



ECOLE
POLYTECHNIQUE
DE BRUXELLES

ULB

UNIVERSITÉ LIBRE DE BRUXELLES, UNIVERSITÉ D'EUROPE

Majorization relations in quantum information theory

Mémoire présenté en vue de l'obtention du diplôme
d'Ingénieur Civil physicien à finalité Physique appliquée

Michael G. Jabbour

Directeur
Professeur Nicolas J. Cerf

Co-Promoteur
Evgueni Karpov

Service
Service QuIC

Année académique
2012–2013

Abstract

The theory of majorization is a powerful mathematical tool which naturally arises in the quantum theory as a consequence of its fundamental connection with unitarity. In this work, we look for majorization relations based on entropic inequalities. We begin by introducing the theory of majorization. We then give some notions of quantum mechanics, like the density matrix and some algebra, before presenting the concept of entropy which is closely related to the theory of majorization. We define in the process the Shannon, von Neumann and Rényi entropies, along with some of their properties and applications. We then define quantum entanglement and give some criteria for separability, which turn out to be applications of the concepts of majorization and entropy in the quantum theory. We present some new majorization relations similar to already existing entropic inequalities, and exhibit their advantages and limitations. The rest of the report is dedicated to the study of the field of Gaussian quantum information. We present the theory, before introducing other majorization relations, which turn out to be useful because they allow one to know if it is possible to transform one quantum state into another when there exists majorization relations between them.

Key Words: Majorization, Entropy, Rényi entropy, Entanglement, Separability, Catalysis, Bell diagonal states, Gaussian states.

Résumé

La théorie de la *majorization* est un puissant outil mathématique qui apparaît naturellement dans la théorie quantique comme une conséquence de sa connexion avec l'unitarité. Dans ce rapport, nous recherchons des relations de *majorization* basées sur des inégalités entropiques. Nous commençons par introduire la théorie de la *majorization*. Nous donnons ensuite quelques notions de mécanique quantique, comme la matrice densité et de l'algèbre, avant de présenter le concept d'entropie qui est étroitement lié à la théorie de la *majorization*. Nous définissons dans le processus les entropies de Shannon, von Neumann et Rényi, avec quelques-unes de leur propriétés et applications. Nous définissons ensuite l'intrication quantique et donnons quelques critères de séparabilité, qui s'avèrent être des applications des concepts de *majorization* et entropie dans la théorie quantique. Nous présentons quelques nouvelles relations de *majorization* similaires à des inégalités entropiques déjà existantes, et exposons leurs avantages et inconvénients. Le reste du rapport est dédié à l'étude du domaine de l'information quantique Gaussienne. Nous présentons la théorie, avant d'introduire d'autres relations de *majorization*, qui s'avèrent être utiles car elles permettent de savoir s'il est possible de transformer un état quantique en un autre lorsqu'il existe des relations de *majorization* entre eux.

Mots clés: *Majorization*, Entropie, Entropy de Rényi, Intrication, Séparabilité, Catalyse, États Bell diagonaux, États Gaussiens.

Remerciements

Je souhaiterais adresser ici mes remerciements aux personnes qui m'ont apporté leur aide et qui ont ainsi contribué à l'élaboration de ce mémoire.

Je tiens à remercier sincèrement Monsieur Cerf, qui, en tant que Directeur de mémoire, m'a proposé ce sujet très intéressant, s'est toujours montré très disponible tout au long de la réalisation de ce mémoire, et sans qui ce modeste travail n'aurait pas vu le jour.

Je remercie également Monsieur Karpov pour le temps qu'il a consacré à répondre à mes nombreuses questions, ainsi que pour son précieux conseil durant la concrétisation de ce travail de recherche.

Je n'oublie pas de remercier mes amis, Benjamin, Jonathan, Livio, Mehdi, Nicolas, et beaucoup d'autres, sans qui ces années universitaires n'auraient pas été aussi joyeuses. Je n'oublie pas non plus les discussions intenses et fascinantes à propos de mécanique quantique et sujets variés de physique entretenues avec Jonathan et Mehdi.

Finalement, je remercie aussi tous mes professeurs sans qui je n'aurais jamais accumulé toutes ses connaissances, les membres du jury de ce mémoire pour l'intérêt qu'ils auront porté à ma recherche, ainsi que toutes les personnes qui ont participé de près ou de loin à la réalisation de ce travail.

Contents

Introduction	1
1 Theory of majorization	2
2 Some fundamental notions of quantum mechanics	4
2.1 Density matrix	4
2.2 Some linear algebra	6
2.2.1 Tensor product	6
2.2.2 Transpose and partial transpose	7
2.2.3 Trace and partial trace	8
2.2.4 The Schmidt decomposition	9
2.3 Majorization in quantum mechanics	9
3 Entropy	11
3.1 Shannon entropy	11
3.1.1 Definition	11
3.1.2 Conditional entropy and mutual information	11
3.1.3 Properties	13
3.1.4 Connection with the theory of majorization	13
3.2 von Neumann entropy	14
3.2.1 Definition	14
3.2.2 Conditional entropy and mutual information	14
3.2.3 Properties	15
3.2.4 Connection with the theory of majorization	15
3.3 Rényi entropy	16
4 Quantum Entanglement	17
4.1 Quantum Teleportation	17
4.2 Some important entangled states : The Bell states	19
4.3 Separable states	20
4.4 Characterization of entanglement and separability with majorization	21
5 Separability criteria	22
5.1 Positive partial transpose	22
5.1.1 Arbitrary dimension	22
5.1.2 2×2 and 2×3 states	23
5.2 Conditional entropy criteria	24
5.2.1 Comparison between entropy of a system and entropies of its sub- systems	24
5.2.2 Positivity of the conditional entropy	24
5.2.3 Conditional amplitude operator	25
5.3 Reduction criterion for separability	25
5.4 Separability criterion based on the theory of majorization	25

6	Going beyond entropic inequalities using the theory of majorization	26
6.1	Nielsen and Kempe's result	27
6.2	Positivity of the joint entropy of a separable state in terms of majorization	27
6.3	First adaptation of the subadditivity relation to the theory of majorization	28
6.4	Second adaptation of the subadditivity relation to the theory of majorization	30
6.4.1	Relation and examples	30
6.4.2	Catalysis	36
6.4.3	Bell diagonal states	39
6.5	Conclusion	40
7	Gaussian Quantum Information	40
7.1	Continuous Variables model	41
7.1.1	The quantum harmonic oscillator model	41
7.1.2	The Gaussian state	42
7.2	Thermal states	47
7.2.1	Definition and properties	47
7.2.2	Thermal decomposition of Gaussian state	49
7.3	Majorization in a beam splitter	53
8	Investigating the possible interconversion between Gaussian states	54
8.1	One-mode Gaussian states	55
8.2	N -mode Gaussian states	58
8.3	Conclusion	63
	Conclusion	64
	References	65
A	MATLAB codes	67
A.1	Matlab code <code>majorization_gamma.m</code>	67
A.2	Matlab code <code>xab_majorization.m</code>	67
A.3	Matlab code <code>subadditivity_Renyi.m</code>	69
A.4	Matlab code <code>gamma_catalysis.m</code>	69
A.5	Matlab code <code>alpha_entropy.m</code>	70
A.6	Matlab code <code>Major.m</code>	71
A.7	Matlab code <code>tensor.m</code>	72
A.8	Matlab code <code>TrX.m</code>	72
A.9	Matlab code <code>TrX23.m</code>	74

Introduction

Quantum information science, the study of the information processing tasks that can be accomplished using quantum mechanics systems, has been experiencing a major growth during the last thirty years. Coupling the subject of quantum mechanics with the theory of information made it possible to uncover various fascinating applications, going from quantum cryptography to quantum teleportation. These are the results of some curiosities specific to the quantum world, like the intriguing phenomenon of entanglement, which is surprisingly contrary to our intuition.

In order to understand entanglement better, be able to detect it and quantify it, quantum information scientists have been looking for separability criteria as well as measures of entanglement, a separable state being one which does not exhibit entanglement at all. Thereby, a number of mathematical relations have been proposed by researchers in order to accomplish this task. One of them, proposed by Nielsen and Kempe, is based on an elegant and powerful theory, majorization. The concept of majorization, which is closely related to the concept of disorder, allows to compare vectors based on their "randomness". Naturally, this characteristic of the theory of majorization connects it to the well known measures of disorder provided by entropies. As a consequence, Nielsen and Kempe's majorization criterion for separability is based on an entropic inequality for separability, but is stronger than the latter.

This close relation between the criterion in terms of majorization relation and entropic inequality is at the root of the present work. Taking inspiration from Nielsen and Kempe's result, the aim of this project is to look for majorization relations based on well known entropic inequalities. When it is possible to find such relations, we expect them, like Nielsen and Kempe's criterion, to be stronger than the entropic inequalities on which it is based. This last characteristic makes it interesting and motivating to look for majorization relations. Furthermore, the application of the theory of majorization to the field of quantum information is recent, and still has much to offer.

This report is organized as follows. In the first section, we will introduce the theory of majorization in terms of "classical" probabilities. In the second section, we will exhibit some notions of quantum mechanics which will be of importance for the application of the mathematical subject studied in this work to the quantum theory. The majorization being closely related to the notion of disorder, we will present the concept of entropy in the third section. Quantum entanglement will be introduced in the fourth section, along with some well known separability criteria in section 5. Section 6 will contain the first half of the results investigated in this work. These result will be based on already existing entropic relations, as we already explained. The second half of the results presented in this report will be given in section 8, and will this time be related to the more particular case of Gaussian quantum information, which we will have already introduced in section 7. Notice that apart from section 6 and 8 which contain the results investigated in this project, every section will be related to the theory of majorization through a short subsection exhibiting the connection between the mathematical theory and the section, or simply giving some practical examples.

1 Theory of majorization

We begin this report by introducing the main mathematical theory explored in the present work, the theory of majorization. It is a powerful and elegant mathematical tool which can be applied to a wide variety of problems in quantum mechanics [17], as we are going to see. However, before entering the quantum world with the tool of majorization in our hand, we will present the theory in terms of more "classical objects", such as simple probabilities distributions. Later, after giving some definitions specific to quantum mechanics, we shall clearly exhibit in section 2.3 the close relation between the mathematical tool and the quantum theory.

The theory of majorization is closely related to the notions of "randomness" and "disorder". It indeed allows us to compare two probability distributions, in order for us to know which one of the two is more random [14]. Let us now give the most general definition of majorization.

Definition 1. *Given two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$, we say that \mathbf{a} majorizes (or dominates) \mathbf{b} , written as $\mathbf{a} \succ \mathbf{b}$ iff*

$$\left\{ \begin{array}{l} \sum_{i=1}^k a_i^\downarrow \geq \sum_{i=1}^k b_i^\downarrow \quad \text{for } k = 1, \dots, d-1, \\ \sum_{i=1}^d a_i = \sum_{i=1}^d b_i \end{array} \right. \quad (1.1)$$

where a_i^\downarrow and b_i^\downarrow are the elements of \mathbf{a} and \mathbf{b} , respectively, sorted in decreasing order.

Note that the original order of the elements in the vectors does not play any role in terms of majorization. Furthermore, the second equation of system (1.1) is automatically satisfied when the elements of vectors \mathbf{a} and \mathbf{b} are elements of probability distributions.

An important difference between the notion of inequality and the theory of majorization is the fact that the latter only provides a partial order, since the fact that \mathbf{a} is not majorized by \mathbf{b} does not automatically imply that \mathbf{b} is majorized by \mathbf{a} , i.e.

$$\mathbf{a} \not\succeq \mathbf{b} \quad \not\Rightarrow \quad \mathbf{b} \succ \mathbf{a}. \quad (1.2)$$

If both $\mathbf{a} \not\succeq \mathbf{b}$ and $\mathbf{a} \not\succeq \mathbf{b}$ hold, the two vectors \mathbf{a} and \mathbf{b} are *incomparable*.

