alpha|I|\frac{\log(n+1)}{n}}. \tag{4.32}$$

Let $Z = \sum_{i\in I} P(i) f(|\phi_i\rangle)^{1/\alpha}$ and let P_ϕ be the probability distribution $P_\phi(i) = \frac{P(i)f(|\phi_i\rangle)^{1/\alpha}}{Z}$. Then

$$-D\left(Q||P\right) + \sum_i Q(i)\log f(|\phi_i\rangle)^{1/\alpha} = -D\left(Q\middle|\middle|ZP_\phi\right) = \log Z - D\left(Q\middle|\middle|P_\phi\right). \tag{4.33}$$

Using Eq. (4.33), Eq. (4.32) becomes

$$f(|\psi\rangle) \geq 2^{\left(\log Z - D(Q\|P_\phi)\right)\alpha} 2^{-\alpha|I|\frac{\log(n+1)}{n}}. \tag{4.34}$$

For each $n\in\mathbb{N}$, let Q_n be an n-type with $\mathrm{supp}Q_n = \mathrm{supp}P_\phi$ such that $\lim_n Q_n = P_\phi$. Then $D(Q_n\|P_\phi) \to D(P_\phi\|P_\phi) = 0$. Inserting Q_n in Eq. (4.34) and letting $n\to\infty$ yields

$$f(|\psi\rangle) \geq Z^\alpha = \left[\sum_{i\in I} P(i) f(|\phi_i\rangle)^{1/\alpha}\right]^{\alpha} = f\left(\sum_{i\in I} P(i)\,|\phi_i\rangle\langle\phi_i|\otimes|i\rangle\langle i|\right), \tag{4.35}$$

showing that the extension is monotone under remembering one-step LOCC channels applied to pure states. We use this result to generalize to LOCC channels of the form Eq. (1.27) on conditionally pure states. The action of such channels on a conditionally pure state always looks like this:

$$\sum_{x\in\mathcal{X}} |\psi_x\rangle\langle\psi_x|\otimes|x\rangle\langle x| \xrightarrow{\text{LOCC}} \sum_{x\in\mathcal{X}}\sum_{j\in g^{-1}(x)} |\phi_j\rangle\langle\phi_j|\otimes|j\rangle\langle j|\,, \tag{4.36}$$

where $|\phi_j\rangle = K_j |\phi_{g(j)}\rangle$, in the notation of Proposition 1.1.15 (except for the fact that we use g in place of f, which is already in use). By restricting the protocol to only the Kraus operators acting on $|\psi_x\rangle$ one gets for each $x \in \mathcal{X}$

$$|\psi_x\rangle\langle\psi_x| \xrightarrow{\text{LOCC}} \sum_{j\in g^{-1}(x)} |\phi_j\rangle\langle\phi_j| \otimes |j\rangle\langle j| \,. \tag{4.37}$$

Therefore

$$\begin{aligned} f\left(\sum_{x\in\mathcal{X}} |\psi_x\rangle\langle\psi_x| \otimes |x\rangle\langle x|\right) &= \left(\sum_{x\in\mathcal{X}} f(|\psi_x\rangle)^{1/\alpha}\right)^{\alpha} \\ &\geq \left(\sum_{x\in\mathcal{X}} \sum_{j\in g^{-1}(x)} f(|\phi_j\rangle)^{1/\alpha}\right)^{\alpha} \\ &= f\left(\sum_{x\in\mathcal{X}} \sum_{j\in g^{-1}(x)} |\phi_j\rangle\langle\phi_j| \otimes |j\rangle\langle j|\right). \end{aligned}$$

For the case $\alpha = 0$, note that

$$|\psi_x\rangle\langle\psi_x| \xrightarrow{\text{LOCC}} \sum_{j\in g^{-1}(x)} |\phi_j\rangle\langle\phi_j| \otimes |j\rangle\langle j| \tag{4.38}$$

implies $|\psi_x\rangle \xrightarrow{\text{LOCC}} |\phi_j\rangle$ for each $j \in g^{-1}(x)$, which by monotonicity of f on pure states implies $f(|\psi_x\rangle) \geq \max_{j\in g^{-1}(x)} f(|\phi_j\rangle)$. Therefore

$$\begin{aligned} f\left(\sum_{x\in\mathcal{X}} |\psi_x\rangle\langle\psi_x| \otimes |x\rangle\langle x|\right) &= \max_{x\in\mathcal{X}} f(|\psi_x\rangle) \\ &\geq \max_{j\in J} f(|\phi_j\rangle) \\ &= f\left(\sum_{x\in\mathcal{X}} \sum_{j\in g^{-1}(x)} |\phi_j\rangle\langle\phi_j| \otimes |j\rangle\langle j|\right). \end{aligned}$$

□

Remark 4.2.5. *It is not so hard to see that if ρ is conditionally pure and Π_g is a coarse-graining such that $\Pi_g\rho$ is also conditionally pure, then the extension of f also satisfies $f(\rho) \geq f(\Pi_g\rho)$. By Proposition 4.2.4 this implies that when $\rho \xrightarrow{\text{LOCC}} \sigma$, for conditionally pure states, then $f(\rho) \geq f(\sigma)$.*

Lemma 4.2.6. *Let $f : \mathcal{S}_k \to \mathbb{R}^+$ be a semiring homomorphism with $f(\sqrt{p}\,|0\ldots0\rangle) = p^\alpha$ for some $\alpha \in [0,1]$ which satisfies Eq. (4.24) for any choice of pure state and orthogonal projection. Then*

$$f\Big(|\phi\rangle\Big) \geq \left(f\Big(A|\phi\rangle\Big)^{1/\alpha} + f\Big(B|\phi\rangle\Big)^{1/\alpha}\right)^{\alpha} \tag{4.39}$$

for any $|\phi\rangle \in \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k$, $i \in \{1,\ldots,k\}$ and $A, B \in \mathcal{B}(\mathcal{H}_i)$ with $A^*A + B^*B \leq I$.

Proof. Consider the operator $U = \begin{bmatrix} A \\ B \\ \sqrt{I - A^*A - B^*B} \end{bmatrix} : \mathcal{H}_i \to \mathcal{H}_i^3$. This is an isometry, so $f(|\phi\rangle) = f(|\psi\rangle)$, where $|\psi\rangle = U|\phi\rangle$. Let $\Pi : \mathcal{H}_i^3 \to \mathcal{H}_i^3$ be the projection onto the first summand. Then $[\Pi|\psi\rangle] = [A|\phi\rangle]$ and $[(I-\Pi)|\psi\rangle] \geq [B|\phi\rangle]$, so

$$\begin{aligned} f\Big(|\phi\rangle\Big) = f(|\psi\rangle) &\geq \left(f\Big(\Pi|\psi\rangle\Big)^{1/\alpha} + f\Big((I-\Pi)|\psi\rangle\Big)^{1/\alpha}\right)^{\alpha} \\ &\geq \left(f\Big(A|\phi\rangle\Big)^{1/\alpha} + f\Big(B|\phi\rangle\Big)^{1/\alpha}\right)^{\alpha}. \end{aligned} \tag{4.40}$$

□

Proof of Theorem 4.2.1. Suppose f is monotone, then by Proposition 4.2.2 there is an $\alpha \geq 0$ such that $f(\sqrt{p}|\phi\rangle) = p^{\alpha} f(|\phi\rangle)$ for all $|\phi\rangle$ and $p > 0$. Consider the extension of f to conditionally pure states. Let $|\phi\rangle$ and Π be given as in the statement of the theorem, then

$$|\phi\rangle\langle\phi| \xrightarrow{\text{LOCC}} \Pi|\phi\rangle\langle\phi|\Pi \otimes |0\rangle\langle 0| + (I-\Pi)|\phi\rangle\langle\phi|(I-\Pi) \otimes |1\rangle\langle 1|, \tag{4.41}$$

so by monotonicity of the extension of f we get

$$f\Big(|\phi\rangle\Big) \geq \left(f\Big(\Pi|\phi\rangle\Big)^{1/\alpha} + f\Big((I-\Pi)|\phi\rangle\Big)^{1/\alpha}\right)^{\alpha}.$$

When $|\phi\rangle = |0\dots 0\rangle + |1\dots 1\rangle$, $\Pi = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $I - \Pi = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, we get by Eq. (4.24)

$$2 = f\Big(|\phi\rangle\Big) \geq \left(f\Big(\Pi|\phi\rangle\Big)^{1/\alpha} + f\Big((I-\Pi)|\phi\rangle\Big)^{1/\alpha}\right)^{\alpha} = \left(f\Big(|0\dots 0\rangle\Big)^{1/\alpha} + f\Big(|1\dots 1\rangle\Big)^{1/\alpha}\right)^{\alpha} = 2^{\alpha}, \tag{4.42}$$

showing that $\alpha \leq 1$. This concludes the proof of the "only if" statement.

Conversely, suppose f is a homomorphism satisfying Eq. (4.24). By Lemma 4.2.6, f satisfies Eq. (4.39). Consider the extension of f to conditionally pure states. By Proposition 1.1.15 we need only check that f is monotone under conditional application of local instruments with components of Kraus rank 1 Eq. (1.27), and when tracing out the register of a state of the form $\sum_i a_i |\phi\rangle\langle\phi| \otimes |i\rangle\langle i|$. f is monotone under the latter, since

$$\left(\sum_i f\Big(\sqrt{a_i}|\phi\rangle\Big)^{1/\alpha}\right)^{\alpha} = \Big(\sum a_i\Big)^{\alpha} f(|\phi\rangle) = f\Big(\sum_i \sqrt{a_i}|\phi\rangle\Big). \tag{4.43}$$

Now for monotonicity under conditional application of rank 1 instruments. Like in the proof of Proposition 4.2.4 we first consider the case when the initial state is pure. That is, we need to show:

$$f\Big(|\phi\rangle\Big) \geq \left(\sum_{j=1}^{d} f\Big(K_j |\phi\rangle\Big)^{1/\alpha}\right)^{\alpha} \tag{4.44}$$

whenever $\sum_j K_j^* K_j \leq I$, for some local maps $(K_j)_j^d$. Assume for the sake of induction that Eq. (4.44) is true for $d-1$ and let $(K_j)_{j=1}^d$ be some Kraus operators with $\sum_j K_j^* K_j \leq I$. Set

$$A = \sqrt{\sum_{j=1}^{d-1} K_j^* K_j}, \tag{4.45}$$

$$B = K_d, \tag{4.46}$$

and

$$\tilde{K}_j = K_j A^{-1} \quad j = 1, \dots, d-1. \tag{4.47}$$

Here A^{-1} denotes the Moore–Penrose pseudoinverse. Since

$$\sum_{j=1}^{d-1} \tilde{K}_j^* \tilde{K}_j = A^{-1} \sum_{j=1}^{d-1} K_j^* K_j A^{-1} = A^{-1} A^2 A^{-1} \leq I, \tag{4.48}$$

we may apply the induction hypothesis to the operators $(\tilde{K}_j)_j^{d-1}$ and the vector $A|\phi\rangle$ to obtain

$$\begin{aligned} f\Big(|\phi\rangle\Big) &\geq \left(f\Big(A|\phi\rangle\Big)^{1/\alpha} + f\Big(B|\phi\rangle\Big)^{1/\alpha}\right) \\ &\geq \left(\left(\left(\sum_{j=1}^{d-1} f\Big(\tilde{K}_j A|\phi\rangle\Big)^{1/\alpha}\right)^{\alpha}\right)^{1/\alpha} + f\Big(B|\phi\rangle\Big)^{1/\alpha}\right)^{\alpha} \\ &= \left(\sum_{j=1}^{d-1} f\Big(K_j|\phi\rangle\Big)^{1/\alpha} + f\Big(K_d|\phi\rangle\Big)^{1/\alpha}\right)^{\alpha} \\ &= \left(\sum_{j=1}^{d} f\Big(K_j|\phi\rangle\Big)^{1/\alpha}\right)^{\alpha}, \end{aligned} \tag{4.49}$$

finishing the induction step.

