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

Information Theory

Chapter 2

Majorization Theory

2.1 The majorization relation

2.1.1 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 thesis.

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

$$S_j^\downarrow(p) := \sum_{i=1}^{j} p_i^\downarrow \tag{2.1}$$

where $p_i^\downarrow$ is the ith 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*

$$\begin{cases} S_j^\downarrow(p) \geq S_j^\downarrow(q) & \forall j = 1, ..., d-1 \\ S_d^\downarrow(p) = S_d^\downarrow(q) \end{cases} \tag{2.2}$$

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 use increasing cumulative sums $S_i^\uparrow(p)$ instead and get an equivalent description of majorization. The only effect would be reversed inequalities in (2.2).

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 probability vector $p \in \mathbb{R}^d$, we have*

$$(\frac{1}{d}, ..., \frac{1}{d}) \prec p \prec (1, 0, ..., 0) \tag{2.3}$$

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.*

2.1.2 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*

$$\begin{cases} D_{i,j} \geq 0 & \forall i, j = 1, ..., d \\ \sum_{i=1}^{d} D_{i,j} = 1 & \forall j = 1, ..., d \\ \sum_{j=1}^{d} D_{i,j} = 1 & \forall i = 1, ..., d \end{cases} \tag{2.4}$$

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*

$$D = \sum_j p_j P_j \tag{2.5}$$

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 $p, q \in \mathbb{R}^d$ be probability vectors of size d. Then, p majorizes q if and only if there exists a d-dimensional bistochastic matrix D such that*

$$q = Dp \tag{2.6}$$

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 series of inequalities (2.2). Let's take the same example as above, $p = (0.9, 0.1)$ and $q = (0.6, 0.4)$. Then, $p \succ q$ because there exists a bistochastic matrix D such that $q = Dp$. In this case, $D = \begin{pmatrix} 5/8 & 3/8 \\ 3/8 & 5/8 \end{pmatrix}$.

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.*

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.

2.1.3 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 explicitly 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.

A few examples of Lorenz curves are given in figure X. FAIRE DES EXEMPLES

2.2 Schur-convex and Schur-concave functions

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.

Definition 4 (Schur-convex function) *A function $f \in C(\mathbb{R}^d \to \mathbb{R})$ is Schur-convex if and only if*

$$p \prec q \implies f(p) \leq f(q) \tag{2.7}$$

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

2.2.1   Shannon entropy

2.3   The majorization lattice

2.3.1   The lattice structure

2.3.2   The meet

2.3.3   The join

2.3.4   Properties of the Shannon entropy on the lattice

Chapter 3

Quantum Information

Part II

Results

Chapter 4

Entropic boost

Chapter 5

Volume of entaglement cones

Chapter 6

Conclusion

Appendices

A   Appendix 1

B   Appendix 2