When $\mathbf{a} \succ \mathbf{b}$, it can be said that \mathbf{b} is more disordered than \mathbf{a} . Definition 1 does not really exhibit the relation between majorization and disorder. In order to see this connection, let us introduce a useful property, which is as follows.

Property 1. *Given the two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$, $\mathbf{a} \succ \mathbf{b}$ if and only if*

$$\mathbf{b} = D\mathbf{a} \quad (1.3)$$

for some doubly stochastic matrix D .

Note that a doubly stochastic matrix (also called bistochastic), is a square matrix $D = (d_{ij})$ of nonnegative real numbers, each of whose rows and columns sum to 1, i.e.,

$$\sum_i d_{ij} = \sum_j d_{ij} = 1. \quad (1.4)$$

Thus, a doubly stochastic matrix is both left stochastic and right stochastic. The set of bistochastic matrices of a given dimensions is convex. Now, the extreme points of this convex set are given by the permutation matrices π_n [8]. Therefore, any doubly stochastic matrix can be expressed as a convex combination of permutation matrices. Using property 1, we see that $\mathbf{a} \succ \mathbf{b}$ if and only if there exists a set of d -dimensional permutation matrices π_n and probability distribution $\{t_n\}$ such that

$$\mathbf{b} = \sum_n t_n \pi_n \mathbf{a}. \quad (1.5)$$

Consequently, we see that $\mathbf{a} \succ \mathbf{b}$ if and only if \mathbf{b} can be obtained by randomly permuting the components of vector \mathbf{a} and afterwards taking the average over all permutations [8]. This intuitively shows that \mathbf{b} is indeed more disordered than \mathbf{a} . As an example, let us introduce the following proposition.

Proposition 1. *Suppose the vector \mathbf{s} is any probability distribution on d outcomes, which means that its components are non-negative and sum to one. We then have [20]*

$$\left(\frac{1}{d}, \dots, \frac{1}{d}\right) \prec \mathbf{s}. \quad (1.6)$$

This is indeed true, since the uniform distribution $(1/d, \dots, 1/d)$ can be obtained by averaging over permutations of \mathbf{s} . This agrees with our intuition that the uniform distribution on d elements is at least as disordered as any other probability distribution over d elements [20].

Since Majorization theory is related to the notion of disorder, it should also be connected with measures of this disorder, such as entropies, which we will be defining later in this report. In order to show this connection, let us introduce another property of majorization we will often be using in this work. This property namely relates all the Schur-convex functions (like entropies, as we will state later) to the theory of majorization. Before giving this property and the definition of a Schur-convex function, we relate majorization to the simpler concept of convex functions. The definition of a convex function is as follows.

Definition 2. *Let f be a function from \mathbb{R}^n to $[-\infty, +\infty]$. Then f is convex if and only if [23]*

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y), \quad 0 < \lambda < 1, \quad (1.7)$$

whenever $f(x) < \alpha$ and $f(y) < \beta$.

A property which relates convex functions to the concept of majorization is the following [14].

Property 2. *The inequality*

$$\sum_i g(x_i) \leq \sum_i g(y_i) \quad (1.8)$$

holds for all continuous convex functions $g : \mathbb{R} \rightarrow \mathbb{R}$ if and only if $x \prec y$.

We see that for some majorization relation to hold, some similar inequalities should hold for all the convex functions. However, the entropies are not related to the property of convexity. They are in fact related to the concept of *Schur-convexity*. Let us give the definition of a Schur-convex function [14].

Definition 3. A real-valued function ϕ defined on a set $\mathcal{A} \subset \mathbb{R}^n$ is said to be Schur-convex on \mathcal{A} if

$$x \prec y \text{ on } \mathcal{A} \quad \rightarrow \quad \phi(x) \leq \phi(y). \quad (1.9)$$

We see that, unlike the case of convex functions, the definition of Schur-convex functions is directly related to the concept of majorization. We are now able to state the property which connects majorization to the concept of Schur-convex functions. It is as follows [14].

Property 3. $x \prec y$ if and only if $\phi(x) \leq \phi(y)$ for all Schur-convex functions ϕ .

As we will be saying when introducing the concept of entropy later in this work, all the Rényi's entropies, which are a measure of disorder, are in fact Schur-concave, a concave function being the negative of a convex function. Therefore, for some majorization relation to be true, some similar inequality should be verified for at least all the Rényi's entropies. We conclude that the concept of entropy, which we will be introducing in section 3, is closely related to the theory of majorization.

As we already stated, majorization is a powerful tool when studying quantum systems. We are going to show this, but we need to introduce some general concepts of quantum mechanics before. This is the aim of the next section.

2 Some fundamental notions of quantum mechanics

We are now going to explain some fundamental notions in quantum mechanics, like the density matrix, which is used to describe any quantum system, as well as some mathematical operations, which are used to manipulate quantum states.

2.1 Density matrix

Here we introduce one of the very basic concepts of quantum mechanics. In the quantum mechanics formalism, a system can be described by a mathematical function, called wave function, which completely describes the quantum state of the system. A wave function ψ can be associated with a vector, which is represented by the notation

$$|\psi\rangle \quad (2.1)$$

in "bra-ket" notations. The vector $|\psi\rangle$ is called pure state, since the system is characterized by one unique vector $|\psi\rangle$. However, in the general case, one does not have access to the entire information about the system. Consequently, one does not know in which pure state the quantum system is. That is why it is usually described as a statistical ensemble of several quantum states ($|\psi_1\rangle, |\psi_2\rangle, \dots$), with the use of a certain probability distribution (p_1, p_2, \dots). In this case, the system is described by a density matrix, usually written as ρ (we omit the time dependence of the state here), which is an operator that completely characterises any quantum state [7], and is a matrix associated to the wave function of the state. The density matrix describing the state is defined by

$$\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|, \quad (2.2)$$

with the conditions (characteristic of a probability distribution)

$$\begin{cases} 0 \leq p_i \leq 1 & \forall i \\ \sum_i p_i = 1. \end{cases} \quad (2.3)$$

The density matrix described by equation (2.2) is called a mixed state in general, and generalises the case of the pure state. Indeed, when the state is pure, the density matrix reduces to

$$\rho = |\psi\rangle \langle\psi| \quad (2.4)$$

where ψ is the wave function of the pure state. In this case, the vector $|\psi\rangle$ is sufficient to characterize the quantum state.

Lets us exhibits some important properties of the density matrix that will be useful in the following:

1. The density matrix is a hermitian operator : $\rho^\dagger = \rho$.
2. The density matrix is normalized : $\text{Tr } \rho = 1$ (the definition of the operator Tr will be given later, see section 2.2.3).
3. The density matrix is positive : $\rho \geq 0$.

The last property tells us in particular that all the eigenvalues of ρ are positive. There is also one more property which is different whether we are in the case of pure states or mixed states. In the case of pure states, the density matrix is a projector

$$\rho^2 = \rho. \quad (2.5)$$

The last condition leads to

$$\text{Tr } \rho^2 = 1. \quad (2.6)$$

This property is not verified in the case of a mixed state, for which

$$\text{Tr } \rho^2 < 1. \quad (2.7)$$

Note that the last condition often provides the easiest way to verify if a state is pure or if it is in a statistical mixture. As we shall see, it is also connected to the Rényi entropy of order 2, which we already mentioned before. We will define the Rényi entropy of order α later, as we already said, but let us give the particular definition of the Rényi entropy of order 2 in order to see its relation with condition (2.7). For a discrete random variable X with probabilities p_i , it is defined by

$$I_2(X) = -\log \left(\sum_{i=1}^n p_i^2 \right). \quad (2.8)$$

In order to see the connection with relation (2.7), take a state ρ , which can always be written in its diagonal decomposition

$$\rho = \sum_{i=1}^n \lambda_i |i\rangle \langle i|, \quad (2.9)$$

where the λ_i are its eigenvalues and the $\{|i\rangle\}_{i=1,\dots,n}$ are orthogonal states. The trace of ρ^2 is given by (again, the definition of the trace operation will be given later)

$$\text{Tr}\rho^2 = \sum_{i=1}^n \lambda_i^2. \quad (2.10)$$

We therefore see that condition (2.7) is equivalent to

$$I_2(X) > 0, \quad (2.11)$$

where the discrete random variable X has probabilities λ_i . We already mentioned in the section where we introduced the concept of majorization that the entropy is a measure of disorder. Indeed, we see that if $I_2(X) = 0$, according to condition (2.7), ρ is a pure state. In this case, it is not given by a statistical ensemble. This is consistent with the fact that its measure of disorder is equal to 0. If ρ was a mixed state, its entropy of order 2 would be greater than 0, which is consistent with the fact that its density matrix would be given by a statistical ensemble of pure states.

2.2 Some linear algebra

Let us now introduce some linear algebra that will be useful later. Setting forth some mathematical definitions will allow us to choose them once and for all.

2.2.1 Tensor product

Vectors Suppose \mathbf{u} is a vector in \mathcal{H}_1 and \mathbf{v} is a vector in \mathcal{H}_2 , where \mathcal{H}_1 and \mathcal{H}_2 are Hilbert spaces of dimensions k and n , respectively. The tensor product \mathbf{w} of \mathbf{u} and \mathbf{v} is a vector in a Hilbert space \mathcal{H} of dimension $k \times n$. If $\mathbf{u} = (u_1, \dots, u_k)$ and $\mathbf{v} = (v_1, \dots, v_n)$, \mathbf{w} is given by (we write \mathbf{w} as a column vector for simplicity)

$$\mathbf{w} = \mathbf{u} \otimes \mathbf{v} = \begin{pmatrix} u_1 \mathbf{v} \\ u_2 \mathbf{v} \\ \vdots \\ u_k \mathbf{v} \end{pmatrix} = \begin{pmatrix} u_1 v_1 \\ \vdots \\ u_1 v_n \\ u_2 v_1 \\ \vdots \\ u_2 v_n \\ \vdots \\ u_k v_n \end{pmatrix}$$

Matrices Another useful tensor product is the one between matrices, since we will be performing tensor product between density operators. Suppose $\mathcal{B}_1 : \mathcal{H}_1 \rightarrow \mathcal{H}_1$ the space of the linear operators acting on the Hilbert space \mathcal{H}_1 . Suppose, similarly, $\mathcal{B}_2 : \mathcal{H}_2 \rightarrow \mathcal{H}_2$ the space of the operators acting on the Hilbert space \mathcal{H}_2 . Given $\mathbf{A} \in \mathcal{B}_1$ and $\mathbf{B} \in \mathcal{B}_2$ two matrices, defined by

$$\mathbf{A} = \begin{pmatrix} A_{11} & \cdots & A_{1l} \\ \vdots & \ddots & \vdots \\ A_{k1} & \cdots & A_{kl} \end{pmatrix} \quad (2.12)$$

and

$$\mathbf{B} = \begin{pmatrix} B_{11} & \cdots & B_{1q} \\ \vdots & \ddots & \vdots \\ B_{p1} & \cdots & B_{pq} \end{pmatrix}, \quad (2.13)$$

the tensor product of \mathbf{A} and \mathbf{B} is a matrix \mathbf{C} in the space $\mathcal{B} : \mathcal{H}_1 \otimes \mathcal{H}_2 \rightarrow \mathcal{H}_1 \otimes \mathcal{H}_2$ defined by

$$\mathbf{C} = \mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} A_{11}\mathbf{B} & \cdots & A_{1l}\mathbf{B} \\ \vdots & \ddots & \vdots \\ A_{k1}\mathbf{B} & \cdots & A_{kl}\mathbf{B} \end{pmatrix} \quad (2.14)$$

which results in

$$\mathbf{C} = \begin{pmatrix} A_{11}B_{11} & \cdots & A_{11}B_{1q} & A_{12}B_{11} & \cdots & A_{12}B_{1q} & \cdots & A_{1l}B_{11} & \cdots & A_{1l}B_{1q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{11}B_{p1} & \cdots & A_{11}B_{pq} & A_{12}B_{p1} & \cdots & A_{12}B_{pq} & \cdots & A_{1l}B_{p1} & \cdots & A_{1l}B_{pq} \\ A_{21}B_{11} & \cdots & A_{21}B_{1q} & A_{22}B_{11} & \cdots & A_{22}B_{1q} & \cdots & A_{2l}B_{11} & \cdots & A_{2l}B_{1q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{21}B_{p1} & \cdots & A_{21}B_{pq} & A_{22}B_{p1} & \cdots & A_{22}B_{pq} & \cdots & A_{2l}B_{p1} & \cdots & A_{2l}B_{pq} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{k1}B_{11} & \cdots & A_{k1}B_{1q} & A_{k2}B_{11} & \cdots & A_{k2}B_{1q} & \cdots & A_{kl}B_{11} & \cdots & A_{kl}B_{1q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{k1}B_{p1} & \cdots & A_{k1}B_{pq} & A_{k2}B_{p1} & \cdots & A_{k2}B_{pq} & \cdots & A_{kl}B_{p1} & \cdots & A_{kl}B_{pq} \end{pmatrix}. \quad (2.15)$$

Let us state a useful property of tensor products we will be using later.