Just like in the proof of Proposition 4.2.4, this extends to conditionally pure states. Throughout this proof it has been assumed that $\alpha \neq 0$. For the case $\alpha = 0$, the proof is similar, and obtained by simply replacing every instance of α-anti-norm [44] $\left(\sum_j x_j^{\alpha}\right)^{1/\alpha}$ by $\max_j x_j$ in the above proof. □

Note that for $\alpha = 0$ an LOCC spectral point is in fact a point in the asymptotic spectrum of tensors $\Delta(\mathcal{S}_k)$, see Section 4.4. For $\alpha = 1$ there is just one spectral point, the norm squared:

Proposition 4.2.7. *Let $f : \mathcal{S}_k \to \mathbb{R}$ be a monotone semiring homomorphism with $f(\sqrt{p}\,|0\dots0\rangle) = pf(|0\dots0\rangle)$ for $p > 0$, then*

$$f(|\phi\rangle) = \langle\phi|\phi\rangle.$$

Proof. Given $|\phi\rangle$ of norm 1 we have $\frac{1}{\sqrt{d}}|u_d\rangle \xrightarrow{\text{LOCC}} |\phi\rangle \xrightarrow{\text{LOCC}} |0\ldots0\rangle$ for sufficiently large d. Furthermore

$$f\left(\frac{1}{\sqrt{d}}|u_d\rangle\right) = \frac{1}{d}\sum_{i=1}^{d} f(|i\ldots i\rangle) = \frac{1}{d}\sum_{i=1}^{d} f(|0\ldots0\rangle) = f(|0\ldots0\rangle), \tag{4.50}$$

showing that $f(|\phi\rangle) = f(|0\ldots0\rangle) = 1$. So $f(\sqrt{p}\,|\phi\rangle) = p$ for $p > 0$. $\square$

By Proposition 4.2.2 we can pull scalings of unnormalized states outside spectral evaluation, which allows us to reformulate Eq. (4.23). For globally entangled $|\psi\rangle$:

$$\begin{aligned} E^*_{\mathcal{P}}(r,\psi,\phi) &= \sup\left\{\tau\in\mathbb{R}^+ \middle| \forall f\in\Delta(\mathcal{S}_k) : f(2^{r/2}|\psi\rangle) \geq f(|\phi\rangle)^\tau\right\} \\ &= \sup\left\{\tau\in\mathbb{R}^+ \middle| \forall f\in\Delta(\mathcal{S}_k) : r\alpha(f) + \log f(|\psi\rangle) \geq \tau\log f(|\phi\rangle)\right\} \\ &= \inf_{f\in\Delta(\mathcal{S}_k)} \frac{r\alpha(f)+\log f(|\psi\rangle)}{\log f(|\phi\rangle)}. \end{aligned} \tag{4.51}$$

Here $\alpha(f) = \log f\left(\sqrt{2}\,|0\ldots0\rangle\right)$ is the α from Theorem 4.2.1. For resources which are not globally entangled, the formula expresses the extraction rate, provided a proportionately vanishing amount of entanglement shared between each pair of parties.

4.3. Example: Bipartite states and $\Delta(\mathcal{S}_2)$

When $k = 2$, we may, by the Schmidt decomposition, write any element in $\mathcal{S}_2$ as a finite direct sum of terms of the form $\sqrt{p}\,|00\rangle$. Therefore any monotone semiring homomorphism, f, is entirely determined by the value of $\alpha(f)\in[0,1]$: For $|\phi\rangle = |\psi_P\rangle = \sum_i \sqrt{P(i)}\,|ii\rangle$ a monotone semiring homomorphism, f, must be given by

$$f(|\phi\rangle) = \sum_i P(i)^\alpha = \operatorname{Tr}\left[(\operatorname{Tr}_2|\phi\rangle\langle\phi|)^\alpha\right], \tag{4.52}$$

where Tr_2 is the partial trace of the second system.

The question to answer is then: For which $\alpha\in[0,1]$ does $f_\alpha : |\phi\rangle \mapsto \operatorname{Tr}\left[(\operatorname{Tr}_2|\phi\rangle\langle\phi|)^\alpha\right]$ satisfy equation Eq. (4.24). The answer is all of them.

Theorem 4.3.1. $\Delta(\mathcal{S}_2) = \{f_\alpha | \alpha\in[0,1]\}$ *where*

$$f_\alpha : |\phi\rangle \mapsto \operatorname{Tr}\left[(\operatorname{Tr}_2|\phi\rangle\langle\phi|)^\alpha\right].$$

Proof. When $\alpha = 0$, $f_\alpha(|\psi\rangle)$ is the Schmidt rank, which is monotone. Assume instead that $\alpha\in(0,1]$. Let $|\phi\rangle\in\mathbb{C}^d\otimes\mathbb{C}^d$ and $\Pi\in\mathcal{B}(\mathbb{C}^d)$ be an orthogonal projection. It suffices to verify Eq. (4.24) for projections acting on the first system. Let $X\in\mathcal{B}(\mathbb{C}^d)$ be such that

$$|\phi\rangle = \sum_{i=1}^{d} X|i\rangle\otimes|i\rangle.$$

Since the coefficients of $|\phi\rangle$ are the square roots of the eigenvalues of $\operatorname{Tr}_2 |\phi\rangle\langle\phi|$ and

$$\operatorname{Tr}_2 |\phi\rangle\langle\phi| = \sum_{i=1}^{d} X\,|i\rangle\langle i|\,X^* = XX^*,$$

Eq. (4.24) is equivalent to

$$[\operatorname{Tr}(XX^*)^\alpha]^{1/\alpha} \geq [\operatorname{Tr}(\Pi XX^*\Pi)^\alpha]^{1/\alpha} + [\operatorname{Tr}((I-\Pi)XX^*(I-\Pi))^\alpha]^{1/\alpha}.$$

Since YY^* and Y^*Y always have the same eigenvalues we may formulate it instead as

$$[\operatorname{Tr}(X^*X)^\alpha]^{1/\alpha} \geq [\operatorname{Tr}(X^*\Pi X)^\alpha]^{1/\alpha} + [\operatorname{Tr}(X^*(I-\Pi)X)^\alpha]^{1/\alpha}.$$

For $\alpha = 1$ this inequality holds since $X^*X = X^*IX = X^*(\Pi+(I-\Pi))X = X^*\Pi X + X^*(I-\Pi)X$. For $\alpha \in (0,1)$ it follows from [44, Proposition 3.7]. □

Note that the topology on $\Delta(\mathcal{S}_2)$ as described in Theorem 3.1.4 is the Euclidean topology on $[0,1]$, such that $\Delta(\mathcal{S}_2)$ can topologically be identified with the unit interval.

Since $\Delta(\mathcal{S}_2)$ is known we get the following formula for the asymptotic extraction rate between normalized states given converse error exponent r, using the α-Rényi entropy from Definition 1.2.1.

$$E^*_{\mathcal{P}}(r, \psi_P, \psi_Q) = \inf_{\alpha\in[0,1)} \frac{r\alpha + \log\sum P(i)^\alpha}{\log\sum Q(i)^\alpha} = \inf_{\alpha\in[0,1)} \frac{r\frac{\alpha}{1-\alpha} + H_\alpha(P)}{H_\alpha(Q)}. \tag{4.53}$$

When $|\psi_Q\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is the maximally mixed state we retrieve the result [39, eq. (114)] mentioned in Eq. (4.1):

$$E^*_{\mathcal{P}}(r, \psi_P, \psi_Q) = \inf_{\alpha\in[0,1)} \frac{r\alpha + \log\sum P(i)^\alpha}{1-\alpha}. \tag{4.54}$$

4.4. Known points in the asymptotic spectrum of LOCC transformation for 3-partite states

We finish this chapter by reviewing what we know about the LOCC spectral points in the tripartite case. Firstly, note that because $|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle \implies |\psi\rangle \xrightarrow{\text{SLOCC}} |\phi\rangle$ for any unnormalized states $|\psi\rangle$ and $|\phi\rangle$, we have a surjective, order preserving semiring homomorphism $\mathcal{S}_k \to \mathcal{T}_k$, where $\mathcal{T}_k$ is the semiring of tensors discussed in Chapter 3. This implies that any $f \in \Delta(\mathcal{T}_k)$, can naturally be considered as an element of $\Delta(\mathcal{S}_k)$ by simply composing with the map $\mathcal{S}_k \to \mathcal{T}_k$. The $\alpha(f) \in [0,1]$ from Theorem 4.2.1 for such spectral points is necessarily $\alpha(f) = 0$ as these spectral points are invariant under scaling. On the other hand, $|\psi\rangle \xrightarrow{\text{SLOCC}} |\phi\rangle$, then $|\psi\rangle \xrightarrow{\text{LOCC}} p\,|\phi\rangle$ for some $p > 0$. So if $f \in \Delta(\mathcal{S}_k)$ such that $\alpha(f) = 0$, then f respects the SLOCC ordering of $\mathcal{T}_k$. In this way, the $\alpha(f) = 0$ part of $\Delta(\mathcal{S}_k)$ is simply $\Delta(\mathcal{T}_k)$:

$$\Delta(\mathcal{T}_k) \subset \Delta(\mathcal{S}_k) \tag{4.55}$$

In the $k = 3$ case we therefore might consider the quantum functionals F_θ of Theorem 3.2.5 also as members of $\Delta(\mathcal{S}_k)$.

For any partition $\bigsqcup_{j\in J} I_j = [k]$ the corresponding flattening map $\bigotimes_{i\in[k]} \mathcal{H}_i \to \bigotimes_{j\in J} \left(\bigodot_{i\in I_j} \mathcal{H}_i\right)$ induces an order preserving semiring homomorphism $\mathcal{S}_k \mapsto \mathcal{S}_{|J|}$. Composing with this homomorphism, gives a map $\Delta(\mathcal{S}_{|J|}) \hookrightarrow \Delta(\mathcal{S}_k)$. In the special case of $k = 3$ we have three fattenings onto bipartite systems, by grouping two of the three tensor legs together. So from Theorem 4.3.1 it follows that

$$f^i_\alpha : |\psi\rangle \mapsto \operatorname{Tr}\left[(\operatorname{Tr}_{[3]\setminus\{i\}} |\phi\rangle\langle\phi|)^\alpha\right], \tag{4.56}$$

also belongs to $\Delta(\mathcal{S}_3)$, for $\alpha \in [0, 1]$ and $i \in [3]$. Note that for $\alpha = 0$, $f^i_0 = F_{\theta_i}$, where θ_i is the distribution on [3] with all it's weight at i, i.e. $\theta_i(j) = \delta_{j=i}$. For $\alpha = 1$, $f^i_1(|\psi\rangle) = \|\psi\|^2$, independent of i. Perhaps Theorem 4.2.1 will be useful in finding other points of $\Delta(\mathcal{S}_3)$ in the future.

Chapter 5

Asymptotic majorization of probability distributions

In Chapter 4 we considered conversion rates given a converse error exponent for pure states.

$$E^*_{\mathcal{P}}(r, \rho, \sigma) = \sup \left\{ \tau \in \mathbb{R}^+ \middle| 2^{nr+o(n)} \rho^{\odot n} \xrightarrow{\text{LOCC}} \sigma^{\odot \lfloor \tau n \rfloor} \text{ for } n \gg 1 \right\}. \tag{5.1}$$

In this chapter we consider the case when the probability of success goes to 1 rather than to 0

$$E_{\mathcal{P}}(\rho, \sigma) = \sup \left\{ \tau \in \mathbb{R}^+ \middle| (1 + o(1))\, \rho^{\odot n} \xrightarrow{\text{LOCC}} \sigma^{\odot \lfloor \tau n \rfloor} \text{ for } n \gg 1 \right\}, \tag{5.2}$$

and the conversion rate when success probability equals 1 for large n

$$E_{exact}(\rho, \sigma) = \sup \left\{ \tau \in \mathbb{R}^+ \middle| \rho^{\odot n} \xrightarrow{\text{LOCC}} \sigma^{\odot \lfloor \tau n \rfloor} \text{ for } n \gg 1 \right\}. \tag{5.3}$$

Again we only consider the pure case and in this chapter only the bipartite case. The connection between this chapter and entanglement transformation rests on Nielsen's theorem.

Theorem 5.0.1 (Nielsen, [45])**.** *Let P and Q be finitely supported probability distributions. Then*

$$|\psi_P\rangle \xrightarrow{\text{LOCC}} |\psi_Q\rangle \iff P \preceq Q. \tag{5.4}$$

Here $P \preceq Q$ reads as Q majorizes P, as defined in Definition 5.1.1 below. So majorization of probability distributions is inverse to the LOCC ordering of the corresponding pure bipartite states.

