ECOLE
POLYTECHNIQUE
DE BRUXELLES

ULB

UNIVERSITÉ LIBRE DE BRUXELLES

Majorization Lattice in the Theory of Entanglement

Mémoire présenté en vue de l'obtention du diplôme
d'Ingénieur Civil Physicien à finalité spécialisée

Alexander Stévins

Directeur
Professeur Nicolas Cerf

Superviseur
Serge Deside

Service
QuIC

Année académique
2024 - 2025

Abstract

Keywords:

The theory of majorization has been used in a variety of fields to compare probability distributions, and has been the key to several theorems in quantum information. In this master thesis, the lattice generated by the majorization relations on the set of n-dimensional probability vectors has been explored, and properties of the lattice have proved to be useful in the study of quantum entanglement. The first result of this thesis is a corollary to Nielsen and Kempe's majorization criterion for separability, which guarantees that a bipartite quantum state is entangled if the vector of eigenvalues of its density matrix has a lower entropy than the meet of the vectors of eigenvalues of its subsystems. This result has been shown to be true not only for the Shannon entropy, but for all Renyi entropies. This is particularly interesting in the case of the collision entropy, which is directly related to the purity of the quantum state, which is easier to access and to compute than the Shannon entropy of the state. The second result of this thesis is the definition of two new complementary measures of comparability between probability vectors defined using Cicalese and Vaccaro's notion of distance between probability vectors on the majorization lattice. This might be useful in a cryptographic setting, where the amount of comparability between the eigenvalues of the two reduced density matrices of a bipartite system might be the source of a new resource theory.

Author First name Last name
Section name
Title
Academic Year

Contents

Part I

Theoretical Background

Chapter 1

Majorization Theory

1.1 Majorization relations

1.1.1 Cumulative sum definition

The theory of majorization gives a framework for what it means for a probability distribution (or more general objects) to be more disordered, more random than another one. A rigorous notion of disorder such as the one that majorization suggests is interesting because it should be linkable to other objects that attempt to quantify randomness in information theory, and most notably Shannon entropy. In order to see that this is the case, we first need to rigorously define the majorization relation.

Let p be a probability vector, *i.e.* a vector such that $\sum_i^d p_i = 1$ for a vector of dimension d. Let $p^{\downarrow}$ be the non-increasing reordering of p, meaning that the entries of $p^{\downarrow}$ are the same as those of p, but sorted such that $p_1^{\downarrow} \geq p_2^{\downarrow} \geq ... \geq p_d^{\downarrow}$. We will always be considering reordered vectors in this Ms thesis, and we will therefore denote the set of n-dimesional probability vectors sorted in nonincreasing order $\mathcal{P}^n$. Formally, $\mathcal{P}^n = \{p \in \mathbb{R}^n | \sum_i^n p_i = 1, p_i \geq p_{i+1}\}$.

Definition 1 (Cumulative sum)**.** Let p be a vector in $\mathcal{P}^d$. The j^{th} decreasing[1] cumulative sum of p is the sum of the j biggest components of p, which can be written as

$$S_j^{\downarrow}(p) \coloneqq \sum_{i=1}^{j} p_i^{\downarrow} \tag{1.1}$$

where $p_i^{\downarrow}$ is the i^{th} largest entry of p.

Definition 2 (Majorization relation)**.** Let $p, q \in \mathbb{R}^d$ be two vectors of dimension d. We say that p majorizes q, written $p \succ q$, if and only if

$$\begin{cases} S_j^{\downarrow}(p) \geq S_j^{\downarrow}(q) & \forall j = 1, ..., d-1 \\ S_d^{\downarrow}(p) = S_d^{\downarrow}(q) \end{cases} \tag{1.2}$$

If the dimensions of the vectors do not match, one can always append zeroes to the one of lower dimension and apply the definition on the enlarged vector.

Remark 1. If p and q are probability vectors, then the last equality is automatically verified because $S_d^{\downarrow}(p) = S_d^{\downarrow}(q) = 1$.

[1] We could also have defined $\mathcal{P}^d$ to be increasing vectors and use increasing cumulative sums $S_i^{\uparrow}(p)$ instead. We would get an equivalent description of majorization, the only effect being reversed inequalities in (1.2).

This definition encapsulates the idea of disorder : if $p \succ q$, then the largest probability of p ($p_1^\downarrow$) is larger than the largest probability of q ($q_1^\downarrow$), and the rest of the distribution in p is lowered for the sum of probabilities to give one, whereas q has a distribution that is more spread out. As an example, consider $p = (0.9, 0.1)$ and $q = (0.6, 0.4)$, and applying definition 2 we see that $p \succ q$. This feels fairly intuitive : p is not as uncertain as q, because a process described by probability distribution p sees one of the 2 events happen most of the time, whereas the outcome of a process with probability distribution q is almost equally likely to be one or the other. It is clear from this example that the *majorizer* is more ordered than the *majorized.*

Remark 2. For any $p \in \mathcal{P}^d$, we have

$$(\frac{1}{d}, ..., \frac{1}{d}) \prec p \prec (1, 0, ..., 0) \tag{1.3}$$

which is fairly intuitive, considering that $(\frac{1}{d}, ..., \frac{1}{d})$ is the most uncertain distribution and $(1, 0, ..., 0)$ is the most certain distribution.