Property 4. Suppose \mathbf{u} and \mathbf{v} are two vectors, and \mathbf{A} and \mathbf{B} are two matrices. Then

$$(\mathbf{A}\mathbf{u}) \otimes (\mathbf{B}\mathbf{v}) = (\mathbf{A} \otimes \mathbf{B})(\mathbf{u} \otimes \mathbf{v}). \quad (2.16)$$

The tensor product is a central operation in the field of quantum information, and more generally in the field of quantum mechanics, because it allows us to combine different systems. For example, if system A is characterized by density matrix ρ^A and system B is characterized by density matrix ρ^B , then the total system resulting from the combination of systems A and B will be characterized by the density operator

$$\rho^{AB} = \rho^A \otimes \rho^B. \quad (2.17)$$

The order with which we perform the tensor product is important because it will affect the way an operator only acting on one of the spaces of the two subsystems A and B , will act on the space of the total system AB .

2.2.2 Transpose and partial transpose

Transposition One basic operation in the field of quantum mechanics is the transpose T of a matrix in a Hilbert space \mathcal{H} . Using the basis $|n\rangle$ of \mathcal{H} , the transposition is defined as the mapping which transforms $\rho = \sum_{n,m} \rho_{nm} |n\rangle \langle m|$ into [2]

$$\mathsf{T}(\rho) = \rho^{\mathsf{T}} \equiv \sum_{n,m} \rho_{nm} |m\rangle \langle n| \quad (2.18)$$

The elements of the transpose ρ^{T} of ρ are

$$[\rho^{\mathsf{T}}]_{nm} = \rho_{mn}. \quad (2.19)$$

Partial Transposition Having shown earlier how to combine two systems, we can introduce another important operation of linear algebra, known as the partial transpose. It consists namely in performing the transpose operation on only one of the two subsystems (here we take the case of only two subsystems, but the total system could be composed of more). Suppose subsystem A is related to Hilbert space \mathcal{H}^A with basis $|n^A\rangle$ and subsystem B is related to Hilbert space \mathcal{H}^B with basis $|\mu^B\rangle$, then the partial transposition in \mathcal{H}^A , $(\mathbb{T}^A \otimes \mathcal{I}^B)\rho^{AB} = (\rho^{AB})^{\mathbb{T}_A}$ is given by the mapping of ρ^{AB} : [2]

$$\rho^{AB} \leftrightarrow \rho_{m\mu, n\nu} = \langle m^A, \mu^B | \rho^{AB} | n^A, \nu^B \rangle \quad (2.20)$$

onto

$$(\rho^{AB})^{\mathbb{T}_A} \leftrightarrow \rho_{m\mu, n\nu}^{\mathbb{T}_A} = \rho_{n\mu, m\nu}. \quad (2.21)$$

2.2.3 Trace and partial trace

Trace In section 2.1, one of the properties of the density matrix was based on the trace (Tr), which acts on an operator (the density matrix in the case of section 2.1). Here we define it in order to introduce another object, the partial trace. The trace is a complex-valued function of a linear operator. Its action on an operator A is defined by [2]

$$\text{Tr}[A] \equiv \sum_i \langle i | A | i \rangle, \quad (2.22)$$

where $|i\rangle$ is any orthonormal basis.

Partial trace The partial trace acts on only one of the subspaces of the system. For example, the partial trace over the space \mathcal{H}^A of an operator Z^{AB} is defined by [2]

$$\text{Tr}_A[Z^{AB}] \equiv \sum_n \langle n^A | Z^{AB} | n^A \rangle, \quad (2.23)$$

where $|n^A\rangle$ is any orthonormal basis in Hilbert space \mathcal{H}^A . The order in which one applies the partial traces over different subsystems is irrelevant. Indeed,

$$\begin{aligned} \text{Tr}_B[\text{Tr}_A[Z^{AB}]] &= \sum_n \langle n^B | (\text{Tr}_A[Z^{AB}]) | n^B \rangle \\ &= \sum_n \langle n^B | \left(\sum_m \langle m^A | Z^{AB} | m^A \rangle \right) | n^B \rangle \\ &= \sum_m \langle m^A | \left(\sum_n \langle n^B | Z^{AB} | n^B \rangle \right) | m^A \rangle \\ &= \sum_m \langle m^A | (\text{Tr}_B[Z^{AB}]) | m^A \rangle \\ &= \text{Tr}_A[\text{Tr}_B[Z^{AB}]] \end{aligned} \quad (2.24)$$

Furthermore, if one applies all the partial traces relative to all the subsystems, it is equivalent to applying the trace. Therefore,

$$\text{Tr}[Z^{AB}] = \text{Tr}_A[\text{tr}_B[Z^{AB}]] = \text{Tr}_B[\text{tr}_A[Z^{AB}]]. \quad (2.25)$$

The importance of the partial trace reside in the fact that it allows us to recover one of the subsystems starting from the total system. In other words, if one wants the density matrix ρ^A of subsystem A when one has the density matrix ρ^{AB} of the total system, one applies the partial trace over system B , namely

$$\rho_A = \text{Tr}_B[\rho^{AB}]. \quad (2.26)$$

2.2.4 The Schmidt decomposition

The density matrix, the partial trace and the partial transposition are powerful tools for the study of composite quantum systems, which are at the root of the field of quantum information. There are additional tools of great value for the study of quantum systems. One of them is the Schmidt decomposition.

Suppose, like previously, that we have a composite system AB in the Hilbert space \mathcal{H}^{AB} . This system is the result of the composition of the two subsystems A and B , which are in the Hilbert spaces \mathcal{H}^A and \mathcal{H}^B , respectively, such that

$$\mathcal{H}^{AB} = \mathcal{H}^A \otimes \mathcal{H}^B. \quad (2.27)$$

Suppose furthermore that the dimensions of the subsystems are given by $\dim\mathcal{H}^A = a$ and $\dim\mathcal{H}^B = b$. Let $|\psi^{AB}\rangle$ be a pure state in the space AB . The density matrix corresponding to this pure state is given by

$$\rho^{AB} = |\psi^{AB}\rangle\langle\psi^{AB}|, \quad (2.28)$$

and, like previously, the reduced density matrices of the subsystems A and B are given by $\rho^A = \text{tr}_B[\rho^{AB}]$ and $\rho^B = \text{tr}_A[\rho^{AB}]$, respectively. Then the following results are true: [2]

1. The vector $|\psi^{AB}\rangle$ can be written in the form of a Schmidt decomposition

$$|\psi^{AB}\rangle = \sum_{n=1}^k \sqrt{p_n} |u_n^A\rangle |v_n^B\rangle \quad \text{with } p_n > 0 \forall n \quad (2.29)$$

with $k \leq \min(a, b)$, where $|u_n^A\rangle$ and $|v_n^B\rangle$ are the orthonormalised eigenvectors of $\rho^A \in \mathcal{H}^A$ and $\rho^B \in \mathcal{H}^B$, respectively.

2. ρ^A and ρ^B have the same positive eigenvalues p_1, \dots, p_k .

An interesting consequence of this result is the fact that for a bipartite pure state, all the measures of entropies on subsystem A are respectively equal to the same measures of entropy on subsystem B . Indeed, as we are going to see in the next section, the measure of entropy of a state only depends on its eigenvalues. Since the density matrices of the two subsystems A and B of a pure state have the same eigenvalues, one can conclude that they also have the same measures of entropy.

2.3 Majorization in quantum mechanics

What connections are there between majorization and quantum mechanics? Having introduced the notions of sections 2.1 and 2.2, we are now able to answer this question. In

the classical case of section 1, we used the majorization theory in order to analyze probability distributions. The different values of the probability distributions were arranged in vectors, which were compared using majorization relations. In the field of quantum mechanics, it is interesting to compare the eigenvalues of density matrices. The eigenvalues form probability distributions, since they are all non-negative and sum to 1 when the density matrix represents a state.

The main reason for the close relation between majorization and quantum mechanics may be appreciated by inspection of two elegant results [17] which we now state. The first result is called *Horn's lemma* and is as follows.

Proposition 2. *Horn's lemma:* *For two vectors \mathbf{r} and \mathbf{s} , $\mathbf{r} \prec \mathbf{s}$ if and only if $r_i = \sum_j |u_{ij}|^2 s_j$ for some unitary matrix $\mathbf{u} = (u_{ij})$ of complex numbers.*

The second result is called *Uhlmann's theorem* and is as follows.

Proposition 3. *Uhlmann's theorem:* *For the two vectors $\lambda(R)$ and $\lambda(S)$ composed of the eigenvalues of the two respective Hermitian matrices R and S , the relation $R \prec S$ is true if and only if there exist unitary matrices U_j and a probability distribution $\{p_j\}$ such that*

$$R = \sum_j p_j U_j S U_j^\dagger. \quad (2.30)$$

Unitarity is of fundamental importance in the field of quantum mechanics. As we will explain later, unitary transformations are an important type of quantum operations, since they are in fact reversible. This ensures that relations of the type featuring in Horn's lemma and Uhlmann's theorem arise frequently [17], and it is this which accounts for many of the applications of majorization to quantum mechanics.

Let us exhibit an interesting application of the theory of majorization in the field of quantum information. This application is related to one kind of transformations that can be performed on quantum systems. These transformations are the processing of local operations and classical communication (LOCC). Suppose we have two pure bipartite states $|\psi\rangle$ and $|\phi\rangle$ of Alice and Bob's system. The theory of majorization allows us to know if it is possible to perform LOCC in order to transform $|\psi\rangle$ into $|\phi\rangle$, by comparing the vectors of eigenvalues of the states corresponding to one of the subsystems. Here, denote the states of Alice's system by

$$\rho_\psi \equiv \text{Tr}_B (|\psi\rangle \langle \psi|) \quad (2.31)$$

and

$$\rho_\phi \equiv \text{Tr}_B (|\phi\rangle \langle \phi|). \quad (2.32)$$

Let λ_ψ and λ_ϕ be the vectors of eigenvalues of ρ_ψ and ρ_ϕ , respectively. Nielsen's theorem about LOCC is the following [16].