Majorization of probability distributions (see Definition 5.1.1) is an important notion in the field of information theory. Given probability distributions P and Q, we ask whether $P^{\otimes n} \preceq Q^{\otimes n}$ for large n, and we ask how large $r \in \mathbb{R}$ is allowed to be for $P^{\otimes n} \preceq Q^{\otimes \lfloor nr \rfloor}$ to be true for large n. We denote the supremum of such r by $E(P, Q)$. Since the squared Schmidt coefficients of

$|\psi_P\rangle^{\odot n}$ are $P^{\otimes n}$, this implies, by Nielsen's theorem, that in the context of LOCC conversion of bipartite pure states, $E(P,Q)$ is the optimal rate at which one can extract exact copies of a pure state with squared Schmidt coefficients Q from copies of a pure state with squared Schmidt coefficients P:

$$E_{exact}\left(\psi_P, \psi_Q\right) = E(P,Q). \tag{5.5}$$

In Theorem 5.2.9, one of the main result of this chapter, it is shown that

$$E(P,Q) = \min_{\alpha\in[0,\infty]} \frac{H_\alpha(P)}{H_\alpha(Q)}. \tag{5.6}$$

This formula was conjectured in [16, Example 8.26].

That the rate cannot be larger follows from the well-known fact that H_α is Schur-concave (see Proposition 5.2.6). The main tool for showing that the rate is attainable is a description of the growth exponents defined in Definition 5.1.2. This description is found in Proposition 5.1.5. One immediate consequence of Eq. (5.6) is that the asymptotic resource theory of exact entanglement transformations is irreversible, in the sense that $E(P,Q)E(Q,P) < 1$ for generic P and Q.

Using the method for showing Eq. (5.6) and the result from the previous chapter in Eq. (4.53), we arrive at the second main result of this chapter: A formula for the conversion rate, for exact, *probabilistic* LOCC transformations of bipartite pure quantum states with success probability going to 1, which is denoted by $E_{\mathcal{P}}$, and defined in Eq. (5.2) above. This will be done in Section 5.3. The formula for $E_{\mathcal{P}}$ is the same as Eq. (5.6), except that the minimum is taken only over the interval $\alpha \in [0,1]$.

$$E_{\mathcal{P}}\left(\psi_P, \psi_Q\right) = \min_{\alpha\in[0,1]} \frac{H_\alpha(P)}{H_\alpha(Q)}. \tag{5.7}$$

The attainability of this probabilistic rate rests on the techniques for showing Eq. (5.6). The fact that it is optimal follows from Eq. (4.53). The two rates $E_{\mathcal{P}}$ and E then coincide precisely when the minimal ratio of Rényi entropies is obtained at $\alpha \in [0,1]$, which is sometimes the case.

Other resource theories where majorization plays a role include the resource theory of coherence, of purity and thermodynamics (see e.g. [46] [47] [48] [49]). It might be possible to interpret the results of this chapter in those contexts. In particular the result Eq. (5.6) bears some resemblance with the result in [50] on catalytic majorization. The exact relation between the two results, would be interesting to study further.

5.1. Asymptotic exponents

Given a probability distribution $P : \mathcal{X} \to [0,1]$ with finite support $\left|\operatorname{supp}(P)\right| = d$, we let $P^{\downarrow} : [d] = \{1, \ldots, d\} \to [0,1]$ be P ordered non-increasingly. We may naturally extend to $P^{\downarrow} : \mathbb{N} \to [0,1]$ by letting $P(i) = 0$ for $i > d$. Like in previous chapters, all probability distributions will have finite support.

Definition 5.1.1. *Given two probability distributions P, Q, we say that Q majorizes P, written $P \preceq Q$, if*

$$\sum_{i=1}^{N} P^{\downarrow}(i) \leq \sum_{i=1}^{N} Q^{\downarrow}(i), \tag{5.8}$$

for all $N \in \mathbb{N}$.

For $n \in \mathbb{N}$, $P^{\otimes n} : \mathcal{X}^n \to [0,1]$ is the n'th product distribution given by $P^{\otimes n}(I) = \prod_{j=1}^{n} P(I_j)$. We wish to study majorization of $P^{\otimes n}$ by $Q^{\otimes n}$ for large n. To this end, given a value v, we are interested in the size of the set of multiindicies I, such that $P^{\otimes n}(I) \geq v$ and the sum of these probabilities. In order to asymptotically compare these for different probability distributions, it is useful to let v depend exponentially on n and look at asymptotic growth rates.

Definition 5.1.2. *For $V \in \left[\log P(d), \log P(1)\right]$ let*

$$m_n^P(V) = \sum_{\substack{I \in [d]^n \\ P^{\otimes n}(I) \geq 2^{nV}}} P^{\otimes n}(I), \tag{5.9}$$

$$m_{n*}^P(V) = \sum_{\substack{I \in [d]^n \\ P^{\otimes n}(I) \leq 2^{nV}}} P^{\otimes n}(I), \tag{5.10}$$

$$s_n^P(V) = \left|\left\{I \in [d]^n \middle| P^{\otimes n}(I) \geq 2^{nV}\right\}\right|, \tag{5.11}$$

$$s_{n*}^P(V) = \left|\left\{I \in [d]^n \middle| P^{\otimes n}(I) \leq 2^{nV}\right\}\right|. \tag{5.12}$$

We define asymptotic exponents of these functions as follows:

$$M^P(V) = \lim_{n \to \infty} \frac{1}{n} \log m_n^P(V), \tag{5.13}$$

$$M_*^P(V) = \lim_{n \to \infty} \frac{1}{n} \log m_{n*}^P(V), \tag{5.14}$$

$$S^P(V) = \lim_{n \to \infty} \frac{1}{n} \log s_n^P(V), \tag{5.15}$$

$$S_*^P(V) = \lim_{n \to \infty} \frac{1}{n} \log s_{n*}^P(V). \tag{5.16}$$

It is not immediately clear that the limits describing M^P, M^P_*, S^P and S^P_* are well defined, but this will follow from Lemma 5.1.3 below. The letters chosen stand for value, mass and size. V, M^P and S^P might be called the value, mass and size exponents, respectively. M^P_* and S^P_* might then be called the converse mass and size exponents. The first step to creating tangible formulas for these exponents is to relate them to the sizes of type classes.

Lemma 5.1.3. *Given a probability distribution $P = P^{\downarrow}$ with $\lvert \mathrm{supp}(P) \rvert = d$ and $V \in \left[\log P(d), \log P(1)\right]$,*

$$M^P(V) = \max_{-H(Q)-D(Q||P)\geq V} -D(Q||P), \tag{5.17}$$

$$M^P_*(V) = \max_{-H(Q)-D(Q||P)\leq V} -D(Q||P), \tag{5.18}$$

$$S^P(V) = \max_{-H(Q)-D(Q||P)\geq V} H(Q), \tag{5.19}$$

$$S^P_*(V) = \max_{-H(Q)-D(Q||P)\leq V} H(Q). \tag{5.20}$$

Proof. We show Eq. (5.19), as the other equations are shown in the same way: By Eq. (1.43)

$$s^P_n(V) = \left| \left\{ I \in [d]^n \middle| P^{\otimes n}(I) \geq 2^{nV} \right\} \right| = \sum_{\substack{Q \in \mathcal{P}_n \\ -H(Q)-D(Q||P)\geq V}} |T^n_Q|. \tag{5.21}$$

By Eq. (1.40) and Lemma 1.2.3,

$$\begin{aligned} \sum_{\substack{Q \in \mathcal{P}_n \\ -H(Q)-D(Q||P)\geq V}} |T^n_Q| &\leq (n+1)^d \max_{\substack{Q \in \mathcal{P}_n \\ -H(Q)-D(Q||P)\geq V}} |T^n_Q| \\ &\leq (n+1)^d \max_{\substack{Q \in \mathcal{P}_n \\ -H(Q)-D(Q||P)\geq V}} 2^{nH(Q)} \\ &\leq \max_{\substack{Q \in \mathcal{P} \\ -H(Q)-D(Q||P)\geq V}} (n+1)^d 2^{nH(Q)}. \end{aligned} \tag{5.22}$$

Applying "$\lim_{n\to\infty} \frac{1}{n} \log$" to both sides shows that $S^P(V) \leq \max_{-H(Q)-D(Q||P)\geq V} H(Q)$. Similarly, by applying the other inequality from Lemma 1.2.3,

$$\begin{aligned} \sum_{\substack{Q \in \mathcal{P}_n \\ -H(Q)-D(Q||P)\geq V}} |T^n_Q| &\geq \max_{\substack{Q \in \mathcal{P}_n \\ -H(Q)-D(Q||P)\geq V}} |T^n_Q| \\ &\geq \max_{\substack{Q \in \mathcal{P}_n \\ -H(Q)-D(Q||P)\geq V}} \frac{1}{(n+1)^d} 2^{nH(Q)}. \end{aligned} \tag{5.23}$$

Since $\bigcup_n \mathcal{P}_n$ is dense in $\mathcal{P}$ and H is continuous we get $S^P(V) \geq \max_{\substack{Q \in \mathcal{P} \\ -H(Q)-D(Q||P) \geq V}} H(Q)$. We conclude that

$$S^P(V) = \max_{-H(Q)-D(Q||P) \geq V} H(Q). \tag{5.24}$$

Eq. (5.20) is shown in a similar manner. By using Lemma 1.2.5 in place of Lemma 1.2.3 in the above argument, one obtains Eq. (5.17) and Eq. (5.18) □

Once Proposition 5.1.5 has been established it becomes clearer how the Rényi entropy enters the picture. In order to prove Proposition 5.1.5 we need Lemma 5.1.4. It should be said that the proof of Proposition 5.1.5 has been extracted from [39].

Lemma 5.1.4. *Let $X \subset \mathbb{R}^n$ be a compact, convex set. Let $g : X \to \mathbb{R}$ be continuous and $h : X \to \mathbb{R}$ be continuous and strictly concave. Suppose h takes its maximum value at $x_2 \in X$. If g takes its minimum value at $x_1 \in X$, then*

$$y \mapsto \max_{x:g(x)=y} h(x) \quad y \in [g(x_1), g(x_2)] \tag{5.25}$$

is strictly monotonely increasing.
If g takes its maximum value at $x_1 \in X$, then

$$y \mapsto \max_{x:g(x)=y} h(x) \quad y \in [g(x_2), g(x_1)] \tag{5.26}$$

is strictly monotonely decreasing.

Proof. Assume g takes its minimum value at x_1. Let $g(x_1) \leq y' < y'' \leq g(x_2)$. Let $x' \in g^{-1}(y')$ such that $\max_{x:g(x)=y'} h(x) = h(x')$. By continuity of g we may find x'' on the line segment between x' and x_2, such that $g(x'') = y''$. That is

$$x'' = \lambda x_2 + (1-\lambda)x' \tag{5.27}$$

for some $\lambda \in (0, 1]$. Since h is strictly concave

$$h(x'') \geq \lambda h(x_2) + (1-\lambda)h(x') > h(x'). \tag{5.28}$$

So

$$\max_{x:g(x)=y'} h(x) = h(x') < h(x'') \leq \max_{x:g(x)=y''} h(x). \tag{5.29}$$

The second part of the lemma follows from the first by replacing g with $-g$. □

Given a probability distribution P with support $[d]$, we let

$$F_P(\alpha) = \log \sum_{i \in [d]} P(i)^\alpha. \tag{5.30}$$

Note that $F_P(\alpha)$ is just $(1-\alpha)H_\alpha(P)$. In order to make things simpler, we shall only consider F_P for probability distributions that are non-uniform (such that F_P is strictly convex) and ordered non-increasingly (such that we may simply write $P(1)$ instead of $\max_{x\in\mathcal{X}} P(x)$ and $P(d)$ instead of $\min_{x\in\mathcal{X}} P(x)$).

The function F_P will be central to the rest of the chapter. Note that

$$F_P'(\alpha) = \frac{\sum_i P(i)^\alpha \log P(i)}{\sum_i P(i)^\alpha} \tag{5.31}$$

is negative and monotone increasing $F_P' : \mathbb{R} \to \big(\log P(d), \log P(1)\big)$. We shall define

$$F_P'(\infty) = \lim_{\alpha\to\infty} F_P'(\alpha) = \log P(1) \tag{5.32}$$

and

$$F_P'(-\infty) = \lim_{\alpha\to-\infty} F_P'(\alpha) = \log P(d). \tag{5.33}$$

F_P is decreasing and strictly convex. Two important values to keep in mind are

$$\begin{aligned} F(0) &= H_0(P) = \log d, \\ F(1) &= 0. \end{aligned} \tag{5.34}$$

Also note the following bijections

$$F_P' \begin{cases} [-\infty, 0] & \longleftrightarrow \quad \left[\log P(d), \frac{\sum_i \log P(i)}{d}\right], \\ [0,1] & \longleftrightarrow \quad \left[\frac{\sum_i \log P(i)}{d}, -H(P)\right], \\ [1,\infty] & \longleftrightarrow \quad \left[-H(P), \log P(1)\right]. \end{cases} \tag{5.35}$$

We are now ready to give explicit formulas for the exponent functions Eq. (5.13), Eq. (5.14), Eq. (5.15), Eq. (5.16). The following proposition is in fact also true for uniform distribution, the statement just becomes trivial as $\big[\log P(d), \log P(1)\big]$ becomes a one-point set and the exponent functions become constant.