The majorization relation creates a preorder on probability vectors. When either $p \succ q$ or $p \prec q$ is verified, we will say that p and q are *comparable* (sometimes written $p \sim q$). However, there also exist cases where both $p \nsucc q$ and $p \nprec q$ are true, and we will then say that p and q are *incomparable* (sometimes written $p \nsim q$). For example, $(0.5, 0.25, 0.25) \nsim (0.4, 0.4, 0.2)$.

Remark 3. Two vectors of dimension 2 are always comparable. Incomparability is only possible from dimension 3 and above.

1.1.2 Link with bistochastic matrices

A second equivalent definition for majorization can be obtained by a different intuition involving bistochastic matrices. The idea is the following : if the probability vector q can be obtained by randomly mixing the components of p together, then p is more ordered than q. It turns out, this simple idea is equivalent to the majorization relation. Let us express this rigorously.

Definition 3 (Bistochastic matrix)**.** A bistochastic matrix D is a matrix in $\mathbb{R}^{d\times d}$ such that

$$\begin{cases} D_{i,j} \geq 0 \quad \forall i, j = 1, ..., d \\ \sum_{i=1}^{d} D_{i,j} = 1 \quad \forall j = 1, ..., d \\ \sum_{j=1}^{d} D_{i,j} = 1 \quad \forall i = 1, ..., d \end{cases} \tag{1.4}$$

A bistochastic matrix (also called doubly stochastic matrix) is basically a matrix with nonnegative entries such that the columns and rows each sum to one. The following theorem formalizes the way that bistochastic matrices mix entries of vectors together.

Theorem 1 (Birkhoff's theorem)**.** *If D is a d-dimensional bistochastic matrix, then there exists a probability distribution $\{p_j\}$ and a set of d-dimensional permutation matrices P_j such that*

$$D = \sum_j p_j P_j \tag{1.5}$$

From this theorem, we can see that the effect of a bistochastic matrix on a vector is essentially to perform a convex mixture of the vector's entries, producing a vector that is more smoothed out, and thus more disordered.

Theorem 2 (Hardy, Littlewood and Pólya)**.** *Let $p, q \in \mathcal{P}^d$. Then, $p \succ q$ if and only if there exists a d-dimensional bistochastic matrix D such that*

$$q = Dp \tag{1.6}$$

One could take theorem 2 as a definition of the majorization relation, as the notion of a vector being more disordered is perhaps clearer in the bistochastic matrix picture. Then, the nature of the convex combination of entries would have given us the set of inequalities (1.2). Let's take the same example as earlier, $p = (0.9, 0.1)$ and $q = (0.6, 0.4)$. Then we can find that $D = \left(\begin{smallmatrix} 5/8 & 3/8 \\ 3/8 & 5/8 \end{smallmatrix}\right)$ gives us $q = Dp$, and therefore $p \succ q$.

Whichever picture one prefers intuitively, the fact that there is a full equivalence between the two is quite powerful when working on proofs, as sometimes one picture is easier to work with than the other depending on the context.

Remark 4. The bistochastic matrix $\begin{pmatrix} \frac{1}{d} & \cdots & \frac{1}{d} \\ \vdots & & \vdots \\ \frac{1}{d} & \cdots & \frac{1}{d} \end{pmatrix}$ produces the probability vector $(\frac{1}{d}, ..., \frac{1}{d})$ regardless of the probability vector on which it is applied, hence $(\frac{1}{d}, ..., \frac{1}{d})$ is majorized by every probability vector.

1.1.3 Lorenz curves

Lorenz curves were first introduced in 1905 by economist Max Lorenz as a way to represent income inequality. The idea is the following : given a distribution, plot the cumulative sum of its largest entries. Doing this for two distributions, we can then say that if one curve is above the other at all times, then it is more unequal (more disordered) than the second. By normalizing the distributions for total income, this graphical tool could then be used to quickly compare income inequality between different regions or countries.

It turns out that this notion of unequality is precisely the same as majorization, as at each coordinate the plotted curves explicitely show one of the inequalities of definition 2. More formally, the Lorenz curve of a vector $p \in \mathbb{R}^d$ is obtained by linear interpolation between the points of the set $\{(i, S_i^{\downarrow}(p)) | i \in \mathbb{N}, i \leq d\}$. The alias $L_p^{\downarrow}(x) \coloneqq S_x^{\downarrow}(p)$ is an alternative notation used for Lorenz curves, where the abscissa goes in the argument of the function. In this report, Lorenz curves will mostly be useful in sections 1.3.3 and 1.3.4 to understand how the algorithm for constructing key vectors in the lattice work.

A few examples of Lorenz curves are given in figures 1.1 and 1.2.

1.2 Schur-convex and Schur-concave functions

1.2.1 Definition

Some functions preserve the ordering defined by the majorization relation and are therefore interesting to study because a majorization relation directly implies an inequality on these functions. The following few definitions will be useful.