Theorem 1. *$|\psi\rangle$ transforms to $|\phi\rangle$ using local operations and classical communication if and only if λ_ψ is majorized by λ_ϕ . More succinctly,*

$$|\psi\rangle \rightarrow |\phi\rangle \quad \text{iff} \quad \lambda_\psi \prec \lambda_\phi \quad (2.33)$$

There are many applications of the theory of majorization in the field of quantum mechanics. Theorem 1 is one of them. Another one is one of the separability criteria we will be introducing later. Furthermore, as we already explained, the goal in this work will be to introduce new majorization relations, which will have more applications in the field of quantum information.

3 Entropy

We will now introduce a key concept in the field of quantum information, the entropy. Entropy measures the amount of uncertainty in the state of a physical system [18]. It is a very powerful tool, since it allowed Shannon to establish a mathematical theory around the concept of information. Indeed, before Shannon's contribution, it was not possible to quantify information. We will see how this concept also applies to the field of quantum mechanics, generating the field of quantum information itself.

3.1 Shannon entropy

3.1.1 Definition

The Shannon entropy was introduced by Shannon himself in the field of classical information. There are two ways of viewing the Shannon entropy. Suppose we have a random variable X , and we learn its value. In one point of view, the Shannon entropy quantifies the amount of information we gain when we learn the value of X (*after* measurement). In another point of view, the Shannon entropy tells us the amount of uncertainty about the variable X before we learn its value (*before* measurement).

Definition 4. *Suppose there is a probability distribution p_1, \dots, p_n associated to the random variable X . The Shannon entropy associated with it is defined by*

$$H(X) \equiv - \sum_x p_x \log p_x, \quad (3.1)$$

where the log function is taken to base 2.

Note that there is no problem with the definition in the case of a zero probability, since

$$\lim_{x \rightarrow 0} x \log x = 0. \quad (3.2)$$

The entropy is usually measured in bits of information, which means that when we learn the value of a random variable X , we acquire $H(X)$ bits of information about X in the process. This also means that there is less uncertainty about the random variable X after the measurement, consistent with the two interpretations of the entropy.

3.1.2 Conditional entropy and mutual information

Suppose we have two random variables X and Y . Now what would be interesting is to relate the information content of X to the information content of Y . This can be done using two powerful concepts of information theory, the conditional entropy and the mutual information. Before introducing these two concept, we set forth a new definition, which is the definition of the joint entropy of a pair of random variables.

Definition 5. *The joint entropy of X and Y is defined by [18]*

$$H(X, Y) \equiv - \sum_{x,y} p(x, y) \log p(x, y). \quad (3.3)$$

The joint entropy measures the total uncertainty about the pair (X, Y) .

Suppose now that we only learn the value of Y . There is still some unknown information about X , but this information content is not given by $H(X)$ in general. Indeed, some information could be common to X and Y . Therefore, in this case, if we know all the information about Y , then we could know part of the information of X . The remaining unknown information about X is called entropy of X conditional on knowing Y .

Definition 6. *The entropy of X conditional on knowing Y is defined as [18]*

$$H(X|Y) \equiv H(X, Y) - H(Y). \quad (3.4)$$

This quantity is smaller than $H(X)$ in general. Again, the conditional entropy $H(X|Y)$ is a measure of the uncertainty on the value of X , given that we know the value of Y .

We talked about the fact that there could be some information shared by the two random variables X and Y . This information is called mutual information content of X and Y .

Definition 7. *The mutual information content of X and Y is defined by [18]*

$$H(X : Y) \equiv H(X) + H(Y) - H(X, Y). \quad (3.5)$$

Indeed, when we compute the value $H(X) + H(Y)$, we take the information content that X and Y have in common twice. In addition to that, we have the remaining information about X , and the remaining information about Y . That is why, in order to compute the mutual information of X and Y , we need to remove this remaining information in addition to the common information content, and this corresponds to the joint entropy $H(X, Y)$. This explains equation (3.5).

In order to understand this further, suppose we're in the case of independent variables X and Y . In this case, we have $p(x, y) = p(x)p(y)$, and the joint entropy reduces to

$$\begin{aligned} H(X, Y) &= - \sum_{x,y} p(x, y) \log p(x, y) \\ &= - \sum_{x,y} p(x)p(y) \log[p(x)p(y)] \\ &= - \sum_{x,y} p(x)p(y) (\log[p(x)] + \log[p(y)]) \\ &= - \sum_y p(y) \sum_x p(x) \log[p(x)] - \sum_x p(x) \sum_y p(y) \log[p(y)] \\ &= - \sum_x p(x) \log[p(x)] - \sum_y p(y) \log[p(y)] \\ &= H(X) + H(Y) \end{aligned} \quad (3.6)$$

since $p(x)$ and $p(y)$ are probability distributions. Since in this case, there is no common information between the two variables, the conditional entropy of X should simply be equal to the entropy of X . This is indeed the case:

$$\begin{aligned} H(X|Y) &= H(X, Y) - H(Y) \\ H(X|Y) &= H(X) + H(Y) - H(Y) \\ H(X|Y) &= H(X) \end{aligned} \quad (3.7)$$

Furthermore, it is obvious that the mutual information should be zero. Again, this is verified since

$$\begin{aligned} H(X : Y) &= H(X) + H(Y) - H(X, Y) \\ H(X : Y) &= H(X) + H(Y) - H(X) - H(Y) \\ H(X : Y) &= 0 \end{aligned} \tag{3.8}$$

Note that an interesting equality relating the conditional entropy to the mutual information is given by

$$H(X : Y) = H(X) - H(X|Y). \tag{3.9}$$

3.1.3 Properties

Here we give some basic properties of the Shannon entropy that might be useful [18].

1. The Shannon entropy $H(X)$ is a concave function of the probability distribution corresponding to its argument X .
2. $H(X, Y) = H(Y, X)$, $H(X : Y) = H(Y : X)$.
3. $H(Y|X) \geq 0$ and thus $H(X : Y) \leq H(Y)$.
4. $H(X) \leq H(X, Y)$.
5. $H(Y|X) \leq H(Y)$ and thus $H(X : Y) \geq 0$.

The last property is in fact known as the subadditivity, and is often written in the equivalent way

$$H(X, Y) \leq H(X) + H(Y) \tag{3.10}$$

with equality iff X and Y are independent variables. This property will be at the root of some majorization relations we will be studying.

3.1.4 Connection with the theory of majorization

Here we show how the Shannon entropy is connected to the theory of majorization. They are both linked to the measure of disorder in a system. However, the theory of majorization usually gives stronger criteria than the entropic inequalities. In order to see this, let us give the following proposition.

Proposition 4. *Take a random variable X associated to probabilities $\{p_i\}$ and a random variable Y associated to probabilities $\{q_i\}$. If the vector $\mathbf{p} \equiv (p_1, p_2, \dots)$ is majorized by the vector $\mathbf{q} \equiv (q_1, q_2, \dots)$, i.e. $\mathbf{p} \prec \mathbf{q}$, then [17]*

$$H(X) \geq H(Y), \tag{3.11}$$

where H is the Shannon entropy.

Proposition 4 is in fact a particular case of property 2. Indeed, the Shannon entropy is a concave function, as stated in its properties. Proposition 4 can be proved very easily using the property of concavity of the Shannon entropy. Suppose $p \prec q$, then according to relation (1.5), there exists a set of permutation matrices π_n and probability distribution $\{t_n\}$ such that

$$\mathbf{p} = \sum_n t_n \pi_n \mathbf{q}. \tag{3.12}$$

From the concavity of the Shannon entropy, it follows that

$$H(X) \geq \sum_n t_n H(Y) \quad (3.13)$$

since permuting the elements of \mathbf{q} does not change its Shannon entropy. Knowing that $\{t_n\}$ is a probability distribution, we have

$$H(X) \geq H(Y). \quad (3.14)$$

We therefore see that some relations of order based on the majorization theory imply some relations of order based on the entropy. However, the opposite is not true. Indeed,

$$H(X) \geq H(Y) \quad (3.15)$$

does not imply $\mathbf{p} \prec \mathbf{q}$ in general. This implies that measures based on the entropy are essentially weaker than the notion of majorization.

3.2 von Neumann entropy

3.2.1 Definition

The Shannon entropy we described in the previous sections measures the uncertainty about some random variables associated with some classical probability distributions. The von Neumann entropy does the same, but in the field of quantum mechanics. Here, the classical probability distributions are replaced by density operators.

Definition 8. *The von Neumann entropy of a quantum state ρ is defined by [18]*

$$S(\rho) \equiv -\text{Tr}(\rho \log \rho). \quad (3.16)$$

If the λ_x are the eigenvalues of ρ , then the von Neumann entropy becomes

$$S(\rho) = -\sum_x \lambda_x \log \lambda_x. \quad (3.17)$$

The last equation is exactly the same as equation (3.1), but with the eigenvalues of the state ρ replacing the classical probabilities.

3.2.2 Conditional entropy and mutual information

By analogy with what we did in section 3.1.2, it is possible to define joint entropies, conditional entropies and mutual information for composite quantum systems. Suppose we have a composite system AB in the Hilbert space \mathcal{H}^{AB} , resulting from the tensor product of the two Hilbert spaces \mathcal{H}^A and \mathcal{H}^B , which correspond to subsystems A and B , respectively. Let ρ^{AB} be a density matrix in \mathcal{H}^{AB} , such that

$$\rho^{AB} = \rho^A \otimes \rho^B \quad (3.18)$$

where $\rho^A \in \mathcal{H}^A$ and $\rho^B \in \mathcal{H}^B$.

Definition 9. *The joint entropy for the composite system is then obviously given by [18]*

$$S(A, B) \equiv -\text{Tr}(\rho^{AB} \log(\rho^{AB})). \quad (3.19)$$

Definition 10. *The conditional entropy is given by [5]*

$$S(A|B) \equiv S(A, B) - S(B), \quad (3.20)$$

Definition 11. *The mutual information is defined by [5]*

$$S(A : B) \equiv S(A) + S(B) - S(A, B). \quad (3.21)$$

3.2.3 Properties

Some of the properties we introduced in section 3.1.3 are still valid in the field of quantum information, but this is not the case for some others. Here we give some useful properties of the von Neumann entropy [18].

1. The von Neumann entropy $S(\rho)$ is concave.
2. The von Neumann entropy is non-negative. The entropy is zero iff the state is pure.
3. In a Hilbert space of dimension d , the entropy is at most $\log d$. The entropy is equal to $\log d$ iff the system is in the completely mixed state \mathcal{I}/d .
4. If a composite system AB is in a pure state, then the entropies corresponding to the reduced density matrices are equal, $S(A) = S(B)$.
5. **Joint entropy theorem:** Suppose p_i are probabilities, $|i\rangle$ are orthogonal states for a system A , and ρ_i is any set of density operators for another system, B . Then

$$S\left(\sum_i p_i |i\rangle\langle i| \otimes \rho_i\right) = H(p_i) + \sum_i p_i S(\rho_i), \quad (3.22)$$

where $S(\rho)$ is the von Neumann entropy of ρ and $H(\rho)$ is the Shannon entropy of ρ .

One interesting feature of the quantum systems is that the von Neumann entropy of the composite system can be lower than the entropy of the subsystems. This is the case when the composite system exhibits some entanglement. We therefore have the following property:

Suppose $|\psi^{AB}\rangle$ is a pure state of a composite system in a Hilbert space \mathcal{H}^{AB} . $|\psi^{AB}\rangle$ is entangled iff

$$S(B|A) < 0. \quad (3.23)$$

Finally, in the quantum case, it is also possible to define the subadditivity. Suppose again a composite system AB . The joint entropy for the two systems satisfies the inequality [18]

$$S(A, B) \leq S(A) + S(B). \quad (3.24)$$

3.2.4 Connection with the theory of majorization

Like in the case of the Shannon entropy, the von Neumann entropy is connected to the concept of majorization, since it is a measure of disorder applied to quantum systems. We can state a proposition similar to proposition 4, and which relates the von Neumann entropy to the theory of majorization. The proposition is as follows.