Proposition 5.1.5. *Let P be a non-uniform probability distribution with* $\mathrm{supp}(P) = [d]$ *which is ordered non-increasingly. For $V \in \big[\log P(d), \log P(1)\big]$ let $\alpha_V \in [-\infty, \infty]$ be the unique solution to $F_P'(\alpha) = V$, then*

$$M^P(V) = \begin{cases} 0 & \textit{if } V \in \big[\log P(d), -H(P)\big], \\ F_P(\alpha_V) + (1-\alpha_V)F_P'(\alpha_V) & \textit{if } V \in \big[-H(P), \log P(1)\big]. \end{cases} \tag{5.36}$$

$$M_*^P(V) = \begin{cases} F_P(\alpha_V) + (1-\alpha_V)F_P'(\alpha_V) & \textit{if } V \in \big[\log P(d), -H(P)\big], \\ 0 & \textit{if } V \in \big[-H(P), \log P(1)\big]. \end{cases} \tag{5.37}$$

$$S^P(V) = \begin{cases} \log d & \textit{if } V \in \left[\log P(d), \frac{\sum \log P(i)}{d}\right], \\ F_P(\alpha_V) - \alpha_V F_P'(\alpha_V) & \textit{if } V \in \left[\frac{\sum \log P(i)}{d}, \log P(1)\right]. \end{cases} \tag{5.38}$$

$$S_*^P(V) = \begin{cases} F_P(\alpha_V) - \alpha_V F_P'(\alpha_V) & \textit{if } V \in \left[\log P(d), \frac{\sum \log P(i)}{d}\right], \\ \log d & \textit{if } V \in \left[\frac{\sum \log P(i)}{d}, \log P(1)\right]. \end{cases} \tag{5.39}$$

Whenever $\alpha_V = \pm\infty$ the above formulas are to be interpreted as the limit $\alpha \to \pm\infty$.

Proof. Let $\mathcal{P}([d])$ be the set of probability distributions on $[d]$. The map $h : Q \mapsto H(Q)$ is concave on $\mathcal{P}([d])$ and takes its maximum value at the uniform distribution, where $-H(Q) - D(Q||P) = \frac{\sum \log P(i)}{d}$. The map $g : Q \to -H(Q) - D(Q||P)$ has maximim value $\log P(1)$ and minimum value $\log P(d)$. According to Lemma 5.1.4,

$$V \mapsto \max_{-H(Q)-D(Q||P)=V} H(Q) \tag{5.40}$$

is strictly monotone decreasing on $\left[\frac{\sum \log P(i)}{d}, \log P(1)\right]$ and strictly monotone increasing on $\left[\log P(d), \frac{\sum \log P(i)}{d}\right]$.

Similarly, since $D(Q||P)$ is convex with respect to Q, Lemma 5.1.4 tells us that

$$V \mapsto \max_{-H(Q)-D(Q||P)=V} -D(Q||P) \tag{5.41}$$

is strictly monotone increasing on $[\log P(d), -H(P)]$ and strictly monotone decreasing on $[-H(P), \log P(1)]$. For each $V \in [\log P(d), \log P(1)]$ we wish to find the probability distribution, Q, that solves the maximization problems in Eq. (5.40) and Eq. (5.41).

Given $V \in (\log P(d), \log P(1))$, let α be the solution to $F_P'(\alpha) = V$ and consider the distribution $P_\alpha(i) = \frac{P(i)^\alpha}{\sum_j P(j)^\alpha}$. Note that $-H(P_\alpha) - D(P_\alpha||P) = F_P'(\alpha) = V$. We prove that P_α solves the above optimization problems. Let $V \in (\log P(d), \log P(1))$ be given and choose $\alpha \in (-\infty, \infty)$ such that $-H(P_\alpha) - D(P_\alpha||P) = V$. Let $Q \in \mathcal{P}([d])$ be such that also

$-H(Q) - D(Q||P) = V$. We need to show that $H(P_\alpha) \geq H(Q)$.

$$
\begin{aligned}
&\frac{1}{1-\alpha}\Bigg[D(Q||P_\alpha) - H(P_\alpha) + H(Q)\Bigg] \\
&= \frac{1}{1-\alpha}\Bigg[D(Q||P_\alpha) + D(P_\alpha||P) - D(Q||P)\Bigg] \\
&= \frac{1}{1-\alpha}\Bigg[\sum_i -Q(i)\log\frac{P(i)^\alpha}{\sum_j P(j)^\alpha} + \frac{P(i)^\alpha}{\sum_j P(j)^\alpha}\log\frac{P(i)^\alpha}{\sum_j P(j)^\alpha} - \frac{P(i)^\alpha}{\sum_j P(j)^\alpha}\log P(i) + Q(i)\log P(i)\Bigg] \\
&= \frac{1}{1-\alpha}\Bigg[(1-\alpha)\sum_i \left(Q(i) - \frac{P(i)^\alpha}{\sum_j P(j)^\alpha}\right)\log P(i)\Bigg] \\
&= \sum_i \left(Q(i) - \frac{P(i)^\alpha}{\sum_j P(j)^\alpha}\right)\log P(i) = H(P_\alpha) + D(P_\alpha||P) - H(Q) - D(Q||P) = 0.
\end{aligned}
\tag{5.42}
$$

So

$$
H(P_\alpha) - H(Q) = D(Q||P_\alpha) \geq 0, \tag{5.43}
$$

which proves that P_α solves the optimization problems with the values

$$
\begin{aligned}
H(P_\alpha) &= -\sum_i \frac{P(i)^\alpha}{\sum_j P(j)^\alpha}\log\frac{P(i)^\alpha}{\sum_j P(j)^\alpha} = \log\sum P(i)^\alpha - \alpha\frac{\sum P(i)^\alpha \log P(i)}{\sum P(i)^\alpha} \\
&= F_P(\alpha) - \alpha F_P'(\alpha), \\
-D(P_\alpha||P) &= H(P_\alpha) - \Big(H(P_\alpha) + D(P_\alpha||P)\Big) = H(P_\alpha) + F_P'(\alpha) \\
&= F_P(\alpha) + (1-\alpha)F_P'(\alpha).
\end{aligned}
\tag{5.44}
$$

By Lemma 5.1.3 we obtain for $V \in \left[\frac{\sum \log P(i)}{d}, \log P(1)\right)$

$$
\begin{aligned}
S^P(V) &= \max_{-H(Q)-D(Q||P)\geq V} H(Q) = \max_{-H(Q)-D(Q||P)=V} H(Q) \\
&= H(P_\alpha) = F_P(\alpha) - \alpha F_P'(\alpha).
\end{aligned}
\tag{5.45}
$$

Similarly, for $V \in \left(\log P(d), \frac{\sum \log P(i)}{d}\right]$

$$
\begin{aligned}
S_*^P(V) &= \max_{-H(Q)-D(Q||P)\leq V} H(Q) = \max_{-H(Q)-D(Q||P)=V} H(Q) \\
&= H(P_\alpha) = F_P(\alpha) - \alpha F_P'(\alpha).
\end{aligned}
\tag{5.46}
$$

For $V \in \big(\log P(d), -H(P)\big]$

$$
\begin{aligned}
M_*^P(V) &= \max_{-H(Q)-D(Q||P)\leq V} -D(Q||P) = \max_{-H(Q)-D(Q||P)=V} -D(Q||P) \\
&= -D(P_\alpha||P) = F_P(\alpha) + (1-\alpha)F_P'(\alpha).
\end{aligned}
\tag{5.47}
$$

And for $V \in \left[-H(P), \log P(1)\right)$

$$\begin{aligned} M^P(V) &= \max_{-H(Q)-D(Q||P)\geq V} -D(Q||P) = \max_{-H(Q)-D(Q||P)=V} -D(Q||P) \\ &= -D(P_\alpha||P) = F_P(\alpha) + (1-\alpha)F_P'(\alpha). \end{aligned} \tag{5.48}$$

We may take α to $-\infty$ or ∞ and get the results at the boundary. $\square$

Remark 5.1.6. *Define $\overline{m_n^P}(V)$ on $\left[\log P(d), \log P(1)\right]$ to be equal to $m_n^P(V)$ at the endpoints, but for $V \in \left(\log P(d), \log P(1)\right)$ we use a strict inequality and define*

$$\overline{m_n^P}(V) = \sum_{\substack{I\in[d]^n \\ P^{\otimes n}(I)>2^{nV}}} P^{\otimes n}(I). \tag{5.49}$$

Define $\overline{m_{n*}^P}$, $\overline{s_n^P}$ and $\overline{s_{n*}^P}$ similarly. By continuity of M^P, M_*^P, S^P and S_*^P one sees that we could replace $m_n^P, m_{n*}^P, s_n^P, s_{n*}^P$ in equations Eq. (5.13), Eq. (5.14), Eq. (5.15), Eq. (5.16) with respectively $\overline{m_n^P}, \overline{m_{n*}^P}, \overline{s_n^P}, \overline{s_{n*}^P}$, without the limit changing. Furtermore since all functions are monotone and the limit functions are monotone, continuous and bounded, the convergences are all uniform. This will be important later.

A few nice values to keep in mind for S^P, M^P and M_*^P are the following

$$\begin{aligned} &M^P(-H(P)) = 0 \\ &M^P(\log P(1)) = \log P(1) + \log\left|\{i\in[d] \mid P(i) = P(1)\}\right| \\ &M_*^P(-H(P)) = 0 \\ &S^P(\log P(1)) = \log\left|\{i\in[d] \mid P(i) = P(1)\}\right| \\ &S^P(-H(P)) = H(P) \\ &S^P\left(\sum \frac{\log P(i)}{d}\right) = \log d = H_0(P). \end{aligned} \tag{5.50}$$

5.2. A sufficient and almost necessary condition for asymptotic majorization

Lemma 5.2.1. *Let P and Q be non-uniform probability distributions with*

$$\min_{\alpha\in[0,1]} \frac{H_\alpha(P)}{H_\alpha(Q)} > 1. \tag{5.51}$$

For sufficiently small $\varepsilon > 0$ and all $V \in \left[\frac{\sum \log P(i)}{d_P}, -H(P)\right]$ and $W \in \left[\frac{\sum \log Q(i)}{d_Q}, -H(Q)\right]$

$$S^P(V) \leq S^Q(W) + \varepsilon \implies M_*^P(V) \geq M_*^Q(W) + \varepsilon. \tag{5.52}$$

Proof. We wish to apply Lemma A.1.1 to $F = F_P : [0,1] \to \mathbb{R}$ and $G = F_Q : [0,1] \to \mathbb{R}$. The conditions of the lemma are satisfied, since for $\alpha \in [0,1)$, $H_\alpha(P) > H_\alpha(Q)$ is equivalent to

$F_P(\alpha) > F_Q(\alpha)$ and $F_P'(1) = -H(P) < -H(Q) = F_Q'(1)$. By Lemma A.1.1, there is an $\varepsilon > 0$ such that

$$F_P(x) - xF_P'(x) \leq F_Q(y) - yF_Q'(y) + \varepsilon \implies F_P(x) + (1-x)F_P'(x) \geq F_Q(y) + (1-y)F_Q'(y) + \varepsilon. \tag{5.53}$$

By Eq. (5.37) and Eq. (5.38) applied to P and Q, this is precisely the same as

$$S^P(V) \leq S^Q(W) + \varepsilon \implies M_*^P(V) \geq M_*^Q(W) + \varepsilon. \tag{5.54}$$