Definition 4 (Convex function)**.** A function $f \in C(\mathcal{I} \to \mathbb{R})$ is convex on an interval $\mathcal{I}$ iff for all $x_1, x_2 \in \mathcal{I}$

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2) \quad \forall \alpha \in [0, 1]. \tag{1.7}$$

Definition 5 (Schur-convex function)**.** A function $f \in C(\mathbb{R}^d \to \mathbb{R})$ is Schur-convex iff

$$p \prec q \implies f(p) \leq f(q) \tag{1.8}$$

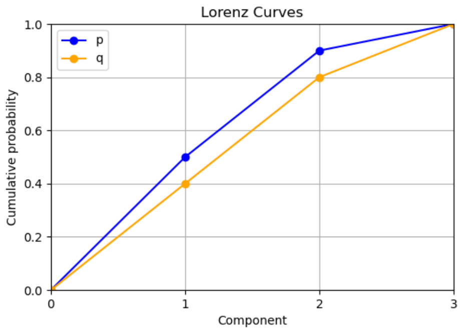

Figure 1.1: Lorenz curve for an example with comparable vectors $p = (0.5, 0.4, 0.1)$, $q = (0.4, 0.4, 0.2)$ example. Notice that comparable vectors must have Lorenz curves that do not cross each other.

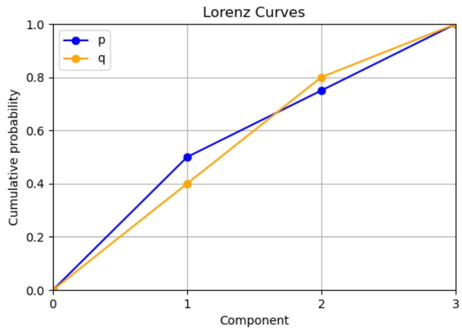

Figure 1.2: Lorenz curve for the incomparable vector $p = (0.5, 0.25, 0.25)$, $q = (0.4, 0.4, 0.2)$ example. Notice that incomparable vectors must have Lorenz curves that cross each other at least once.

Definition 6 (Concave function)**.** A function $f \in C(\mathcal{I} \to \mathbb{R})$ is concave on an interval $\mathcal{I}$ iff for all $x_1, x_2 \in \mathcal{I}$

$$f(\alpha x_1 + (1-\alpha)x_2) \geq \alpha f(x_1) + (1-\alpha) f(x_2) \quad \forall \alpha \in [0,1]. \tag{1.9}$$

Definition 7 (Schur-concave function)**.** A function $f \in C(\mathbb{R}^d \to \mathbb{R})$ is Schur-concave iff $-f$ is Schur-convex.

Additional properties links with convex and concave functions exist, such as inequality (C.1) from Ref. [6, p. 92], often called Karamata's inequality.

Lemma 1 (Karamata's inequality)**.** *If $f \in C(\mathcal{I} \to \mathbb{R})$ is convex on $\mathcal{I}$, and if $p \succ q$, then*

$$\sum_{i=1}^{d} f(p_i) \geq \sum_{i=1}^{d} f(q_i). \tag{1.10}$$

1.2.2 Shannon entropy

The Shannon entropy function $H(p) := -\sum_{i=1}^{d} p_i \log p_i$ is a foundational quantity in classical information theory (cf. section 2.1). $H(p)$ essentially captures the uncertainty content of a random process distributed over a probability distribution p. We therefore expect it to be linked to majorization theory which gives a partial ordering on the uncertainty of different probability distributions. The following lemma establishes this link through Schur-concavity.

Lemma 2 (Schur-concavity of the Shannon entropy)**.** *For any $p, q \in \mathcal{P}^d$,*

$$p \succ q \implies H(p) \leq H(q). \tag{1.11}$$

This property of Shannon entropy is fairly intuitive: Shannon entropy captures the uncertainty content of a probability distribution, and $p \succ q$ means that p is unequivocally[2] more ordered (so more certain) than q, so we would expect Schur-concavity to hold. However a majorization relation is stronger than an entropic inequality, and so we should study majorization relations alongside Shannon entropy to gain a better understanding of information theory.

[2] As opposed to incomparable distributions being more ordered than the other each in their own way.

1.3 Majorization lattice

1.3.1 Lattice structures

The majorization relation creates a preorder on probability vectors. When restricted to *ordered* probability vectors, the majorization relation creates a proper partial order. With the ordering induced by the majorization relation, the set of ordered vectors $\{z|z_1 \geq z_2 \geq ... \geq z_d\}$ becomes a lattice.