Proposition 5. *Take two density matrices ρ and σ . If $\rho \prec \sigma$, then [17]*

$$S(\rho) \geq S(\sigma), \quad (3.25)$$

where S is the von Neumann entropy.

This proposition can be proved using the concavity of the von Neumann entropy. Suppose $\rho \prec \sigma$, then according to Uhlmann's theorem there exist probabilities p_j and unitaries U_j such that

$$\rho = \sum_j p_j U_j \sigma U_j^\dagger. \quad (3.26)$$

From the concavity of the von Neumann entropy it follows that

$$S(\rho) \geq \sum_j p_j S(U_j \sigma U_j^\dagger). \quad (3.27)$$

Since $S(U_j \sigma U_j^\dagger) = S(\sigma)$, we end up with

$$S(\rho) \geq S(\sigma). \quad (3.28)$$

Here we see the connection between the theory of majorization and the von Neumann entropy, as we did for the Shannon entropy. In fact, one can show that some similar relations can be found for all the Rényi entropies we are going to define in the next section. This is a consequence of property 3, as we already said.

3.3 Rényi entropy

Here we briefly introduce a notion which generalises the Shannon entropy. It will be useful later when studying some majorization relations. Suppose X is a random variable associated with a probability distribution p_k .

Definition 12. *Rényi's entropy of order α of X is defined by [22]*

$$I_\alpha(X) = \frac{1}{1-\alpha} \log \left(\sum_k p_k^\alpha \right) \quad (3.29)$$

with $\alpha \neq 1$ and $\alpha \geq 0$.

The Rényi entropy is another measurement of the uncertainty there is about the random variable X . In the limit where $\alpha \rightarrow 1$, it can be shown that I_α converges to the Shannon entropy (3.1).

Let us exhibit some properties of the Rényi entropy [22], since it is of great importance in this work, as we will see later.

1. The Rényi entropy $I_\alpha(X)$ is a continuous function of all the probabilities p_k associated to the random variable X .
2. The Rényi entropy $I_\alpha(X)$ is permutationally symmetric: that is, the position change of any two or more probabilities p_k associated to the random variable X will not change the entropy value.
3. Additivity: If the random variable X is associated to the probabilities p_k and if the random variable Y is associated to the probabilities q_k , where the p_k and the q_k are two independent probability distributions, then the joint entropy of X and Y is given by

$$I_\alpha(X, Y) = I_\alpha(X) + I_\alpha(Y). \quad (3.30)$$

4. The Rényi entropy $I_\alpha(X)$ is a Schur-concave function of the random variable X .

An additional interesting property is the fact that for all the values of α , the Rényi entropy of the uniform distribution W on d -elements is equal to $\log d$. Indeed, we have

$$I_\alpha(W) = \frac{1}{1-\alpha} \log \left(\sum_{k=1}^d \frac{1}{d^\alpha} \right) \quad (3.31)$$

with $\alpha \neq 1$ and $\alpha \geq 0$. Therefore,

$$I_\alpha(W) = \frac{1}{1-\alpha} \log \left(\frac{d}{d^\alpha} \right) = \frac{1}{1-\alpha} \log (d^{1-\alpha}) = \log d, \quad (3.32)$$

for all the $\alpha \geq 0$, with $\alpha \neq 1$. We consequently see that for the uniform distribution, all the measures of order provided by the Rényi entropies with different values of α are the same.

4 Quantum Entanglement

After introducing majorization, the main mathematical theory explored in the present work, along with the concept of entropy, it would be fitting to exhibit a fascinating physical phenomenon for which majorization is a powerful modelling tool. This phenomenon is nowadays known as entanglement and is considered as one of the main "resources" in the field of quantum information.

Quantum entanglement is a purely quantum mechanical resource [18]. It is a kind of quantum correlation between the states of different systems, which interacted at least at one moment of their existence. Entanglement is fascinating in the sense that even if the systems become separated by an arbitrarily important distance, they still can affect one another through this quantum effect, and the speed at which they do it is arbitrarily high. It is the effect which is responsible for a large quantity of interesting phenomena that would make the field of quantum information much less captivating if they didn't exist. One of them is quantum teleportation, which we are going to explain in the next section, in order to show how entanglement can be used as a resource for quantum information processes.

4.1 Quantum Teleportation

In order to illustrate the utility of quantum entanglement, let us introduce the concept of quantum teleportation [3] which, as the name suggests, would be impossible to perform if the only accessible resources were classical (no entanglement). Entanglement is indeed fundamental if we wish to perform quantum teleportation. Suppose Alice wants to send some unknown quantum bit to Bob. Suppose, further, that they can only communicate via classical means, and that they have at their disposal one pair of so called entangled qubits. [12] What Alice could do is to measure the qubit and try to guess the state knowing the results of the measurement, and then describe it to Bob via their classical channel. However, this way there is a certain probability that the state described by Alice won't be the real qubit. Therefore, Alice needs a way to transfer the state with certainty, and this can be done using entanglement as a resource.

Suppose that the state which is to be teleported is [18]

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad (4.1)$$

where α and β are unknown. The entangled state that Alice and Bob have at their disposal is one of the four EPR states, or Bell states, which are defined as maximally entangled quantum states of two qubits. Suppose the EPR state Alice and Bob possess is of the form

$$|\psi_+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \quad (4.2)$$

In order to perform teleportation, Alice and Bob begin by taking the tensor product of the state $|\psi\rangle$ and the EPR state

$$|\psi_0\rangle = |\psi\rangle |\psi_+\rangle = \frac{1}{\sqrt{2}} [\alpha|0\rangle (|00\rangle + |11\rangle) + \beta|1\rangle (|00\rangle + |11\rangle)] \quad (4.3)$$

In our convention, the first two qubits on the left belong to Alice (It is Alice who wants to teleport her state $|\psi\rangle$ to Bob, and she possesses a part of the EPR state) and the third qubit belongs to Bob (Bob possesses the other part of the EPR state). Alice then send her two qubits through a CNOT gate. Without going into details, we will only show the effect of the CNOT gate on the states $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$. The action of the CNOT gate on these states is the following

$$\begin{aligned} |00\rangle &\mapsto |00\rangle \\ |01\rangle &\mapsto |01\rangle \\ |10\rangle &\mapsto |11\rangle \\ |11\rangle &\mapsto |10\rangle \end{aligned} \quad (4.4)$$

After sending her two qubits into the CNOT gate, Alice obtains

$$|\psi_1\rangle = \frac{1}{\sqrt{2}} [\alpha|0\rangle (|00\rangle + |11\rangle) + \beta|1\rangle (|10\rangle + |01\rangle)] \quad (4.5)$$

Another important gate in the field of quantum information is the Hadamard gate, which has the following effect on the particular states $|0\rangle$, $|1\rangle$, $(1/\sqrt{2})(|0\rangle - |1\rangle)$ and $(1/\sqrt{2})(|0\rangle + |1\rangle)$

$$\begin{aligned} |0\rangle &\mapsto \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ |1\rangle &\mapsto \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \\ \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) &\mapsto |1\rangle \\ \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) &\mapsto |0\rangle \end{aligned} \quad (4.6)$$

Alice sends her first qubits through the Hadamard gate, obtaining

$$|\psi_2\rangle = \frac{1}{2} [\alpha(|0\rangle + |1\rangle)(|00\rangle + |11\rangle) + \beta(|0\rangle - |1\rangle)(|10\rangle + |01\rangle)] \quad (4.7)$$

The state $|\psi_2\rangle$ can be rearranged into

$$|\psi_2\rangle = \frac{1}{2} [|00\rangle (\alpha|0\rangle + \beta|1\rangle) + |01\rangle (\alpha|1\rangle + \beta|0\rangle) + |10\rangle (\alpha|0\rangle - \beta|1\rangle) + |11\rangle (\alpha|1\rangle - \beta|0\rangle)] \quad (4.8)$$

We see from expression (4.8) that there are four terms, in which we have different combinations of the qubits $|0\rangle$ and $|1\rangle$, and of the amplitudes α and β (third qubit). For example, in the first term, we find the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, but in the other terms, some operations should be done on the third qubit in order to obtain the wanted state $|\psi\rangle$. Therefore, if Alice performs some measurement on her two qubits and obtains the result 00, then Bob's system will be in the state $|\psi\rangle$. Similarly, we can read Bob's state according to Alice's measurement:

$$\begin{aligned} 00 &\longmapsto |\psi_3(00)\rangle \equiv [\alpha|0\rangle + \beta|1\rangle] \\ 01 &\longmapsto |\psi_3(01)\rangle \equiv [\alpha|1\rangle + \beta|0\rangle] \\ 10 &\longmapsto |\psi_3(10)\rangle \equiv [\alpha|0\rangle - \beta|1\rangle] \\ 11 &\longmapsto |\psi_3(11)\rangle \equiv [\alpha|1\rangle - \beta|0\rangle] \end{aligned} \tag{4.9}$$

If Alice's measurement results in a different state than $|00\rangle$, Bob needs to perform some quantum operation on his state in order to recover the state $|\psi\rangle$. In order to explain this, we need to introduce two other quantum gates known as the sign flip and the bit flip. The effect of the sign flip is the following:

$$a|0\rangle + b|1\rangle \longmapsto a|0\rangle - b|1\rangle \tag{4.10}$$

while the effect of the bit flip is the following

$$a|0\rangle + b|1\rangle \longmapsto a|1\rangle + b|0\rangle \tag{4.11}$$

for any two amplitudes a and b . For example, the combined effects of the two doors gives, if we begin by applying the bit flip, followed by the sign flip

$$a|0\rangle + b|1\rangle \longmapsto -a|1\rangle + b|0\rangle \tag{4.12}$$

for any two amplitudes a and b . If we apply the two doors in the opposite order, we get

$$a|0\rangle + b|1\rangle \longmapsto a|1\rangle - b|0\rangle \tag{4.13}$$

for any two amplitudes a and b . Going back to Alice and Bob's problem, we see that if Alice gets the measurement result 01, Bob needs to apply the bit flip to his state. If Alice's measurement result is 10, he needs to apply the sign flip to his states. Finally, if the measurement result is 11, Bob needs to apply both gates to his state. Of course, Alice needs to send her measurement result to Bob in order for teleportation to be possible, therefore the process of teleportation cannot be faster than light. However, teleportation wouldn't have possible without the entanglement resource. This example therefore shows the importance of entanglement in the field of quantum information.

4.2 Some important entangled states : The Bell states

Let us now introduce some well known entangled states, which will be useful in the following because they will allow us to verify some majorization relations.