□

Lemma 5.2.2. *Let P and Q be non-uniform probability distributions with*

$$\min_{\alpha\in[1,\infty]} \frac{H_\alpha(P)}{H_\alpha(Q)} > 1. \tag{5.55}$$

For sufficiently small $\varepsilon > 0$ and all $V \in \big[-H(P), \log P(1)\big]$ and $W \in \big[-H(Q), \log Q(1)\big]$

$$S^P(V) \leq S^Q(W) + \varepsilon \implies M^P(V) + \varepsilon \leq M^Q(W). \tag{5.56}$$

Proof. The proof of this lemma is essentially the same as the previous proof, using Lemma A.1.2 in place of Lemma A.1.1, again with $F = F_P$ and $G = F_Q$ and by using Eq. (5.36) in place of Eq. (5.37). The lemma applies since for $\alpha > 1$, $H_\alpha(P) > H_\alpha(Q)$ is equivalent to $F_P(\alpha) < F_Q(\alpha)$, because $1-\alpha$ is negative and $\lim_{\alpha\to\infty} F_P'(\alpha) = -H_\infty(P) < -H_\infty(Q) = \lim_{\alpha\to\infty} F_Q'(\alpha)$. Furthermore the limits $\lim_{\alpha\to\infty} F_P(\alpha) - \alpha F_P'(\alpha)$ and $\lim_{\alpha\to\infty} F_Q(\alpha) - \alpha F_Q'(\alpha)$ exist according to Proposition 5.1.5 and are equal to respectively $S^P(\log P(1))$ and $S^Q(\log Q(1))$. □

Proposition 5.2.3. *Let $P = P^\downarrow : [d_P] \to [0,1]$ and $Q = Q^\downarrow : [d_Q] \to [0,1]$ be non-uniform probability distributions with $P(1) > P(2)$ such that*

$$\min_{\alpha\in[0,1]} \frac{H_\alpha(P)}{H_\alpha(Q)} > 1. \tag{5.57}$$

Let $V^* \in \big[\log P(d_P), \log P(1)\big]$ be such that $S^P(V^*) \in \big(H(Q), H(P)\big)$. Then for all sufficiently large n, and all N such that $V = \frac{1}{n}\log(P^{\otimes n\downarrow}(N)) \in \big[\log P(d_P), V^*\big]$,

$$\sum_{i=1}^{N-1} P^{\otimes n\downarrow}(i) \leq \sum_{i=1}^{N-1} Q^{\otimes n\downarrow}(i). \tag{5.58}$$

Proof. Let $\varepsilon > 0$ be small enough that Lemma 5.2.1 applies. By the assumption that $P(1) > P(2)$, S^P surjects $\big[\log P(d_P), \log P(1)\big]$ decreasingly onto $\big[0, H_0(P)\big]$. Assuming $\varepsilon < H_0(P) - H_0(Q)$, we have $H_0(Q) + \varepsilon < H_0(P)$ and we therefore have $V_1 \in \big[\log P(d_P), \log P(1)\big]$ such that $S^P(V_1) = H_0(Q) + \varepsilon$. Now let n be large enough that for both P and Q and all V and W

Eq. (5.13), Eq. (5.14), Eq. (5.15) and Eq. (5.16) are good approximations, and also good approximations when replaced by the alternative versions in Remark 5.1.6 (this may be done since the convergences are uniform). Note that $P^{\otimes n\downarrow}(N) = 2^{nV}$ such that $N \geq \overline{s_n^P}(V)$. We split into three cases: Either $V \in \big[\log P(d_P), V_1\big]$, or $V_1 < -H(P)$ and $V \in \big[V_1, -H(P)\big]$, or $V \in [-H(P), V^*]$.

First assume that $V \in \big[\log P(d_P), V_1\big]$. Then

$$\frac{1}{n}\log N \geq \frac{1}{n}\log \overline{s_n^P}(V) \geq \frac{1}{n}\log \overline{s_n^P}(V_1) \simeq S^P(V_1) = H_0(Q) + \varepsilon > H_0(Q), \tag{5.59}$$

which implies $N > 2^{nH_0(Q)}$, so

$$\sum_{i=1}^{N-1} Q^{\otimes n\downarrow}(i) = 1 \tag{5.60}$$

and Eq. (5.58) holds trivially.

Assume instead that $V \in \big[V_1, -H(P)\big]$, provided that the interval is non-empty. Without loss of generality we can assume that $\varepsilon < H(P) - H(Q)$. Since S^Q maps $\left[\frac{\sum \log Q(i)}{d_Q}, -H(Q)\right]$ onto $\big[H(Q), H_0(Q)\big]$ and $H(Q) + \varepsilon < H(P) = S^P(-H(P))$ and $H^0(Q) + \varepsilon = S^P(V_1)$, we may find $W \in \left[\frac{\sum \log Q(i)}{d_Q}, -H(Q)\right]$ such that $S^Q(W) + \varepsilon = S^P(V)$. By Lemma 5.2.1, this implies

$$M_*^P(V) \geq M_*^Q(W) + \varepsilon. \tag{5.61}$$

And

$$\frac{1}{n}\log N \geq \frac{1}{n}\log \overline{s_n^P}(V) \simeq S^P(V) = S^Q(W) + \varepsilon > S^Q(W) \simeq \frac{1}{n}\log s_n^Q(W), \tag{5.62}$$

which implies $N > s_n^Q(W)$. Therefore

$$\begin{aligned}
&\frac{1}{n}\log \sum_{i=N}^{\infty} P^{\otimes n\downarrow}(i) \geq \frac{1}{n}\log \overline{m_{n*}^P}(V) \simeq M_*^P(V) \geq M_*^Q(W) + \varepsilon \\
&> M_*^Q(W) \simeq \frac{1}{n}\log m_{n*}^Q(W) = \frac{1}{n}\log \sum_{\substack{I \in [d_Q]^n \\ Q^{\otimes n}(I) \leq 2^{nW}}} Q^{\otimes n}(I) \\
&\geq \frac{1}{n}\log \sum_{i=s_n^Q(W)}^{\infty} Q^{\otimes n\downarrow}(i) \geq \frac{1}{n}\log \sum_{i=N}^{\infty} Q^{\otimes n\downarrow}(i),
\end{aligned} \tag{5.63}$$

which implies Eq. (5.58).

Finally, assume that $V \in [-H(P), V^*]$. Let W be such that $S^Q(W) \in \Big(H(Q), S^P(V^*)\Big)$. Note, importantly, that W is chosen independently from V, so that n does not depend on V. Then

$$\frac{1}{n}\log N \geq \frac{1}{n}\log \overline{s_n^P(V)} \simeq S^P(V) \geq S^P(V^*) > S^Q(W) \simeq \frac{1}{n}\log s_n^Q(W), \tag{5.64}$$

So $N > s_n^Q(W)$. Since S^Q is strictly decreasing $W < -H(Q)$, and since $V \geq -H(P)$ it follows from Eq. (5.38) that $M_*^P(V) = 0 > M_*^Q(W)$. We conclude as in Eq. (5.63) that

$$\frac{1}{n} \log \sum_{i=N}^{\infty} P^{\otimes n \downarrow}(i) \geq \frac{1}{n} \log \sum_{i=N}^{\infty} Q^{\otimes n \downarrow}(i). \tag{5.65}$$

$\square$

Proposition 5.2.4. *Let $P = P^{\downarrow} : [d_P] \to [0,1]$ and $Q = Q^{\downarrow} : [d_Q] \to [0,1]$ be non-uniform probability distributions with*

$$\min_{\alpha \in [1,\infty]} \frac{H_\alpha(P)}{H_\alpha(Q)} > 1. \tag{5.66}$$

Let $V^* \in \big[\log P(d_P), \log P(1)\big]$ be such that $S^P(V^*) \in \big(H(Q), H(P)\big)$. Then for all sufficiently large n, and all N such that $V = \frac{1}{n}\log(P^{\otimes n \downarrow}(N)) \in \big[V^*, \log P(1)\big]$,

$$\sum_{i=1}^{N} P^{\otimes n \downarrow}(i) \leq \sum_{i=1}^{N} Q^{\otimes n \downarrow}(i). \tag{5.67}$$

Proof. Like in the proof of Proposition 5.2.3 we let $\varepsilon > 0$ be small enough that Lemma 5.2.2 applies. We split into three cases. Letting $\varepsilon > 0$ be sufficiently small we may let $W_1 \in \big(\log P(1), \log Q(1)\big)$ be the solution to $S^Q(W_1) = S^Q(\log Q(1)) + \varepsilon$.

Firstly, we assume that $S^P(V) \leq S^Q(W_1)$, then

$$\frac{1}{n} \log N \leq \frac{1}{n} \log s_n^P(V) \simeq S^P(V) \leq S^Q(W_1) < S^Q(\log P(1)) \simeq \frac{1}{n} \log s_n^Q(\log P(1)), \tag{5.68}$$

showing that $N \leq s_n^Q(\log P(1))$, which implies that $Q^{\otimes n \downarrow}(i) \geq \log P(1)$ for all $i \in [N]$. So

$$\sum_{i=1}^{N} Q^{\otimes n \downarrow}(i) \geq N \log P(1) \geq \sum_{i=1}^{N} P^{\otimes n \downarrow}(i). \tag{5.69}$$

Secondly, we assume that $S^P(V) \in \Big[S^Q(W_1), H(Q)\Big]$, provided that the interval is non-empty. Let $W \in \big[-H(Q), \log Q(1)\big]$ be such that $S^Q(W) + \varepsilon = S^P(V)$, which is possible by the choice of W_1. By Lemma 5.2.2

$$M^P(V) + \varepsilon \leq M^Q(W). \tag{5.70}$$

And

$$\frac{1}{n} \log N \geq \frac{1}{n} \log \overline{s_n^P}(V) \simeq S^P(V) = S^Q(W) + \varepsilon/2 > S^Q(W) \simeq \frac{1}{n} \log s_n^Q(W), \tag{5.71}$$

showing that $N > s_n^Q(W)$.

$$\begin{aligned}
&\frac{1}{n}\log\sum_{i=1}^{N} P^{\otimes n\downarrow}(i) \leq \frac{1}{n}\log m_n^P(V) \simeq M^P(V) \\
&< M^P(V)+\varepsilon \leq M^Q(W) \simeq \frac{1}{n}\log m_n^Q(W) \\
&= \frac{1}{n}\log\sum_{i=1}^{s_n^Q(W)} Q^{\otimes n\downarrow}(i) \leq \frac{1}{n}\log\sum_{i=1}^{N} Q^{\otimes n\downarrow}(i).
\end{aligned} \tag{5.72}$$

Finally, assume that $S^P(V) \in \Big[H(Q), S^P(V^*)\Big]$. Let $W > -H(Q)$ be such that $M^Q(W) > M^P(V^*)$. Then

$$\frac{1}{n}\log N \geq \frac{1}{n}\log\overline{s_n^P}(V) \simeq S^P(V) \geq H(Q) > S^Q(W) \simeq \frac{1}{n}\log s_n^Q(W), \tag{5.73}$$

showing that $N > s_n^Q(W)$.

$$\begin{aligned}
&\frac{1}{n}\log\sum_{i=1}^{N} P^{\otimes n\downarrow}(i) \leq \frac{1}{n}\log m_n^P(V) \simeq M^P(V) \leq M^P(V^*) < M^Q(W) \\
&\simeq \frac{1}{n}\log m_n^Q(W) = \frac{1}{n}\log\sum_{\substack{I\in[d_Q]^n \\ Q^{\otimes n}(I)\geq 2^{nW}}} Q^{\otimes n}(I) \\
&= \frac{1}{n}\log\sum_{i=1}^{s_n^Q(W)} Q^{\otimes n\downarrow}(i) \leq \frac{1}{n}\log\sum_{i=1}^{N} Q^{\otimes n\downarrow}(i).
\end{aligned} \tag{5.74}$$

,

□

So far we have assumed that all probability distributions are non-uniform. This was mainly a matter of convenience. In the following we no longer make this assumption. If Q is the trivial probability distribution (i.e. $|\mathrm{supp}(Q)| = 1$), then $P^{\otimes n} \preceq Q^{\otimes n}$ holds for any P and n, so this case is rather uninteresting.