Definition 8 (Lattice)**.** A lattice is a quadruple $\langle \mathcal{L}, \sqsubseteq, \wedge, \vee \rangle$ where $\mathcal{L}$ is a set, $\sqsubseteq$ is a partial ordering $\mathcal{L}$, and for all $a, b \in \mathcal{L}$ there is a unique greatest lower bound (glb) $a \wedge b$ and a unique least upper bound (lub) $a \vee b$. That is, $a \wedge b \sqsubseteq a$, $a \wedge b \sqsubseteq b$, $a \sqsubseteq a \vee b$, $b \sqsubseteq a \vee b$, and for each c, d such that $c \sqsubseteq a$, $c \sqsubseteq b$, $a \sqsubseteq d$, $b \sqsubseteq d$ we have that

$$c \sqsubseteq a \wedge b \quad \text{and} \quad a \vee b \sqsubseteq d. \tag{1.12}$$

A simple example of a lattice is the set $\mathbb{N}$ of natural numbers ordered by the relation "divides". Then for $a, b \in \mathbb{N}$ the lub is l.c.m.(a, b) and the glb is g.c.d.(a, b), where l.c.m. = least common multiple and g.c.d. = greatest common divisor.

It was shown that the set $\mathcal{P}^d$ endowed with the partial ordering $\prec$ is a lattice [1]. As the lattice is a central part of this Ms thesis, we will go through the main ideas behind the proof, which hinges on defining the correct greatest lower bound (glb) $p \wedge q$, which we will call the *meet* of p and q, and the correct least upper bound (lub) $p \vee q$, which we will call the *join*.

Let us first go through what it means for a vector to be the greatest lower bound, the meet, of two vectors $p, q \in \mathcal{P}^d$. Intuitively, the glb of p and q in terms of majorization is the least disordered vector that is majorized by both p and q. While this may not seem very interesting at a first glance, recall that two vectors may be incomparable, meaning that both $p \not\prec q$ and $p \not\succ q$ are true. Even if they are incomparable, the subset of vectors that are majorized by both p and q is nonempty[3] and has a supremum, $p \wedge q$, *which majorizes every other vector in the subset.* Note that an equivalent definition of a lattice is to require that every partially ordered subset has a supremum and infimum [6, p. 19]. Since the two definitions are equivalent, we can consider this to be a property of the lattice, guaranteeing the existence of the supremum $p \wedge q$.

In similar fashion, the lub, the join, of two vectors $p, q \in \mathcal{P}^d$ is the least ordered vector that majorizes both p and q. The subset of vectors that majorize both p and q is nonempty[4] and has an infimum, $p \vee q$, *which is majorized by every other vector in the subset.* Moreover, the same property as before guarantees the existence of the infimum.

Remark 5. If two vectors are comparable, say $p \prec q$ (the $p \succ q$ case being symmetric), then the question becomes trivial as the least ordered vector that majorizes both p and q is q and the least disordered vector that is majorized by both p and q is p, i.e. $p \wedge q = p$ and $p \vee q = q$.

1.3.2 Majorization cones and geometric intuition

Before going further with the proof, let us first give a 2D visual representation of the lattice, which is very helpful in building intuition for using the lattice, first proposed in the field of

[3] It is guaranteed to at least contain the uniform distribution $(\frac{1}{d}, ..., \frac{1}{d})$ since it is majorized by every other distribution.

[4] It is guaranteed to at least contain the certain distribution $(1, 0, ..., 0)$ since it majorizes every other distribution.

thermomajorization [5]. More accurate representations exist but are much more complicated, living in the canonical Weyl chamber of a Δ_{d-1} simplex [4]. Nevertheless, the simple 2D depiction is very useful to think about, even if one has to be careful not to draw hasty conclusions from it.

Definition 9 (Majorization cones)**.** The set of states that majorize p is called the *future cone*[5] $\mathcal{T}_+(p)$. The set of states that are majorized by p is called the *past cone* $\mathcal{T}_-(p)$. The set of states that are neither in the past nor the future cone of p is called the *incomparable region* $\mathcal{T}_\emptyset(p)$.

Figure 1.3 shows an example of the 2D representation of a vector $p \in \mathcal{P}^d$ in the lattice, along with some definitions we introduce now, and figure 1.4 shows the intersections of the majorization cones of two vectors $p, q \in \mathcal{P}^d$ and their meet and join.

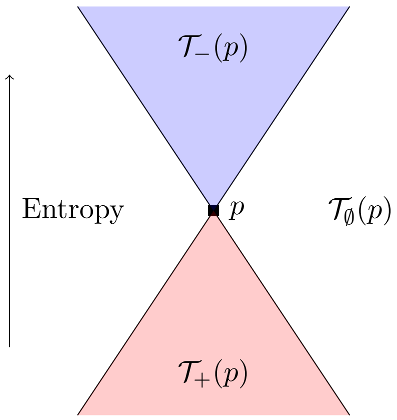

Figure 1.3: Depiction of the majorization cones of a nondescript vector $p \in \mathcal{P}^d$.