A fundamental group of quantum states is the class composed by the Bell states, or EPR states, which are defined as

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B) \tag{4.14}$$

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B - |1\rangle_A \otimes |1\rangle_B) \quad (4.15)$$

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |1\rangle_B + |1\rangle_A \otimes |0\rangle_B) \quad (4.16)$$

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |1\rangle_B - |1\rangle_A \otimes |0\rangle_B) \quad (4.17)$$

These states are named after John S. Bell, who worked on his Bell inequality. The EPR states characterize well the type of systems described by Einstein, Podolsky and Rosen in their EPR paper [18], and are therefore a good example of entanglement. Indeed, it is easy to notice that there are some correlations between the two subsystems composing the total system, and that these correlations go beyond the classical world. In order to see that, take the first Bell state

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B). \quad (4.18)$$

If one measures the first qubits of the state, one either obtains the qubit $|0\rangle_A$ with probability $1/2$, or the qubit $|1\rangle_A$ with probability $1/2$. This is also the case for the second qubit. The first thing we can conclude is that the state gives no knowledge about the subsystems [12], since the subsystems are described by a probability distribution, and are therefore in a mixture of states. However, the state of the total subsystem is pure, and we have maximal knowledge about the whole subsystem. This is disturbing, because it means there is more uncertainty about the subsystems than about the whole system. This would not be possible classically, and is a direct result of quantum entanglement. We will investigate this in more details when we will be talking about entropies, which mathematically characterize the uncertainty about a state. Another interesting thing about the state is the entanglement itself, which can also be easily seen after measuring the first qubit. If we do the measurement and obtain $|0\rangle_A$, then the second qubit is automatically $|0\rangle_B$. However, if we obtain $|1\rangle_A$, then the second qubit is automatically $|1\rangle_B$. This is another feature which cannot be observed classically, and which characterizes entanglement.

4.3 Separable states

We have seen in the previous sections that some states exhibit some kind of quantum correlations called entanglement. These states sometimes don't verify Bell's inequalities. However, all the states that have no entanglement at all (classical states) verify Bell's inequalities. These classical states are called separable states.

Definition 13. *A separable mixed state of n systems is a state that can be written as a convex combination of product states [12]*

$$\rho = \sum_i p_i \rho_1^i \otimes \cdots \otimes \rho_n^i, \quad (4.19)$$

where a mixed product state of n systems is a state that can be written in the form

$$\rho_{\otimes} = \rho_1 \otimes \cdots \otimes \rho_n. \quad (4.20)$$

A mixed entangled state of n systems is a state that cannot be written in the form (4.19). It is important to be able to distinguish between separable state and entangled states, in order to know what we are able to do with the state involved. For example, it is impossible to perform teleportation using only separable states. The problem is that it is often difficult to prove that a state cannot be written in the form (4.19). It is sometimes easy to find such a form for a state (when it is possible) and prove that the state is separable, but proving this way that a state is entangled is very difficult, and the method is not a good approach. Therefore, we need some criteria for separability or entanglement, in the form of some mathematical criteria or inequalities for example. Later, we are going to present some of the most important separability criteria. These criteria will make use of the concept of majorization and of the notion of entropy we already introduced before. Before this, let us introduce some notions which will clarify the connexion between the theory of majorization and the notions of entanglement and separability.

4.4 Characterization of entanglement and separability with majorization

In the previous sections, we have introduced the concepts of entanglement and separability. The theory of majorization is in fact also closely related to these two notions. In section 2.3, we presented the concept of local operations and classical communication (LOCC). We also showed how majorization was a powerful tool in order to investigate the possibility of performing LOCC on bipartite states. In fact, it is also possible to analyze the evolution of entanglement of a bipartite state when performing LOCC. The notion of majorization therefore becomes automatically connected to the entanglement on a state through the concept of LOCC. In order to see this, let us introduce the notion of entanglement monotone [24], a measure of entanglement, through the following definition.

Definition 14. *We call entanglement monotone EM any magnitude $\mu(\rho)$ that does not increase, on average, under local transformations.*

Now consider the ordered Schmidt coefficients $\lambda_1^\downarrow \geq \dots \geq \lambda_d^\downarrow \geq 0$ of a state ψ , where the Schmidt decomposition has been defined in section 2.2.4 For each $l = 1, \dots, d$, define the following entanglement monotone $E_l(\psi)$ as [17]

$$E_l(\psi) \equiv \sum_{i=l}^n \lambda_i. \quad (4.21)$$

We have the following theorem.

Theorem 2. *The pure state ψ can be transformed into one element of the ensemble $\{p_j, \psi_j\}$ using only LOCC if and only if*

$$E_l(\psi) \geq \sum_j p_j E_l(\psi_j), \quad l = 1, \dots, d. \quad (4.22)$$

Now, one generalization of theorem 1 is given by the following theorem.

Theorem 3. *The pure state ψ can be transformed into one element of the ensemble $\{p_j, \psi_j\}$ using only LOCC if and only if*

$$\lambda(\psi) \prec \sum_j p_j \lambda(\psi_j), \quad (4.23)$$

where $\lambda(\psi)$ ($\lambda(\psi_j)$) corresponds to the spectrum of the reduced density matrix $\rho_A^\psi \equiv \text{Tr}_B |\psi\rangle\langle\psi|$ ($\rho_A^{\psi_j} \equiv \text{Tr}_B |\psi_j\rangle\langle\psi_j|$).

We see that in fact, each one of the inequalities in equation (4.22) is just one of the majorization inequalities in equation (4.23). The notion of majorization is therefore here equivalent to the notion of entanglement monotone when it comes to investigating the possibility of performing LOCC.

As for the notion of separability, as we are going to show in the next section, the theory of majorization can be used in particular to find conditions for the separability of quantum states.

5 Separability criteria

In section 4, we introduced the concept of quantum entanglement, which is, as we already showed, central in the field of quantum information. We then remarked, in section 4.3, that some states, called separable states, don't exhibit entanglement at all. Since this property of entanglement is so important, it is useful to have some means to distinguish between entangled states and separable states. Indeed, there are mathematical criteria which allow us to detect entanglement. We will now introduce them, using the helpful concepts of von Neumann entropy and majorization, which we already introduced in section 3.2 and section 1, respectively. Most of these criteria are conditions which are only necessary, but one of them is also sufficient (in some particular cases, as we will later see).

5.1 Positive partial transpose

5.1.1 Arbitrary dimension

The first separability criterion we exhibit is also the most famous. It was proved by Asher Peres in 1996 in his article *Separability Criterion for Density Matrices* [21]. It is a necessary condition for separability of quantum states, and can be stated as follows:

Separability Criterion 1. *If a state ρ is separable, then none of the eigenvalues of its partial transpose σ is negative.*

Note that the partial transpose is defined by (2.21). As an example, take the bipartite 2×2 state introduced by Gisin [9] and defined by the density matrix

$$\rho_1 = \begin{pmatrix} \frac{1-x}{2} & 0 & 0 & 0 \\ 0 & x|a|^2 & xab^* & 0 \\ 0 & xa^*b & x|b|^2 & 0 \\ 0 & 0 & 0 & \frac{1-x}{2} \end{pmatrix}. \quad (5.1)$$

where x is a real number between 0 and 1 and a and b are two complex numbers which verify

$$|a|^2 + |b|^2 = 1. \quad (5.2)$$

This state can be seen as [21] a fraction x of the pure state $a|01\rangle + b|10\rangle$ and fractions $(1-x)/2$ of the pure states $|00\rangle$ and $|11\rangle$. The partial transpose of ρ_1 with respect to the first subsystem is given by

$$\rho_1^{\text{T}_A} = \begin{pmatrix} \frac{1-x}{2} & 0 & 0 & xa^*b \\ 0 & x|a|^2 & 0 & 0 \\ 0 & 0 & x|b|^2 & 0 \\ xab^* & 0 & 0 & \frac{1-x}{2} \end{pmatrix}, \quad (5.3)$$

and the eigenvalues of the partial transpose $\rho_1^{\text{T}_A}$ of ρ_1 are given by the vector

$$\lambda_1 = \left(x|a|^2, x|b|^2, \left(\frac{1-x}{2} - x|ab| \right), \left(\frac{1-x}{2} + x|ab| \right) \right). \quad (5.4)$$

The matrix $\rho_1^{\text{T}_A}$ therefore has a negative eigenvalue when

$$x > \frac{1}{1 + 2|ab|}. \quad (5.5)$$

This limit can be compared with the limit given by the violation of Bell's inequalities. This last limit is given by [9]

$$x > \frac{1}{1 + 2|ab|(\sqrt{2} - 1)}, \quad (5.6)$$

which shows that the Bell inequality test is weaker than the one allowed by Peres' criterion (at least in this case).

5.1.2 2×2 and 2×3 states

R. Horodecki, P. Horodecki and M. Horodecki introduced an extension of Peres' criterion in 1996 in their article *Separability of mixed states: necessary and sufficient conditions* [10]. In the case of 2×2 and 2×3 quantum states, Peres' criterion is also sufficient. It can be stated as follows

Separability Criterion 2. *A state ρ acting on $\mathbb{C}^2 \otimes \mathbb{C}^2$ or $\mathbb{C}^2 \otimes \mathbb{C}^3$ is separable if and only if its partial transposition is a positive operator.*

As an example, take the Werner state

$$\rho_2 = p|\psi^-\rangle\langle\psi^-| + \frac{1-p}{4}\mathcal{I}. \quad (5.7)$$

where p is a real number between 0 and 1. The state can be rewritten

$$\rho_2 = \frac{1}{4} \begin{pmatrix} 1-p & 0 & 0 & 0 \\ 0 & 1+p & -2p & 0 \\ 0 & -2p & 1+p & 0 \\ 0 & 0 & 0 & 1-p \end{pmatrix} \quad (5.8)$$

The partial transpose of ρ_2 with respect to the first subsystem is given by

$$\rho_2^{\text{T}^A} = \frac{1}{4} \begin{pmatrix} 1-p & 0 & 0 & -2p \\ 0 & 1+p & 0 & 0 \\ 0 & 0 & 1+p & 0 \\ -2p & 0 & 0 & 1-p \end{pmatrix}, \quad (5.9)$$

and the eigenvalues of the partial transpose $\rho_2^{\text{T}^A}$ of ρ_2 are given by the vector

$$\lambda_1 = \left(\frac{1}{4}(1+p), \frac{1}{4}(1+p), \frac{1}{4}(1+p), \frac{1}{4}(1-3p) \right). \quad (5.10)$$

Since we have $p \geq 0$, the matrix $\rho_2^{\text{T}^A}$ only has positive eigenvalues when

$$p < \frac{1}{3}. \quad (5.11)$$

Furthermore, it is known [10] that if condition (5.11) is fulfilled, ρ_2 is a separable state. We consequently see that the criteria we used here is both a necessary and sufficient condition.

5.2 Conditional entropy criteria

5.2.1 Comparison between entropy of a system and entropies of its subsystems

The third separability criterion we will present was introduced by R. Horodecki, P. Horodecki and M. Horodecki in 1996 in their article *Quantum α -entropy inequalities: independent condition for local realism?* [11]. It is a necessary condition for separability of quantum states, and can be stated as follows:

Separability Criterion 3. *For any separable state ρ on the finite dimensional Hilbert space, the inequality*

$$I_\alpha(\rho) \geq \max_{i=1,2} I_\alpha(\rho_i) \quad (5.12)$$

where $\alpha \geq 1$ and $\rho_i, i = 1, 2$ are the states corresponding to the subsystems of ρ , is satisfied for $\alpha = 1, 2$.

Note that $I_\alpha(\rho)$ is the Rényi entropy of order α of state ρ , already defined in section 3.3. As we are going to see in the next section with the particular case of the von Neumann entropy, criterion 3 is in fact a condition on the mutual entropies.

5.2.2 Positivity of the conditional entropy

We already introduced the von Neumann entropy of A conditional on B , given by

$$S(A|B) = S(A, B) - S(B). \quad (5.13)$$

From this expression, we see that relation (5.12) is similar to the following relations

$$\begin{cases} S(A|B) \geq 0 \\ S(B|A) \geq 0 \end{cases} \quad (5.14)$$

in terms of von Neumann entropies. We therefore see that if a state ρ_{AB} is separable, than relations (5.14) are satisfied. This leads to the next criterion we will give.