Proposition 5.2.5. *Let $P = P^{\downarrow} : [d_P] \to \mathbb{R}$ and $Q = Q^{\downarrow} : [d_Q] \to \mathbb{R}$ be two probability distributions with $d_Q > 1$ and assume that*

$$\min_{\alpha\in[0,\infty]} \frac{H_\alpha(P)}{H_\alpha(Q)} > 1. \tag{5.75}$$

For sufficiently large n

$$P^{\otimes n} \preceq Q^{\otimes n} \tag{5.76}$$

Proof. If $d_P = 1$ then $H_\alpha(P) = 0$ for all α, so we may assume that $d_P > 1$. For small $\delta > 0$, let

$$P_\delta(i) = \begin{cases} P(1)+\delta & \text{if } i = 1, \\ P(i) & \text{if } 1 < i < d_P, \\ P(d_P)-\delta & \text{if } i = d_P. \end{cases} \tag{5.77}$$

$$Q_\delta(i) = \begin{cases} Q(1) - \delta & \text{if } i = 1, \\ Q(i) & \text{if } 1 < i < d_Q, \\ Q(d_Q) + \delta & \text{if } i = d_Q. \end{cases} \tag{5.78}$$

When δ is sufficiently small

$$\min_{\alpha \in [0,\infty]} \frac{H_\alpha(P_\delta)}{H_\alpha(Q_\delta)} > 1. \tag{5.79}$$

By applying Propositions 5.2.3 and 5.2.4 to P_δ and Q_δ, we get for large n

$$P^{\otimes n} \preceq P_\delta^{\otimes n} \preceq Q_\delta^{\otimes n} \preceq Q^{\otimes n}. \tag{5.80}$$

□

We have now established a sufficient condition for asymptotic majorization. In fact this condition is almost necessary, since for $\alpha \in (0,\infty)$ the α-Rényi entropy is strictly Schur-concave:

Proposition 5.2.6. *Let P and Q be two probability distribution with $P \preceq Q$. Then either*

$$P^\downarrow = Q^\downarrow \tag{5.81}$$

or

$$H_\alpha(P) > H_\alpha(Q) \quad \textit{for all } \alpha \in (0,\infty). \tag{5.82}$$

I.e. H_α is strictly Schur-convcave for $\alpha \in (0,\infty)$; see [51, 3.A.1].

For a proof of Proposition 5.2.6 see e.g. [51, 3.C.1.a], which applied to the map $p \mapsto -p \log p$ shows that H_1 is strictly Schur-concave. Applying [51, 3.C.1.a] to the map $p \mapsto p^\alpha$, shows that $P \mapsto \sum_i P(i)^\alpha$ is strictly Schur-concave for $\alpha \in (0,1)$ and strictly Schur-convex for $\alpha \in (1,\infty)$. So H_α is strictly Schur-concave for all $\alpha \in (0,\infty)$.

Using the fact that $H_\alpha(P^{\otimes n}) = nH_\alpha(P)$, we may sum up the contents of Propositions 5.2.5 and 5.2.6 as follows: When $P^\downarrow \neq Q^\downarrow$;

$$\begin{gathered} \forall \alpha \in [0,\infty] : H_\alpha(P) > H_\alpha(Q) \\ \Downarrow \textit{Proposition } 5.2.5 \\ \exists n \in \mathbb{N} : P^{\otimes n} \preceq Q^{\otimes n} \\ \Downarrow \textit{Proposition } 5.2.6 \\ \forall \alpha \in (0,\infty) : H_\alpha(P) > H_\alpha(Q). \end{gathered} \tag{5.83}$$

Remark 5.2.7. *It is natural to ask if we can make requirements at* 0 *and* ∞ *in order to get a biimplication, that is, if we can determine* $\exists n \in \mathbb{N} : P^{\otimes n} \preceq Q^{\otimes n}$ *entirely from comparing Rényi entropies. It seems that in order to do so, we would have to be more careful with our estimations. I cautiously conjectures that requiring a weak inequality at* ∞ *is sufficient, and that the requirement of a sharp inequality at* 0 *could be replaced by a similar condition regarding the* α*-Rényi entropies for negative* α.

Definition 5.2.8. *When* P *and* Q *are probability distributions with finite support, we let*

$$E(P,Q) = \sup\left\{ r \in \mathbb{R}_{\geq 0} \middle| \text{ for large } n \ P^{\otimes n} \preceq Q^{\otimes \lfloor nr \rfloor} \right\}. \tag{5.84}$$

When $|\mathrm{supp}\ Q| = 1$, $E(P,Q) = \infty$.

Theorem 5.2.9. *Given finitely supported probability distributions* P *and* Q*, with* $|\mathrm{supp}(Q)| > 1$,

$$E(P,Q) = \min_{\alpha \in [0,\infty]} \frac{H_\alpha(P)}{H_\alpha(Q)}. \tag{5.85}$$

Proof. Let $r < \min_{\alpha \in [0,\infty]} \frac{H_\alpha(P)}{H_\alpha(Q)}$. Then for large n

$$\min_{\alpha \in [0,\infty]} \frac{H_\alpha(P^{\otimes n})}{H_\alpha(Q^{\otimes \lfloor nr \rfloor})} = \min_{\alpha \in [0,\infty]} \frac{n}{\lfloor nr \rfloor} \frac{H_\alpha(P)}{H_\alpha(Q)} > 1. \tag{5.86}$$

By Proposition 5.2.5, $P^{\otimes n} \preceq Q^{\otimes \lfloor nr \rfloor}$.

Let $r > \min_{\alpha \in [0,\infty]} \frac{H_\alpha(P)}{H_\alpha(Q)}$ and choose some α_r such that $r > \frac{H_{\alpha_r}(P)}{H_{\alpha_r}(Q)}$. Then for large n

$$\frac{H_{\alpha_r}(P^{\otimes n})}{H_{\alpha_r}(Q^{\otimes \lfloor nr \rfloor})} = \frac{n}{\lfloor nr \rfloor} \frac{H_{\alpha_r}(P)}{H_{\alpha_r}(Q)} < 1. \tag{5.87}$$

By Proposition 5.2.6 $P^{\otimes n} \not\preceq Q^{\otimes \lfloor nr \rfloor}$. $\square$

5.3. Success probability going to 1

In Chapter 4 we considered optimal extraction rates where the success probability was allowed to go to 0. Setting $r = 0$ in equation Eq. (4.53) gives the optimal extraction rate between the two states, where the success rate is allowed to go to 0, but not exponentially fast. This is a good candidate for the optimal extraction rate, when we demand that the success probability goes to 1, the rate which we called $E_{\mathcal{P}}$ in Eq. (5.2).

Indeed, as is shown in Theorem 5.3.3 below, setting $r = 0$ in Eq. (4.53) gives a formula for $E_{\mathcal{P}}(\psi, \phi)$.

Proposition 5.3.1. *Let $P = P^{\downarrow} : [d_P] \to \mathbb{R}$ and $Q = Q^{\downarrow} : [d_Q] \to \mathbb{R}$ be probability distributions, such that*

$$\min_{\alpha\in[0,1]} \frac{H_\alpha(P)}{H_\alpha(Q)} > 1. \tag{5.88}$$

Then for sufficiently large n

$$\sqrt{x_n}\,|\psi_P\rangle^{\odot n} \xrightarrow{\text{LOCC}} |\psi_Q\rangle^{\odot n} \tag{5.89}$$

for some sequence of $x_n \geq 1$ with $x_n \to 1$. That is, one can asymptotically transform n copies of $|\psi\rangle$ to n copies of $|\phi\rangle$ with probability of success going to 1 as $n \to \infty$.

Proof. Analogue to the proof of Proposition 5.2.5, if we assume that we have proven the statement of Proposition 5.3.1 for all non-uniform P and Q with $P(1) > P(2)$. Then for general P and Q satisfying Eq. (5.88), we let P_δ and Q_δ be non-uniform probability distributions with $P \preceq P_\delta$ and $Q_\delta \preceq Q$, such that $\min_{\alpha\in[0,1]} \frac{H_\alpha(P)}{H_\alpha(Q)} > \min_{\alpha\in[0,1]} \frac{H_\alpha(P_\delta)}{H_\alpha(Q_\delta)} > 1$. Then by Theorem 5.0.1 and Proposition 5.3.1 for P_δ and Q_δ, the statement follows for P and Q. So without loss of generality, we can assume that P and Q are satisfy the conditions of Proposition 5.2.3.

$$|\psi_P\rangle^{\odot n} = |\psi_{P^{\otimes n}}\rangle = \sum_{I\in[d_P]^n} \sqrt{P^{\otimes n}(I)}\,|II\rangle\,. \tag{5.90}$$

From Eq. (5.88) we conclude that $H(P) > H(Q)$. Let $V^* > -H(P)$ be chosen such that Proposition 5.2.3 applies. Set $t_n = 2^{nV^*}$ and note that

$$|\psi_P\rangle^{\odot n} \xrightarrow{\text{LOCC}} \sum_{I\in[d_P]^n} \min(\sqrt{P^{\otimes n}(I)}, \sqrt{t_n})\,|II\rangle\,. \tag{5.91}$$

Let $x_n = \left(\sum_{I\in[d_P]^n} \min(P^{\otimes n}(I), t_n)\right)^{-1}$ such that

$$|\eta_n\rangle = \sqrt{x_n} \sum_{I\in[d_P]^n} \min(\sqrt{P^{\otimes n}(I)}, \sqrt{t_n})\,|II\rangle \tag{5.92}$$

is normalized and

$$\sqrt{x_n}\,|\psi_P\rangle^{\odot n} \xrightarrow{\text{LOCC}} |\eta_n\rangle\,. \tag{5.93}$$

The proof is complete, when it is shown that $x_n \to 1$ and $|\eta_n\rangle \xrightarrow{\text{LOCC}} |\psi_Q\rangle^{\odot n}$ for large n.

To see that $x_n \to 1$, first note that

$$\sum_{I\in[d_P]^n} \min(P^{\otimes n}(I), t_n) \geq 1 - \sum_{\substack{I\in[d_P]^n \\ P^{\otimes n}(I)\geq t_n}} P^{\otimes n}(I). \tag{5.94}$$

By Eq. (5.36), since P is non-uniform and $V^* > -H(P)$,

$$\lim_{n\to\infty} \frac{1}{n} \log \sum_{\substack{I\in[d_P]^n \\ P^{\otimes n}(I)\geq t_n}} P^{\otimes n}(I) = M_P(V^*) < 0, \tag{5.95}$$

which implies

$$\lim_{n\to\infty} \sum_{\substack{I\in[d_P]^n \\ P^{\otimes n}(I)\geq t_n}} P^{\otimes n}(I) = 0. \tag{5.96}$$

So by Eq. (5.94), $x_n \to 1$.

To see that $|\eta_n\rangle \xrightarrow{\text{LOCC}} |\psi_Q\rangle^{\odot n}$ for large n, let $P_n \in \mathcal{P}([d_P]^n)$ be the probability distribution given by $P_n(I) = x_n \min(P^{\otimes n}(I), t_n)$, such that $|\eta_n\rangle = |\psi_{P_n}\rangle$. By Theorem 5.0.1, it must be shown that $P_n \preceq Q^{\otimes n}$ for large n. When N is large enough that $P^{\otimes n\downarrow}(N) \leq 2^{nV^*}$, then for all $i \geq N$, $P_n^{\downarrow}(i) = x_n P^{\otimes n\downarrow}(i) \geq P^{\otimes n\downarrow}(i)$, so

$$\sum_{i=1}^{N-1} P_n^{\downarrow}(i) \leq \sum_{i=1}^{N-1} P^{\otimes n\downarrow}(i). \tag{5.97}$$

And by Proposition 5.2.3, we have for large n

$$\sum_{i=1}^{N-1} P_n^{\downarrow}(i) \overset{Eq.\ (5.97)}{\leq} \sum_{i=1}^{N-1} P^{\otimes n\downarrow}(i) \overset{Proposition\ 5.2.3}{\leq} \sum_{i=1}^{N-1} Q^{\otimes n\downarrow}(i). \tag{5.98}$$

What remains is to deal with all N such that $P^{\otimes n\downarrow}(N) > 2^{nV^*}$. To this end, let N^* be the largest number such that $P^{\otimes n\downarrow}(N^*) > 2^{nV^*}$. By Eq. (5.98)

$$\sum_{i=1}^{N^*} P_n^{\downarrow}(i) \leq \sum_{i=1}^{N^*} Q^{\otimes n\downarrow}(i), \tag{5.99}$$

and since $P_n^{\downarrow}(i) = x_n t_n$ is constant for $i \in [N^*]$, we have

$$\sum_{i=1}^{N} P_n^{\downarrow}(i) \leq \sum_{i=1}^{N} Q^{\otimes n\downarrow}(i) \tag{5.100}$$

for all $N \leq N^*$. □

Using $H_\alpha(P^{\otimes n}) = nH_\alpha(P)$, we obtain the following:

Corollary 5.3.2. *Given $n, m \in \mathbb{N}$ with*

$$\frac{m}{n} > \min_{\alpha\in[0,1]} \frac{H_\alpha(P)}{H_\alpha(Q)}. \tag{5.101}$$

Then

$$\sqrt{x_k}\,|\psi\rangle^{\otimes nk} \xrightarrow{\text{LOCC}} |\phi\rangle^{\otimes mk} \tag{5.102}$$

for large k and some sequence $x_k \to 1$.