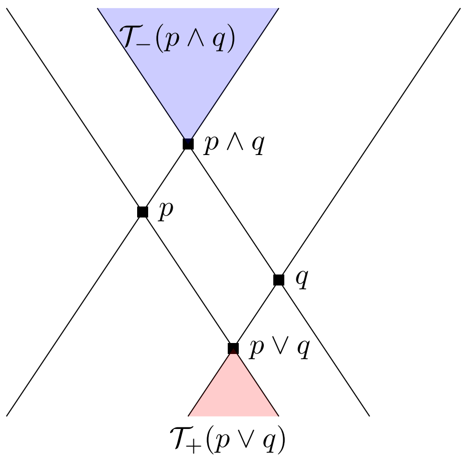

Figure 1.4: Depiction of the intersection of the majorization cones of incomparable vectors $p, q \in \mathcal{P}^d$. Note that by definition of the meet and join, $\mathcal{T}_+(p) \cap \mathcal{T}_+(q) = \mathcal{T}_+(p \vee q)$ and $\mathcal{T}_-(p) \cap \mathcal{T}_-(q) = \mathcal{T}_-(p \wedge q)$.

In quantum information, the lattice is usually drawn with more disordered vectors at the top and more ordered vectors at the bottom because the entropy of a state is considered a

[5]It is important to note that our naming conventions and notations are for the most part entirely opposite to the thermomajorization conventions for reasons that will be made clear in section 2.4.2. Most notably, our future cone corresponds to their past cone and vice versa.

resource (cf. section 2.4.1). This means that the uniform distribution sits at the very top of the diagram, and the certain distribution at the very bottom. One should therefore be careful, because the majorization relation runs from top to bottom, which is the opposite of what one would picture at first (a greatest lower bound looks like a least upper bound and vice-versa). While this is inconvenient for this chapter, this convention will make more sense later down the road.

1.3.3 Constructing the meet

In order to show that the quadruple $\langle \mathcal{P}^d, \prec, \wedge, \vee \rangle$ is indeed a lattice, we need to find an appropriate definition for the glb $p \wedge q$ and lub $p \vee q$ of two arbitrary elements p and q in $\mathcal{P}^d$. Let us start with the meet $p \wedge q$. Thinking in terms of Lorenz curves is actually very helpful here and can lead us very intuitively to the correct algorithm for constructing the meet. Figure 1.5 shows the vectors $p = (0.6, 0.2, 0.2)$ and $q = (0.45, 0.45, 0.1)$, both in $\mathcal{P}^3$.

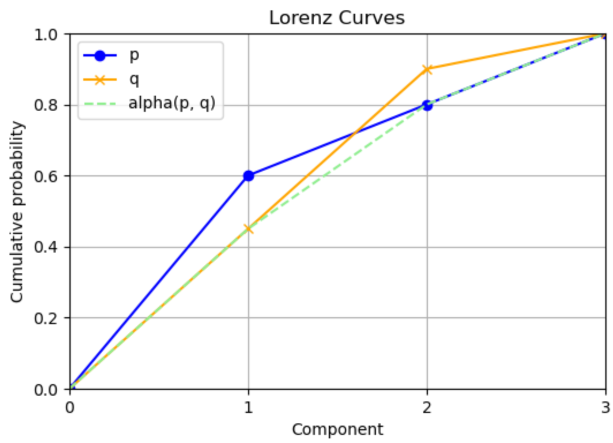

Figure 1.5: Representation of the $p = (0.6, 0.2, 0.2)$ and $q = (0.45, 0.45, 0.1)$ example, along with the proposed $\alpha(p, q)$ lub vector.

Intuitively, the most ordered state that is majorized by both p and q should have a Lorenz curve that hugs the curves of p and q from below. Let $\alpha(p, q) = (a_1, ..., a_d)$ be this state. We can construct it the following way:

$$a_i = \min\Big\{\sum_{j=1}^{i} p_j, \sum_{j=1}^{i} q_j\Big\} - \sum_{j=1}^{i-1} a_j \tag{1.13}$$

$$= \min\Big\{\sum_{j=1}^{i} p_j, \sum_{j=1}^{i} q_j\Big\} - \min\Big\{\sum_{j=1}^{i-1} p_j, \sum_{j=1}^{i-1} q_j\Big\}. \tag{1.14}$$

Lemma 3 (Cicalese and Vaccaro, 2002 [1])**.** *For all $p, q \in \mathcal{P}^d$ we have $\alpha(p, q) = p \wedge q$.*

1.3.4 Constructing the join

We can proceed in a similar way for the join of two vectors $p \vee q$, and thinking in terms of Lorenz curves will once again give us the right idea to move forward. Figure X shows the same vectors $p = (0.6, 0.2, 0.2)$ and $q = (0.45, 0.45, 0.1)$, both in $\mathcal{P}^3$. This time, the least ordered vector that still majorizes both p and q should have a Lorenz curve that hugs their curves from above. Let us define the vector $\beta(p, q) = (b_1, ..., b_d)$ as

$$b_i = \max\left\{\sum_{j=1}^{i} p_j, \sum_{j=1}^{i} q_j\right\} - \sum_{j=1}^{i-1} b_j \tag{1.15}$$

$$= \max\left\{\sum_{j=1}^{i} p_j, \sum_{j=1}^{i} q_j\right\} - \max\left\{\sum_{j=1}^{i-1} p_j, \sum_{j=1}^{i-1} q_j\right\}. \tag{1.16}$$