5.2.3 Conditional amplitude operator

The fourth separability criterion we give was introduced by N. J. Cerf and C. Adami in 1999 in their article *Quantum extension of conditional probability* [6]. It is a necessary condition for separability of quantum states based on a new object, the conditional amplitude operator. This criterion is in fact based on relations (5.14). It can be stated as follows.

Separability Criterion 4. *Any separable bipartite state satisfies the condition $\rho_{A|B} \leq 1$, where the conditional amplitude operator of A conditional on B $\rho_{A|B}$ is defined as*

$$\begin{aligned} \rho_{A|B} &= \exp_2 [\log_2 \rho_{AB} - \log_2 (\mathcal{I}_A \otimes \rho_B)] \\ &= \lim_{n \rightarrow \infty} \left[\rho_{AB}^{1/n} (\mathcal{I}_A \otimes \rho_B)^{1/n} \right] \end{aligned} \quad (5.15)$$

The conditional entropy operator $\rho_{A|B}$ is not a density operator in general, since its eigenvalues can exceed one. It is however worth defining, since it gives a criterion which is in fact stronger than criterion 3.

5.3 Reduction criterion for separability

The fifth separability criterion we give was introduced by N. J. Cerf, C. Adami and R. M. Gingrich in 1998 in their article *Reduction criterion for separability* [4]. In order to state it, let us first define the linear map Λ which maps Hermitian operators on \mathcal{H}_{AB} into Hermitian operators on \mathcal{H}_{AB} by

$$\Lambda : \rho_{AB} \rightarrow \lambda_{AB} \equiv \mathcal{I}_A \otimes \rho_B - \rho_{AB} \quad \text{with } \rho_B = \text{Tr}_A [\rho_{AB}]. \quad (5.16)$$

The separability criterion is given by

Separability Criterion 5. *A necessary condition for the separability of the state ρ_{AB} of a bipartite system AB is that it is mapped by Λ into a positive semidefinite operator, namely*

$$\Lambda \rho_{AB} \geq 0. \quad (5.17)$$

5.4 Separability criterion based on the theory of majorization

The final separability criterion we show was proved by M. A. Nielsen and J. Kempe in 2001 in their article *Separable States Are More Disordered Globally than Locally* [19]. It is a necessary condition for separability of quantum states, based on the majorization theory. The criterion can be stated as follows

Separability Criterion 6. *If the state ρ_{AB} is separable, then*

$$\lambda(\rho_{AB}) \prec \lambda(\rho_A) \quad \text{and} \quad \lambda(\rho_{AB}) \prec \lambda(\rho_B), \quad (5.18)$$

where $\lambda(\rho_A)$, $\lambda(\rho_B)$ and $\lambda(\rho_{AB})$ are the vectors of eigenvalues of ρ_{AB} and the corresponding reduced density matrices. [By convention we append zeros to the vectors $\lambda(\rho_A)$ and $\lambda(\rho_B)$ so they have the same dimension as $\lambda(\rho_{AB})$.]

This criterion is based on the fact that if a state is separable, then the relations given by system (5.14) are satisfied. The relations are indeed equivalent to

$$S(AB) \geq S(A) \quad \text{and} \quad S(AB) \geq S(B). \quad (5.19)$$

The last inequalities are very similar to the majorization relations (5.18), but they allow us to establish a stronger criterion than the one based on the von Neumann entropies.

Take for example the pure bipartite state

$$\psi_{AB} = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle). \quad (5.20)$$

Because it is a pure state, its only eigenvalue is equal to 1. The state corresponding to the subsystem A is given by

$$\begin{aligned} \rho_A &= \text{Tr} [\rho_{AB}] \\ &= \text{Tr} [|\psi_{AB}\rangle \langle \psi_{AB}|] \\ &= \frac{1}{2} (|0\rangle \langle 0| + |1\rangle \langle 1|) \\ &= \mathcal{I}_A \end{aligned} \quad (5.21)$$

which has eigenvalues $(1/2, 1/2)$. Using the theory of majorization, we easily see that condition (5.18) is not satisfied. This is the case because the bipartite state is entangled.

6 Going beyond entropic inequalities using the theory of majorization

The main objective of the present project is to seek new applications of the majorization theory in the field of quantum information. In this section, we are going to present the work that has been done applying the majorization theory to the field of quantum information. We will of course exhibit the main interesting results we have found, but we will also present the way we came to find these results, as well as the examples that allowed us to sometimes illustrate the results, and sometimes to show that what we were trying to prove was in fact not true.

The first area of research we will be investigating is the one related to the entropic inequalities. Indeed, in this section, we will try to establish some majorization relations based on already existing entropic inequalities. We will be working with quantum systems which can be described by states evolving in Hilbert spaces of finite dimensions. The Hilbert spaces involved will often be 2-dimensional, which means that any state of these Hilbert spaces can be written as a linear combination of two orthonormal vectors. In the following, we will always define these two orthonormal vectors as the two vectors (in bracket notation) $|0\rangle$ and $|1\rangle$. Since these two vectors are orthonormal, we have the properties

$$\begin{cases} \langle 0|1\rangle = \langle 1|0\rangle = 0 \\ \langle 0|0\rangle = \langle 1|1\rangle = 1 \end{cases} \quad (6.1)$$

These two vectors in fact allow us, as we already said before, to define the well know qubit, which can be written as in equation (7.1), and which verifies property (7.2). We will often work with bipartite states. Here, we repeat some general definitions already

given in the previous sections just to fix ideas. A bipartite state is a state which represent a system which is the result of the combination of two subsystems. If the total system AB is given by the combination of system A represented by states evolving in a Hilbert space \mathcal{H}^A and system B represented by states evolving in a Hilbert space \mathcal{H}^B , then it is represented by states evolving in a Hilbert space

$$\mathcal{H}^{AB} = \mathcal{H}^A \otimes \mathcal{H}^B. \quad (6.2)$$

The operations of trace, partial trace, transposition and partial transposition described earlier will allow us to swich from one system to the other.

In this section, we are going to present some majorization relations we found. We will then study the last one in details, giving examples for which the relation is satisfied and examples for which it is not. We will then explain why it is not verify in general using the concept of Schur-concavity of the Rényi entropies, and try to investigate it further when it is not verified, using the concept of catalysis. Finally, we will introduce an interesting class of states, the Bell diagonal states, for which the last majorization relation is always verified.

6.1 Nielsen and Kempe's result

One very important result using majorization upon which we can rely is the separability criterion 6. Indeed, Nielsen and Kempe's idea was to find a majorization relation which resembles a well known inequality, which is at the root of a property of separability. This property states that if a state is separable, than the entropy of the total system AB is greater or equal to the entropy of both subsystems A and B , namely

$$S(A) \leq S(A, B) \quad \text{and} \quad S(B) \leq S(A, B), \quad (6.3)$$

where $S(A)$, $S(B)$ and $S(A, B)$ are the von Neumann entropies of system A , system B and of the joint system, respectively. If we think "classically", this result is pretty intuitive. Indeed, if we have a bipartite system AB , and if that system does not exhibit any entanglement, which means that the system can be described as "classical" (no "quantum" phenomenon), then obviously, the incertitude about the total system cannot be smaller than the incertitude of one of the two subsystems. We showed before that the incertitude on a system (or state) can be described by the von Neumann entropy. This means that the entropy of the total system AB cannot be smaller than the entropy of one of the two subsystems A or B . This is in fact the separability criterion 3, with $\alpha = 1$. We can therefore say that Nielsen and Kempe's criterion is based on criterion 3. It is however stronger. Indeed, It has been shown that criterion 6 implies criterion 3.

In the following, we are going to show how we tried to rely on some existing entropy inequalities, and how we tried to find majorization relation based on these inequalities. We will try to explain why the relation is wrong when this is the case, or simply find a counterexample.

6.2 Positivity of the joint entropy of a separable state in terms of majorization

The first idea was to take an inequality equivalent to (6.3) and find some similar majorization relation based on an inequality involving the conditional entropies. The inequality

is in fact at the root of the proposition which is the same as criterion 3, but which is expressed differently. It simply says that if the joint system composed of subsystems A and B is separable, then

$$S(B|A) = S(A, B) - S(A) \geq 0 \quad (6.4)$$

and

$$S(A|B) = S(A, B) - S(B) \geq 0. \quad (6.5)$$

The idea also involved the fact that the entropy of a pure state is null, and that we could therefore represent the zero entropy by any pure state. In order to find a state corresponding to the conditional entropy, we use the definition of the conditional amplitude operator introduced by Cerf and Adami, and which is defined by equation (5.15) in criterion 4. This idea first led to the following proposition.

Proposition 6. *If the state ρ_{AB} is separable, then*

$$\lambda(\rho_{B|A}) \prec \lambda(\rho_P) \quad \text{and} \quad \lambda(\rho_{A|B}) \prec \lambda(\rho_P), \quad (6.6)$$

for any pure state ρ_P , where $\lambda(\rho)$ is the vector of eigenvalues of ρ , and where $\rho_{A|B}$ and $\rho_{B|A}$ are the conditional amplitude operators defined by

$$\begin{aligned} \rho_{A|B} &= \exp_2[\log_2 \rho_{AB} - \log_2(\mathbb{I}_A \otimes \rho_B)] \\ &= \lim_{n \rightarrow \infty} [\rho_{AB}^{1/n} (\mathbb{I}_A \otimes \rho_B)^{-1/n}]^n \end{aligned} \quad (6.7)$$

Conclusion An important characteristic of the majorization relations we are going to study is the fact that they often involve density matrices. Thereby, this implies that if one of the sides of the majorization relation contains a density matrix, then the other side should also be a density matrix. Indeed, when we look at system (1.1), we notice that the last equality is the sum of the eigenvalues of the vector involved in the majorization relation. Knowing that the sum of the eigenvalues of a density matrix is always 1 since the trace of a density matrix is equal to 1, this means that the last equality of system (1.1) should always be equal to one when a state is involved in the majorization relation, which means that the other vector of the majorization relation should also be a density matrix, with trace equal to 1. Therefore, one can say that proposition 6 is not verified in general, since $\rho_{B|A}$ is not a state in general, knowing that its eigenvalues can exceed 1 [5]. Furthermore, it is not very interesting to compare some states to pure states using the majorization theory, since for pure states, there is only one eigenvalue which is 1. In fact, any state is majorized by any pure state. This can easily be seen from system (1.1).

6.3 First adaptation of the subadditivity relation to the theory of majorization

The second idea consisted in finding a relation similar to the subadditivity relation

$$S(A, B) \leq S(A) + S(B). \quad (6.8)$$

The first possibility was to try to take the sum of the two vectors of eigenvalues which correspond better to $S(A)$ and $S(B)$. Since each side of the majorization relation should be a vector of trace 1, a possible relation was the following.

Proposition 7. *If A and B are two systems, then*

$$\lambda(\rho_{AB}) \succ \frac{1}{2}(\lambda^\downarrow(\rho_A) + \lambda^\downarrow(\rho_B)) \quad (6.9)$$

where $\lambda(\rho)$ is the vector of eigenvalues of ρ .