Corollary 5.3.2 and Eq. (4.53) with $r = 0$, yield respectively a lower and upper bound on $E_{\mathcal{P}}(\psi, \phi)$, which can be summed up as:

Theorem 5.3.3. *Given probability distributions P and Q*

$$E_{\mathcal{P}}\left(\psi_P, \psi_Q\right) = \min_{\alpha\in[0,1]} \frac{H_\alpha(P)}{H_\alpha(Q)}. \tag{5.103}$$

5.4. Minimizing the ratio of Rényi entropies

In Theorem 5.2.9 and Theorem 5.3.3, we are minimizing the same ratio $\frac{H_\alpha(P)}{H_\alpha(Q)}$, but over different intervals. So $E_{\mathcal{P}}(\psi_P, \psi_Q) = E(\psi_P, \psi_Q)$ if and only if the minimum is attained at $\alpha \in [0,1]$ and one might ask when this happens. Let us first study what happens in the simplest case of binary probability distributions; $\mathrm{supp}P^{\downarrow} = \mathrm{supp}Q^{\downarrow} = \{1,2\}$, such that the corresponding quantum states are qubits. If $P^{\downarrow}(1) \leq Q^{\downarrow}(1)$, then $P \preceq Q$, which implies that $E(P,Q) \geq 1$. But $\frac{H_0(P)}{H_0(Q)} = \frac{\log(2)}{\log(2)} = 1$, so the minimum is attained at $\alpha = 0$. To see what happens when $P^{\downarrow}(1) > Q^{\downarrow}(1)$ we need the following lemma.

Lemma 5.4.1.

$$f(x) = \frac{\log\left(x^\alpha + (1-x)^\alpha\right)}{\log x} \tag{5.104}$$

is an increasing function for $x \in (0,1)$, when $\alpha > 1$ and a decreasing function for $x \in (0,1)$, when $\alpha \in (0,1)$

Proof. Let us assume that $\alpha > 1$.

$$\begin{aligned} f(x) = \frac{\log\left(x^\alpha + (1-x)^\alpha\right)}{\log x} &= \frac{\log\left(x^\alpha\right) + \log\left(1 + (\frac{1}{x}-1)^\alpha\right)}{\log x} \\ = \alpha + \frac{\log\left(1 + (\frac{1}{x}-1)^\alpha\right)}{\log x} &= \alpha - \frac{\log\left(1 + (\frac{1}{x}-1)^\alpha\right)}{\log \frac{1}{x}}, \end{aligned} \tag{5.105}$$

by substituting $y = \frac{1}{x} - 1$, it suffices to show that $g : y \mapsto \frac{\log(1+y^\alpha)}{\log(1+y)}$ is increasing for $y > 0$. Taking the derivative of g gives

$$g'(y) = \frac{\alpha y^\alpha (1+y)\log(1+y) - y(1+y^\alpha)\log(1+y^\alpha)}{y(1+y)(1+y^\alpha)\log(1+y)^2}, \tag{5.106}$$

which is positive if and only if the numerator is positive. This is equivalent to

$$\alpha y^\alpha (1+y)\log(1+y) > y(1+y^\alpha)\log(1+y^\alpha) \tag{5.107}$$

which is equivalent to

$$\frac{(1+y)\log(1+y)}{y\log(y)} > \frac{(1+y^\alpha)\log(1+y^\alpha)}{y^\alpha\log(y^\alpha)}. \tag{5.108}$$

Since the maps $z \mapsto \frac{1+z}{z}$ and $z \mapsto \frac{\log(1+z)}{\log(z)}$ are decreasing for $z \in (0,1)$ and $z \in (1,\infty)$ and since $y^\alpha > y$ and y^α and y belong to the same interval, we conclude that Eq. (5.108) holds and therefore f is increasing. When $\alpha \in (0,1)$ we have $y^\alpha < y$ for $y > 0$, so the inequalities flip and f is therefore decreasing. □

Proposition 5.4.2. *When P and Q are probability distributions with $|\mathrm{supp}(P)| = |\mathrm{supp}(Q)| = 2$, the minimum $\min_{\alpha\in[0,\infty]} \frac{H_\alpha(P)}{H_\alpha(Q)}$ is attained at either $\alpha = 0$ or $\alpha = \infty$.*

Proof. As we already saw, when $P^{\downarrow}(1) \leq Q^{\downarrow}(1)$, the minimum is attained at $\alpha = 0$. Assume instead that $p = P^{\downarrow}(1) > Q^{\downarrow}(1) = q$. For any $\alpha \notin \{0, 1, \infty\}$,

$$\frac{H_\alpha(P)}{H_\infty(P)} = \frac{\frac{1}{1-\alpha} \log\left(p^\alpha + (1-p)^\alpha\right)}{-\log p}, \tag{5.109}$$

which by Lemma 5.4.1 is larger than

$$\frac{\frac{1}{1-\alpha} \log\left(q^\alpha + (1-q)^\alpha\right)}{-\log q} = \frac{H_\alpha(Q)}{H_\infty(Q)}. \tag{5.110}$$

So

$$\frac{H_\alpha(P)}{H_\alpha(Q)} > \frac{H_\infty(P)}{H_\infty(Q)}. \tag{5.111}$$

By continuity in α, there is at least weak inequality for $\alpha \in \{0, 1\}$. $\square$

In order to get $\alpha \in (0, \infty)$ into play, we thus need to consider probability distributions with support larger than 2. When the support is larger than 2, we sometimes get minimizations on the interior of the interval, and the minimizing α can be both larger and smaller than 1. For instance when $(P(1), P(2), P(3)) = \left(\frac{5}{8}, \frac{1}{3}, \frac{1}{24}\right)$ and $(Q(1), Q(2), Q(3)) = \left(\frac{2}{3}, \frac{1}{5}, \frac{2}{15}\right)$, the minimizing α is around $\alpha \approx 0.802$, with a rate $\frac{H_\alpha(P)}{H_\alpha(Q)} \approx 0.9168$. Or when $(P(1), P(2), P(3)) = \left(\frac{1}{2}, \frac{5}{12}, \frac{1}{12}\right)$ and $(Q(1), Q(2), Q(3)) = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$, the minimizing α is around $\alpha \approx 1.96$, with a rate $\frac{H_\alpha(P)}{H_\alpha(Q)} \approx 0.8591$.

Appendix A

A.1. Inequalities

Lemma A.1.1. *Let $F, G : [0,1] \to \mathbb{R}$ be decreasing, strictly convex and continuously differentiable functions with $F(x) > G(x)$ for all $x \in [0,1)$, $F(1) = G(1) = 0$ and $F'(1) < G'(1)$. Then there exists an $\varepsilon > 0$ such that for all $x, y \in [0,1]$*

$$F(x) - xF'(x) \leq G(y) - yG'(y) + \varepsilon \implies F(x) + (1-x)F'(x) \geq G(y) + (1-y)G'(y) + \varepsilon \quad \text{(A.1)}$$

Proof. We start by proving (A.1) without the ε and with a sharp inequality on the right-hand-side, since then (A.1) follows by compactness.

Let $x, y \in [0,1]$ and let $g : [0,1] \to \mathbb{R}$ be the affine function

$$g(t) = G(y) + (t-y)G'(y) - F(x) - (t-x)F'(x) \quad \text{(A.2)}$$

We wish to prove that $g(0) \geq 0 \implies g(1) < 0$. If $x = y = 1$ then $g(0) = -G'(1) + F'(1) < 0$, so in this case the implication is true. Now assume that x and y are not both 1. By convexity of G

$$g(x) = G(y) + (x-y)G'(y) - F(x) \leq G(x) - F(x) \leq 0, \quad \text{(A.3)}$$

with equality if and only if $x = y = 1$, which we assumed was not the case. Since g is affine and $g(x) < 0$ we have $g(0) \geq 0 \implies g(1) < 0$ as wanted. This is equivalent to

$$F(x) - xF'(x) \leq G(y) - yG'(y) \implies F(x) + (1-x)F'(x) > G(y) + (1-y)G'(y). \quad \text{(A.4)}$$

By Lemma A.1.3 below, with $X = [0,1] \times [0,1]$, $S(x,y) = G(y) - yG'(y) - F(x) + xF'(x)$ and $R(x,y) = F(x) + (1-x)F'(x) - G(y) - (1-y)G'(y)$, we get (A.1). $\square$

Lemma A.1.2. *Let $F, G : [1,\infty) \to \mathbb{R}$ be decreasing, strictly convex and continuously differentiable functions with $F(x) < G(x)$ for all $x \in (1,\infty)$. Assume further that $F(1) = G(1) = 0$,*

$F'(1) < G'(1)$, that $\lim_{x\to\infty} F(x) - xF'(x)$ and $\lim_{y\to\infty} G(y) - yG'(y)$ both exist and that $\lim_{x\to\infty} F'(x) < \lim_{y\to\infty} G'(y)$. Then there exists and $\varepsilon > 0$ such that for all $x, y \in [1,\infty]$

$$F(x) - xF'(x) \leq G(y) - yG'(y) + \varepsilon \implies F(x) + (1-x)F'(x) + \varepsilon \leq G(y) + (1-y)G'(y). \quad (A.5)$$

Here $x = \infty$ or $y = \infty$ is to be interpreted in the sense of limits.

Proof. Like in the proof of Lemma A.1.1, we start by proving (A.5) without the ε and with a sharp inequality on the right-hand-side.

The case $y = \infty$ follows from the fact that $F'(x) < \lim_{y\to\infty} G'(y)$ for all $x \in [1,\infty]$.

Let $x \in [1,\infty]$ and $y \in [1,\infty)$. Like in the proof of Lemma A.1.1, let $g : [0,1] \to \mathbb{R}$ be the affine function

$$g(t) = G(y) + (t-y)G'(y) - F(x) - (t-x)F'(x). \quad (A.6)$$

We wish to prove that $g(0) \geq 0 \implies g(1) > 0$. Again, we don't need to consider the case $x = y = 1$ since then $g(0) = -G'(1) + F'(1) < 0$.

By strict convexity of F,

$$g(y) = G(y) - F(x) - (y-x)F'(x) \geq G(y) - F(y) \geq 0, \quad (A.7)$$

with equality if and only if $x = y = 1$, which we can assume is not the case. Since g is affine and $g(y) > 0$ we have $g(0) \geq 0 \implies g(1) > 0$, which implies

$$F(x) - xF'(x) \leq G(y) - yG'(y) \implies F(x) + (1-x)F'(x) < G(y) + (1-y)G'(y). \quad (A.8)$$

We consider $[1,\infty]$ with the one-point compactification topology. By Lemma A.1.3 with $X = [1,\infty] \times [1,\infty]$,

$$S(x,y) = G(y) - yG'(y) - F(x) + xF'(x)$$

and

$$R(x,y) = G(y) + (1-y)G'(y) - F(x) - (1-x)F'(x),$$

we get (A.5). □

Lemma A.1.3. *Let X be a compact topological space and $S, R : X \to \mathbb{R}$ be continuous functions. If*

$$S(x) \geq 0 \implies R(x) > 0 \quad (A.9)$$

then there is an $\varepsilon > 0$ such that

$$S(x) \geq -\varepsilon \implies R(x) \geq \varepsilon \quad (A.10)$$

Proof. The set $A = \{x \in X | S(x) \geq 0\} \subset X$ is closed and therefore compact, so R takes a minimum value, $\varepsilon_1 > 0$ on A. So

$$S(x) \geq 0 \implies R(x) \geq \varepsilon_1. \tag{A.11}$$

Contraposing, we get

$$S(x) < 0 \impliedby R(x) < \varepsilon_1 \impliedby R(x) \leq \varepsilon_1/2. \tag{A.12}$$

The set $B = \{x \in X | R(x) \leq \varepsilon_1/2\} \subset X$ is compact, so S takes a maximum value $-\varepsilon_2 < 0$ on B. So

$$S(x) < -\varepsilon_2 \impliedby R(x) \leq \varepsilon_1/2. \tag{A.13}$$

Contraposing again yields

$$S(x) > -\varepsilon_2 \implies R(x) > \varepsilon_1/2. \tag{A.14}$$

For $\varepsilon = \min\{\varepsilon_2, \varepsilon_1/2\}$ we now have

$$S(x) \geq -\varepsilon \implies R(x) \geq \varepsilon. \tag{A.15}$$

□

A.2. Bra-ket notation

Given a vector space V and a vector $\psi \in V$ over a field $\mathbb{F}$, we consider the linear map $|\psi\rangle : \mathbb{F} \to V$

$$|\psi\rangle : z \mapsto z\psi. \tag{A.16}$$

This association is a bijection between V and $\mathrm{Hom}(\mathbb{F}, V)$ through the canonical isomorphisms $V \cong \mathbb{F} \otimes V \cong \mathrm{Hom}(\mathbb{F}, V)$. Often we shall write $|\psi\rangle \in V$, understood via this isomorphism between V and $\mathrm{Hom}(\mathbb{F}, V)$. This seems a little ridiculous at first, but turns out to be very convenient notation.