While this is a step in the right direction, things are unfortunately more complicated than for the meet, as this vector is not guaranteed to be sorted in nonincreasing order and is thus not necessarily in $\mathcal{P}^d$. Let us define $\beta'(p, q) = (\beta(p, q))^{\downarrow} \in \mathcal{P}^d$. Let us take $p = (0.6, 0.15, 0.15, 0.1)$ and $q = (0.5, 0.25, 0.20, 0.05)$ and then

$$\beta(p, q) = (0.6, 0.15, 0.2, 0.05) \notin \mathcal{P}^4 \tag{1.17}$$

$$\beta'(p, q) = (0.6, 0.2, 0.15, 0.05) \in \mathcal{P}^4. \tag{1.18}$$

Figure 1.6 shows the Lorenz curves of $p, q, \beta(p, q)$ and $\beta'(p, q)$. It is immediate to see that $p, q \prec \beta'(p, q)$, but while $\beta'(p, q)$ is indeed an upper bound on p and q it does not necessarily majorize all of the vectors in $\mathcal{T}_{-}(p) \cap \mathcal{T}_{-}(q)$. Intuitively, we can still find a Lorenz curve below $\beta'(p, q)$ by choosing the shortest chord (a straight line) between b'_1 and b'_3, i.e. by smoothing over the concave dent in $\beta(p, q)$. This idea yields the vector $\beta''(p, q) = (0.6, 0.175, 0.175, 0.05)$ which clearly majorizes $\beta'(p, q)$, as shown in figure 1.7.

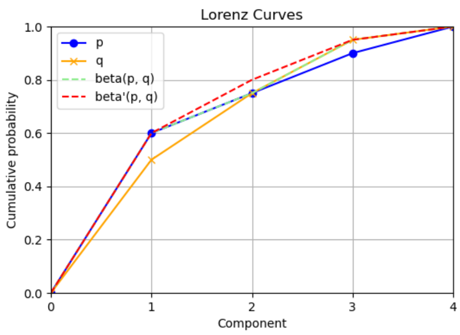

Figure 1.6: Lorenz curves for the vector $p = (0.6, 0.15, 0.15, 0.1)$ and the vector $q = (0.5, 0.25, 0.20, 0.05)$, along with the join prototypes $\beta(p, q)$ and $\beta'(p, q)$.

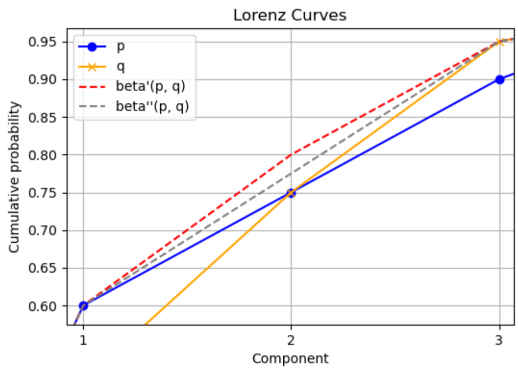

Figure 1.7: Zoom on the region that visually shows that the $\beta'(p, q)$ vector is not the join, as there exist vectors such as $\beta''(p, q)$ that are majorized by it but still majorize both p and q.

We have the following statement:

$$p, q \prec (0.6, 0.175, 0.175, 0.05) \prec \beta'(p, q). \tag{1.19}$$

The following definition provides the smoothing algorithm to apply to $\beta(p, q)$ to obtain the join of p and q.

Definition 10 (Concave dent smoothing algorithm)**.** Let $b = (b_1, ..., b_d)$, and let j be the smallest integer in $\{2, ..., d\}$ such that $b_j > b_{j-1}$. Moreover, let i be the greatest integer in $\{1, 2, ..., j-1\}$ such that

$$b_{i-1} \geq \frac{\sum_{r=1}^{j} b_r}{j - i + 1} =: a. \tag{1.20}$$

Let the probability distribution $c = (c_1, ..., c_d)$ be defined as

$$c_r = \begin{cases} a, & \text{for } r = i, i+1, ..., j \\ b_r & \text{otherwise.} \end{cases} \tag{1.21}$$

This vector is the same as the original vector, but where the *first* concave dent found in the Lorenz curve has been smoothed out.

Lemma 4 (Cicalese and Vaccaro, 2002 [1])**.** *For any $p, q \in \mathcal{P}^d$ we obtain $p \vee q$ in at most $d-1$ iterations*[6] *of the algorithm described in definition 10 applied to the vector $\beta(p, q)$.*

Lemmas 3 and 4 conclude the proof sketch for the set $\mathcal{P}^d$ endowed with the majorization relation being a lattice, as we have found the lub and glb of any two vectors in $\mathcal{P}^d$. Working implementations of the join and meet algorithms in Python are available on the GitHub for the project, which consists of a small library for majorization-related tasks such as plotting Lorenz curves or applying bistochastic matrices on probability vectors. The library can be used with QuTip which is a Python quantum information library for generating quantum state representations.

1.3.5 Properties of the meet and join

In any lattice, the meet and the join can also be seen as two operations[7] acting on elements of a set $\mathcal{L}$, which enjoy some useful properties [3, p. 46].

- Commutativity: $a \wedge b = b \wedge a,\ a \vee b = b \vee a \quad \forall a, b \in \mathcal{L}$.
- Associativity: $a \wedge (b \wedge c) = (a \wedge b) \wedge c,\ a \vee (b \vee c) = (a \vee b) \vee c \quad \forall a, b, c \in \mathcal{L}$.
- Absorption: $a \wedge (a \vee b) = a,\ a \vee (a \wedge b) = a \quad \forall a, b \in \mathcal{L}$.

While it is fairly clear how the algorithms from sections 1.3.3 and 1.3.4 lead to commutativity and associativity, it might not be as clear why they lead to the absorption rule. Recall definition 8 and the interpretation of glb and lub of the meet and join respectively. Then it becomes clearer that the first absorption rule for the majorization lattice states that *the most ordered vector that is majorized by both a and some vector that majorizes a (and b) is none other than a.* Similarly, the second absorption rules states that *the least ordered vector that majorizes both a and some vector that is majorized by a (and b) is a.* Yet another way to understand it is that this boils down to remark 5: a is necessarily comparable to $a \vee b$ (resp. $a \wedge b$) since by definition $a \prec a \vee b$ (resp. $a \succ a \wedge b$), and when two vectors are comparable their meet (resp. join) is simply the least ordered (resp. most ordered) of the two vectors, being a in this case.

These properties are extremely helpful when working on proofs on the lattice, and will appear quite a bit in the second part of this Ms thesis.

1.4 Properties of the Shannon entropy on the lattice

The Schur-concavity of the Shannon entropy, which preserves the ordering of the majorization relation, has already been discussed in section 1.2.2. However, the Shannon entropy enjoys additional properties on the majorization lattice: *supermodularity* and *subadditivity.*

[6] Smoothing a concave dent can lead to a concave dent with the following point, so there is at most $d-1$ concave dents to smoothe.

[7] This is actually another alternative definition of a lattice, which can be defined as a triple $\langle \mathcal{L}, \wedge, \vee \rangle$ with operations $\wedge, \vee$ acting on elements of $\mathcal{L}$. If they satisfy the properties listed above, an ordering $\sqsubseteq$ can then be derived from the triple [3, p. 47].

1.4.1 Supermodularity

Definition 11. A real-valued function f defined on a lattice $\langle \mathcal{L}, \sqsubseteq, \wedge, \vee \rangle$ is *supermodular* iff for any $a, b \in \mathcal{L}$

$$f(a \wedge b) + f(a \vee b) \geq f(a) + f(b). \tag{1.22}$$

Theorem 3 (Supermodularity of the Shannon entropy [1])**.** *The entropy function H is supermodular on the majorization lattice*[8]*, and so for all $p, q \in \mathcal{P}^d$,*

$$H(p \wedge q) + H(p \vee q) \geq H(p) + H(q). \tag{1.23}$$

Note that if $p \sim q$, there is trivially equality.

Proof. Of the 2 properties, this is the most useful one, so let us go through the proof. We will prove the slightly stronger result on $\beta(p,q)$ as defined in (1.16):

$$H(p \wedge q) + H(\beta(p,q)) \geq H(p) + H(q). \tag{1.24}$$

By Shur-concavity of Shannon entropy we have $\beta^{\downarrow}(p,q) \succ p \vee q \implies H(\beta(p,q)) \leq H(p \vee q)$, so showing (1.24) directly gives us the desired result (1.23). The original proof is quite long and hinges on unintuitive index tricks, however the first result of this master thesis is to show an alternative, shorter proof of eq. (1.24). Define

$$r = p \wedge q, s = \beta(p,q), \tag{1.25}$$
$$A = (p_1, p_2, ..., p_d, q_1, q_2, ..., q_d)^{\downarrow} \in \mathbb{R}^{2d}, \tag{1.26}$$
$$B = (r_1, r_2, ..., r_d, s_1, s_2, ..., s_d)^{\downarrow} \in \mathbb{R}^{2d}. \tag{1.27}$$

We have $\sum_{i=1}^{2d} A_i = \sum_{i=1}^{2d} B_i = 2$. We need to show that $A \succ B$. Clearly, the k^{th} cumulative sum of A is made up of the biggest entries of p and q, and the same goes for cumulative sums of B being made of the biggest entries of r and s. Let us additionally introduce the convention that $S_k^{\downarrow}(v) = 1 \ \forall k > d$ for any d-dimensional probability vector v. For all $k \leq 2d$, we have

$$S_k^{\downarrow}(A) = \max_{l \leq k}(S_l^{\downarrow}(p) + S_{k-l}^{\downarrow}(q)), \tag{1.28}$$
$$S_k^{\downarrow}(B) = \max_{m \leq k}(S_{k-m}^{\downarrow}(r) + S_m^{\downarrow}(s)). \tag{1.29}$$

Recall equations (1.13) and (1.15), from which we immediately deduce

$$S_i^{\downarrow}(r) = \min\{S_i^{\downarrow}(p), S_i^{\downarrow}(q)\}, \tag{1.30}$$
$$S_i^{\downarrow}(s) = \max\{S_i^{\downarrow}(p), S_i^{\downarrow}(q)\}. \tag{1.31}$$