When $\mathbb{F} = \mathbb{C}$, as is usually the case in quantum theory and V has an inner product $\langle \cdot, \cdot \rangle_V$ (conjugate linear in *first* entry!), this provides a natural antilinear bijection $V \ni \psi \leftrightarrow \langle \psi, \cdot \rangle \in V^*$. Given a choice of bijection between V and its dual, the image of $\psi \in V$ under this bijection will be denoted $\langle \psi | : V \to \mathbb{F}$.

Now $\langle \psi |$ and $|\psi\rangle$ are composable maps and we may write

$$|\psi\rangle\langle\psi| = \phi \mapsto \left(\psi^*(\phi)\right)\psi \tag{A.17}$$

for the map that projects onto ψ or

$$\langle\psi||\phi\rangle = \psi^*(\phi) \overset{\text{if inner product}}{=} \langle \psi, \phi \rangle\,, \tag{A.18}$$

for ϕ measured along ψ. $\langle\psi||\phi\rangle$ is often shortened as $\langle\psi|\phi\rangle$.

When $V = \mathbb{C}^X = \mathrm{Hom}(X, \mathbb{C})$, and $x' \in X$, we shall denote the natural basis elements $e_x : x' \mapsto \delta_{x,x'}$ by $|e_x\rangle = |x\rangle$. When $V = \mathbb{C}^X \otimes \mathbb{C}^Y \otimes \mathbb{C}^Z$ we shall also use the notation

$$|xyz\rangle = |x\rangle\,|y\rangle\,|z\rangle = |x\rangle \otimes |y\rangle \otimes |z\rangle\,. \tag{A.19}$$

As an example, consider the unnormalized GHZ state, given by $|000\rangle + |111\rangle = |\psi\rangle$, where ψ is the vector $e_0 \otimes e_0 \otimes e_0 + e_1 \otimes e_1 \otimes e_1$ in the space $\mathbb{C}^{\{0,\ldots,d_1-1\}} \otimes \mathbb{C}^{\{0,\ldots,d_2-1\}} \otimes \mathbb{C}^{\{0,\ldots,d_3-1\}}$, for some $d_1, d_2, d_3 \geq 2$.

Bibliography

[1] E. Chitambar and G. Gour, “Quantum Resource Theories,” *arXiv:1806.06107*, Jun 2018.

[2] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, “Concentrating partial entanglement by local operations,” *Phys. Rev. A*, vol. 53, pp. 2046–2052, Apr 1996.

[3] M. Horodecki, P. Horodecki, and R. Horodecki, “Mixed-state entanglement and distillation: Is there a “bound” entanglement in nature?,” *Phys. Rev. Lett.*, vol. 80, pp. 5239–5242, Jun 1998.

[4] M. Walter, D. Gross, and J. Eisert, *Multipartite Entanglement*, pp. 293–330. Apr 2016.

[5] B. Kraus, “Local unitary equivalence of multipartite pure states,” *Phys. Rev. Lett.*, vol. 104, p. 020504, Jan 2010.

[6] J. I. de Vicente, C. Spee, and B. Kraus, “Maximally entangled set of multipartite quantum states,” *Phys. Rev. Lett.*, vol. 111, p. 110502, Sep 2013.

[7] P. Vrana and M. Christandl, “Entanglement distillation from greenberger–horne–zeilinger shares,” *Communications in Mathematical Physics*, vol. 352, pp. 621–627, Jun 2017.

[8] M. Christandl, P. Vrana, and J. Zuiddam, “Universal points in the asymptotic spectrum of tensors,” *arXiv:1709.07851*, Sept. 2017.

[9] G. Gour and N. R. Wallach, “Necessary and sufficient conditions for local manipulation of multipartite pure quantum states,” *New Journal of Physics*, vol. 13, no. 7, p. 073013, 2011.

[10] Y. Xin and R. Duan, “Conditions for entanglement transformation between a class of multipartite pure states with generalized schmidt decompositions,” *Phys. Rev. A*, vol. 76, p. 044301, Oct 2007.

[11] M. Christandl, A. K. Jensen, and J. Zuiddam, “Tensor rank is not multiplicative under the tensor product,” *Linear Algebra and its Applications*, vol. 543, pp. 125 – 139, 2018.

[12] M. Christandl, F. Gesmundo, and A. Jensen, “Border rank is not multiplicative under the tensor product,” *SIAM Journal on Applied Algebra and Geometry*, vol. 3, no. 2, pp. 231–255, 2019.

[13] V. Strassen, “Degeneration and complexity of bilinear maps: some asymptotic spectra.,” *J. Reine Angew. Math.*, vol. 413, p. 127–180, 1991.

[14] A. K. Jensen and P. Vrana, “The asymptotic spectrum of LOCC transformations,” *arXiv:1807.05130*, July 2018.

[15] V. Strassen, “The asymptotic spectrum of tensors,” *J. Reine Angew. Math.*, vol. 384, pp. 102–152, 1988.

[16] T. Fritz, “Resource convertibility and ordered commutative monoids,” *Mathematical Structures in Computer Science*, vol. 27, no. 6, p. 850–938, 2017.

[17] A. K. Jensen, “Asymptotic majorization of finite probability distributions,” *arXiv:1808.05157*, Aug. 2018.

[18] M. M. Wilde, “From Classical to Quantum Shannon Theory,” *arXiv:1106.1445*, Jun 2011.

[19] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition.* New York, NY, USA: Cambridge University Press, 10th ed., 2011.

[20] I. Csiszar and J. Korner, *Information theory : coding theorems for discrete memoryless systems / Imre Csiszar and Janos Korner.* Academic Press ; Akademiai Kiado New York : Budapest, 1981.

[21] V. Strassen, “Relative bilinear complexity and matrix multiplication,” *J. Reine Angew. Math.*, vol. 375/376, pp. 406–443, 1987.

[22] J. M. Landsberg, *Tensors: geometry and applications*, vol. 128. American Mathematical Society Providence, RI, USA, 2012.

[23] M. Bläser, *Fast Matrix Multiplication.* No. 5 in Graduate Surveys, Theory of Computing Library, 2013.

[24] W. Dür, G. Vidal, and J. I. Cirac, “Three qubits can be entangled in two inequivalent ways,” *Phys. Rev. A*, vol. 62, p. 062314, Nov 2000.

[25] J. Landsberg and M. Michalek, “Abelian tensors,” *Journal de mathématiques pures et appliquées*, vol. 108, pp. 333–371, 9 2017.

[26] P. Vrana and M. Christandl, “Asymptotic entanglement transformation between w and ghz states,” *Journal of Mathematical Physics*, vol. 56, no. 2, p. 022204, 2015.

[27] F. Le Gall, “Powers of tensors and fast matrix multiplication,” in *ISSAC 2014—Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*, pp. 296–303, ACM, New York, 2014.

[28] J. Hopcroft and L. Kerr, “On minimizing the number of multiplications necessary for matrix multiplication,” *SIAM Journal on Applied Mathematics*, vol. 20, no. 1, pp. 30–36, 1971.

[29] D. Bini, G. Lotti, and F. Romani, “Approximate solutions for the bilinear form computational problem,” *SIAM J. Comput.*, vol. 9, pp. 692–697, 11 1980.

[30] J. L. Lagrange, “Leçons élémentaires sur les mathématiques données à l’école normale en 1795,”

[31] E. Whittaker, *The Calculus of Observations - A Treatise on Numerical Mathematics.* Read Books, 2012.

[32] F. Gantmacher, *The theory of matrices. Vol. 2. Transl. from the Russian by K. A. Hirsch. Reprint of the 1959 translation.* Providence, RI: AMS Chelsea Publishing, reprint of the 1959 translation ed., 1998.

[33] D. Y. Grigoriev, “Multiplicative complexity of a pair of bilinear forms and of the polynomial multiplication,” in *Mathematical foundations of computer science, 1978 (Proc. Seventh Sympos., Zakopane, 1978)*, vol. 64 of *Lecture Notes in Comput. Sci.*, pp. 250–256, Springer, Berlin-New York, 1978.

[34] J. JáJá, “Optimal evaluation of pairs of bilinear forms,” *SIAM J. Comput.*, vol. 8, no. 3, pp. 443–462, 1979.

[35] L. Teichert, “Die Komplexität von Bilinearformpaaren über beliebigen Körpern..” Mathematisch-Naturwissenschaftliche Fakultät der Technischen Universität Clausthal, 1986.

[36] P. Bürgisser, M. Clausen, and M. A. Shokrollahi, *Algebraic complexity theory*, vol. 315 of *Grundlehren Math. Wiss.* Springer-Verlag, Berlin, 1997.

[37] J. Zuiddam, “The asymptotic spectrum of graphs and the Shannon capacity,” *arXiv:1807.00169*, June 2018.

[38] N. McNew, “On sets of integers which contain no three terms in geometric progression,” *Mathematics of Computation*, 10 2013.

[39] M. Hayashi, M. Koashi, K. Matsumoto, F. Morikoshi, and A. Winter, "Error exponents for entanglement concentration," *J. Phys. A*, vol. 36, no. 2, pp. 527–553, 2003.

[40] C. H. Bennett, S. Popescu, D. Rohrlich, J. A. Smolin, and A. V. Thapliyal, "Exact and asymptotic measures of multipartite pure-state entanglement," *Phys. Rev. A*, vol. 63, p. 012307, Dec 2000.

[41] M. Żukowski, A. Zeilinger, M. A. Horne, and H. Weinfurter, "Quest for ghz states," *Acta Phys. Pol. A*, vol. 93, pp. 187–195, 1998.

[42] D. P. DiVincenzo, C. A. Fuchs, H. Mabuchi, J. A. Smolin, A. Thapliyal, and A. Uhlmann, "Entanglement of assistance," in *Quantum Computing and Quantum Communications* (C. P. Williams, ed.), (Berlin, Heidelberg), pp. 247–257, Springer Berlin Heidelberg, 1999.

[43] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels," *Phys. Rev. Lett.*, vol. 70, pp. 1895–1899, Mar 1993.

[44] J.-C. Bourin and F. Hiai, "Norm and anti-norm inequalities for positive semi-definite matrices," *International Journal of Mathematics*, vol. 22, Jan 2012.

[45] M. A. Nielsen, "Conditions for a class of entanglement transformations," *Phys. Rev. Lett.*, vol. 83, pp. 436–439, Jul 1999.

[46] A. Streltsov, H. Kampermann, S. Wölk, M. Gessner, and D. Bruß, "Maximal coherence and the resource theory of purity," *New Journal of Physics*, vol. 20, p. 053058, may 2018.

[47] F. Brandão, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner, "The second laws of quantum thermodynamics," *Proceedings of the National Academy of Science*, vol. 112, pp. 3275–3279, Mar 2015.

[48] M. Horodecki and O. J., "Fundamental limitations for quantum and nanoscale thermodynamics," *Nat. Commun.*, vol. 4:2059 doi: 10.1038/ncomms3059 (2013).

[49] N. Canosa, R. Rossignoli, and M. Portesi, "Majorization properties of generalized thermal distributions," *Physica A: Statistical Mechanics and its Applications*, vol. 368, pp. 435–441, 08 2006.

[50] M. Klimesh, "Inequalities that Collectively Completely Characterize the Catalytic Majorization Relation," *arXiv:0709.3680*, Sep 2007.

[51] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and its Applications*, vol. 143. Springer, second ed., 2011.