Let us fix m. There are two possible cases, either $S_m^{\downarrow}(p) \geq S_m^{\downarrow}(q)$ or $S_m^{\downarrow}(p) \leq S_m^{\downarrow}(q)$. Let us consider the $S_m^{\downarrow}(p) \geq S_m^{\downarrow}(q)$ case, which gives us

$$S_k^{\downarrow}(B) = S_m^{\downarrow}(p) + \min\{S_{k-m}^{\downarrow}(p), S_{k-m}^{\downarrow}(q)\} \tag{1.32}$$
$$\leq S_m^{\downarrow}(p) + S_{k-m}^{\downarrow}(q) \tag{1.33}$$
$$\leq \max_{l \leq k}(S_l^{\downarrow}(p) + S_{k-l}^{\downarrow}(q)) = S_k^{\downarrow}(A), \tag{1.34}$$

[8] It is important to specify which lattice: on the boolean lattice, which gives a different ordering and different meet and join operations, the Shannon entropy is *submodular* instead.

which is precisely the majorization relation. The $S_m^{\downarrow}(p) \leq S_m^{\downarrow}(q)$ case being symmetric, and this being true for all $k \leq 2d$, we have shown that $A \succ B$. Finally, let $\phi(x) = -x\log(x)$, which is concave on $[0,1]$. Using Lemma 1 on ϕ ($-\phi$ being convex on the same interval), we get

$$\sum_{i=1}^{2d} \phi(A_i) \leq \sum_{i=1}^{2d} \phi(B_i). \tag{1.35}$$

But the LHS of (1.35) is precisely $H(p)+H(q)$, and the RHS is precisely $H(p\wedge q)+H(\beta(p,q))$, so we have proven (1.24). □

1.4.2 Subadditivity

Definition 12 (Subadditivity)**.** A real-valued function f defined on a lattice $\langle \mathcal{L}, \sqsubseteq, \wedge, \vee \rangle$ is *subadditive* iff for any $a, b \in \mathcal{L}$

$$f(a \wedge b) \leq f(a) + f(b). \tag{1.36}$$

Theorem 4 (Subadditivity of the Shannon entropy, [1])**.** *The entropy function H is subadditive on the majorization lattice, so for all $p, q \in \mathcal{P}^d$*

$$H(p \wedge q) \leq H(p) + H(q) \tag{1.37}$$

1.4.3 Entropic distance on the lattice

To conclude this chapter, we will take a look at a definition for an entropic distance on the lattice, which allows us to quantify how far away two probability vectors are from each other on the lattice.

Definition 13. For all $p, q \in \mathcal{P}^d$, we define

$$d(p,q) = H(p) + H(q) - 2H(p \vee q). \tag{1.38}$$

Theorem 5 (Entropic distance [2])**.** *d is a distance on $\mathcal{P}^d$, meaning that it satisfies the following properties:*

- *Symmetry: $d(p,q) = d(q,p)$,*
- *Positivity: $d(p,q) \geq 0$ with equality iff $p = q$,*
- *Triangular inequality: $d(p,q) \leq d(p,t) + d(t,q)$.*

Chapter 2

Quantum Information

2.1 Classical information theory

2.2 Mathematical preliminaries

2.2.1 Density matrices

2.2.2 Schmidt decomposition

2.3 Theory of entanglement

2.3.1 Majorization separability criterion

2.3.2 Distilability criterion

2.4 Entanglement transformations

2.4.1 Quantum Resource Theories

2.4.2 Nielsen's protocol

2.4.3 Vidal's protocol

Part II

Results

Chapter 3

Entropic Boost

Chapter 4

Resource Theory of Incomparability

Chapter 5

Volume of Entaglement Cones

Chapter 6

Conclusion

Bibliography

[1] F. Cicalese and U. Vaccaro. "Supermodularity and subadditivity properties of the entropy on the majorization lattice". In: *IEEE Transactions on Information Theory* 48.4 (2002), pp. 933–938.

[2] Ferdinando Cicalese, Luisa Gargano, and Ugo Vaccaro. "Information theoretic measures of distances and their econometric applications". In: *2013 IEEE International Symposium on Information Theory*. ISSN: 2157-8117. 2013, pp. 409–413.

[3] Sándor Dominich. "Elements of Lattice Theory". en. In: *The Modern Algebra of Information Retrieval*. Springer, Berlin, Heidelberg, 2008, pp. 45–64.

[4] A. de Oliveira Junior et al. "Geometric structure of thermal cones". In: *Physical Review E* 106.6 (2022). Publisher: American Physical Society, p. 064109.

[5] Kamil Korzekwa. "Structure of the thermodynamic arrow of time in classical and quantum theories". In: *Physical Review A* 95.5 (2017). Publisher: American Physical Society, p. 052318.

[6] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: Theory of Majorization and Its Applications*. Springer Series in Statistics. New York, NY: Springer, 2011.

Appendices

A Appendix 1

B Appendix 2