Conclusion Notice that we have chosen to order the vectors $\lambda(\rho_A)$ and $\lambda(\rho_B)$ in the majorization relation. Indeed, when some majorization relation is applied, the order of the elements in the vectors is not important, since the vectors are ordered when verifying the majorization relation, as it can be seen from the definition of the majorization. Here we chose to order the vectors in (6.9) in order to have the "worst situation". Indeed, we can choose to order the vectors $\lambda(\rho_A)$ and $\lambda(\rho_B)$ however we want in the expression

$$\frac{1}{2}(\lambda(\rho_A) + \lambda(\rho_B)). \quad (6.10)$$

Since we would like to have a maximum in the right hand term of (6.9), we should order both $\lambda(\rho_A)$ and $\lambda(\rho_B)$ before as

$$\frac{1}{2}(\lambda^\downarrow(\rho_A) + \lambda^\downarrow(\rho_B)). \quad (6.11)$$

However, choosing to order the vectors before is somehow arbitrary. Furthermore, taking the sum of two vectors of eigenvalues does not make much sense. Also, it is possible to find some counterexample to relation (6.9). Take the Werner state

$$\rho_{AB} = p |\psi^-\rangle \langle \psi^-| + \frac{1-p}{4} \mathcal{I}. \quad (6.12)$$

where p is a real number between 0 and 1. We saw that its eigenvalues are given by the vector

$$\lambda_{AB} = \left(\frac{1}{4}(1+p), \frac{1}{4}(1+p), \frac{1}{4}(1+p), \frac{1}{4}(1-3p) \right). \quad (6.13)$$

The state corresponding to subsystems A and B can be computed using the partial trace. They are given by

$$\rho_A = \text{Tr}_B [\rho_{AB}] = \frac{1}{2} (|0\rangle \langle 0| + |1\rangle \langle 1|) \quad (6.14)$$

and

$$\rho_B = \text{Tr}_A [\rho_{AB}] = \frac{1}{2} (|0\rangle \langle 0| + |1\rangle \langle 1|) \quad (6.15)$$

Their vectors of eigenvalues are therefore given by

$$\lambda_A = \left(\frac{1}{2}, \frac{1}{2} \right) \quad (6.16)$$

and

$$\lambda_B = \left(\frac{1}{2}, \frac{1}{2} \right). \quad (6.17)$$

If we apply definition (1.1) of majorization to relation (6.9), we obtain the following conditions on p :

$$\begin{cases} p \geq \frac{1}{3} \\ p \geq 1 \\ p \geq 1 \end{cases} \quad (6.18)$$

which are of course not verified in general. Therefore, we can say for sure that relation (6.9) is not verified in general.

6.4 Second adaptation of the subadditivity relation to the theory of majorization

6.4.1 Relation and examples

In the previous section, we saw that taking the sum of two vectors before applying some majorization relation didn't make much sense. The idea here is to try to find again a majorization relation similar to the subadditivity relation, but without using a sum of vectors. In order to understand how we thought about the next majorization relation, let us introduce an interesting property of the von Neumann entropy. Suppose we have a joint bipartite system AB . Suppose furthermore that there is no entanglement between the system, and that there is no classical correlations either. In this case, the joint state of system AB is simply a product state given by

$$\rho_{AB} = \rho_A \otimes \rho_B \quad (6.19)$$

where ρ_A and ρ_B are the density matrices corresponding to systems A and B , respectively. Let us compute the von Neumann entropy of the joint system. In order to do that, suppose ρ_A can be written, in its diagonal representation, as

$$\rho_A = \sum_i \lambda_i^A |i^A\rangle \langle i^A| \quad (6.20)$$

where $|i^A\rangle$ are orthogonal states, and λ_i^A are the eigenvalues of ρ_A . Its von Neumann entropy is given by

$$S_A = - \sum_i \lambda_i^A \log \lambda_i^A = H_A \quad (6.21)$$

where H_A is the Shannon entropy $H(\lambda_i^A)$ of the probability distribution λ_i^A , since the eigenvalues can indeed be seen as a probability distribution. We can now compute the von Neumann entropy of the joint system AB ,

$$\begin{aligned} S_{AB} &= S(\rho_{AB}) \\ &= S(\rho_A \otimes \rho_B) \\ &= S\left(\left[\sum_i \lambda_i^A |i^A\rangle \langle i^A|\right] \otimes \rho_B\right) \\ &= H(\lambda_i^A) + \sum_i \lambda_i^A S(\rho_B) \\ &= H(\lambda_i^A) + S(\rho_B) \\ &= S(\rho_A) + S(\rho_B) \end{aligned} \quad (6.22)$$

where we used the Joint entropy theorem introduced in section 3.2.3. We therefore see that when we have a product state of a joint system, the Von Neumann entropy of the joint system is simply equal to the sum of the von Neumann entropies of the subsystems. This property led to the following proposition.

Proposition 8. *If A and B are two systems, then*

$$\lambda(\rho_{AB}) \succ \lambda(\rho_S) \quad (6.23)$$

where

$$\rho_S = \rho_A \otimes \rho_B = \text{Tr}_B(\rho_{AB}) \otimes \text{Tr}_A(\rho_{AB}) \quad (6.24)$$

and ρ_{AB} , ρ_A and ρ_B are the density matrices corresponding to the joint system AB , system A and system B , respectively, and $\lambda(\rho)$ is the vector of eigenvalues of ρ .

Equation (6.23) seems to be the best adaptation of the subadditivity relation in terms of majorization theory, because of the property of the entropy of a product state we introduced just before proposition 8.

Conclusion First notice that relation (6.23) is trivially verified for pure state, since in this case ρ_{AB} only has eigenvalue 1, and for product states, since in this case $\rho_S = \rho_{AB}$. However, the relation is not verified in general, even in the case of separable states ρ_{AB} . In order to verify that, take the state [15]

$$\rho_{AB}^\gamma = g(\gamma)[|00\rangle\langle 00| + |11\rangle\langle 11|] + \frac{\gamma}{2}[|00\rangle\langle 11| + |11\rangle\langle 00|] + (1 - 2g(\gamma))|01\rangle\langle 01|, \quad (6.25)$$

where γ is a real number between 0 and 1, and $g(\gamma)$ is defined as

$$g(\gamma) = \begin{cases} \frac{\gamma}{2}, & \gamma \geq \frac{2}{3} \\ \frac{1}{3}, & \gamma < \frac{2}{3} \end{cases} \quad (6.26)$$

Before studying relation (6.23) using this state, let us try to establish if ρ_{AB}^γ is separable or entangled. In order to do this, we will apply criterion 2. ρ_{AB}^γ can be written in matrix form as

$$\rho_{AB}^\gamma = \begin{pmatrix} g(\gamma) & 0 & 0 & \frac{\gamma}{2} \\ 0 & 1 - 2g(\gamma) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{\gamma}{2} & 0 & 0 & g(\gamma) \end{pmatrix}. \quad (6.27)$$

Its partial transpose is given by

$$\rho_{AB}^{\gamma \text{TA}} = \begin{pmatrix} g(\gamma) & 0 & 0 & 0 \\ 0 & 1 - 2g(\gamma) & \frac{\gamma}{2} & 0 \\ 0 & \frac{\gamma}{2} & 0 & 0 \\ 0 & 0 & 0 & g(\gamma) \end{pmatrix} \quad (6.28)$$

and the eigenvalues of the partial transpose $\rho_{AB}^{\gamma \text{TA}}$ of ρ_{AB}^γ are given by the vector

$$\lambda_{AB}^{\gamma \text{TA}} = \left(g(\gamma), g(\gamma), 1 - 2g(\gamma) + \sqrt{(2g(\gamma) - 1)^2 + \gamma^2}, 1 - 2g(\gamma) - \sqrt{(2g(\gamma) - 1)^2 + \gamma^2} \right). \quad (6.29)$$

The eigenvalues of $\rho_{AB}^{\gamma \text{TA}}$ are therefore all positive if and only if

$$\begin{cases} 1 - 2g(\gamma) + \sqrt{(2g(\gamma) - 1)^2 + \gamma^2} > 0 \\ 1 - 2g(\gamma) - \sqrt{(2g(\gamma) - 1)^2 + \gamma^2} > 0 \end{cases} \quad (6.30)$$

These inequalities lead to the condition $\gamma = 0$, which means that $\rho_{AB}^{\gamma \text{TA}}$ is separable if and only if $\gamma = 0$.

Let us now try to investigate relation (6.23) with ρ_{AB}^γ . We begin by computing states ρ_A^γ and ρ_B^γ of subsystems A and B , respectively, using the partial trace. We obtain

$$\begin{aligned}\rho_A^\gamma &= \text{Tr}^B [\rho_{AB}^\gamma] \\ &= g(\gamma) [|0\rangle\langle 0| + |1\rangle\langle 1|] + (1 - 2g(\gamma)) |0\rangle\langle 0|\end{aligned}\quad (6.31)$$

and

$$\begin{aligned}\rho_B^\gamma &= \text{Tr}^A [\rho_{AB}^\gamma] \\ &= g(\gamma) [|0\rangle\langle 0| + |1\rangle\langle 1|] + (1 - 2g(\gamma)) |1\rangle\langle 1|\end{aligned}\quad (6.32)$$

The tensor product of ρ_A^γ and ρ_B^γ is consequently given by

$$\begin{aligned}\rho_S^\gamma &= \rho_A^\gamma \otimes \rho_B^\gamma \\ &= [g(\gamma) (1 - g(\gamma))] (|00\rangle\langle 00| + |11\rangle\langle 11|) \\ &\quad + [(g(\gamma))^2 - 2g(\gamma) + 1] (|01\rangle\langle 01|) + (g(\gamma))^2 (|10\rangle\langle 10|)\end{aligned}\quad (6.33)$$

The eigenvalues of ρ_{AB}^γ are given by

$$\lambda_{AB}^\gamma = \left(1 - 2g(\gamma), g(\gamma) + \frac{\gamma}{2}, g(\gamma) - \frac{\gamma}{2}, 0\right), \quad (6.34)$$

while the eigenvalues of ρ_S^γ are given by

$$\lambda_S^\gamma = (g(\gamma) (1 - g(\gamma)), (g(\gamma))^2 - 2g(\gamma) + 1, (g(\gamma))^2, g(\gamma) (1 - g(\gamma))). \quad (6.35)$$

Let us first take the case in which ρ_{AB}^γ is separable. We therefore have

$$\lambda_{AB}^\gamma = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0\right), \quad (6.36)$$

and

$$\lambda_S^\gamma = \left(\frac{2}{9}, \frac{2}{9}, \frac{4}{9}, \frac{1}{9}\right). \quad (6.37)$$

If we order the eigenvalues, we get

$$\lambda_{AB}^{\gamma\downarrow} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0\right), \quad (6.38)$$

and

$$\lambda_S^{\gamma\downarrow} = \left(\frac{4}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9}\right). \quad (6.39)$$

Since the first elements of the ordered eigenvalues verify $\lambda_{AB,1}^{\gamma\downarrow} < \lambda_{S,1}^{\gamma\downarrow}$, we can immediately conclude that relation (6.23) is not verified in the case of separable states.

We will now take the case in which ρ_{AB}^γ is entangled. This happens when we have $\gamma \neq 0$. A good way to first test if relation (6.23) is verified in this case is by using numerical methods. The Matlab program `majorization_gamma.m` verifies in which cases the majorization relation (6.23) is verified for state ρ_{AB}^γ . Using the program results in the fact that the majorization relation is verified for $\gamma \in [0.23, 1]$, but it is not for $\gamma \in [0, 0.22]$. We therefore see that even in the case of entangled state, the relation is not verified. In order to be more precise, let us search analytically for which values of γ relation (6.23) is not verified.

First case : $\gamma \geq 2/3$ In this case, we have $g(\gamma) = \gamma/2$. The eigenvalues of ρ_{AB}^γ and ρ_S^γ are therefore respectively given by

$$\lambda_{AB}^\gamma = (1 - \gamma, \gamma, 0, 0), \quad (6.40)$$

and

$$\lambda_S^\gamma = \left(\frac{\gamma}{2} \left(1 - \frac{\gamma}{2}\right), \frac{\gamma}{2} \left(1 - \frac{\gamma}{2}\right), \left(\frac{\gamma}{2} - 1\right)^2, \frac{\gamma^2}{4} \right). \quad (6.41)$$

If we order these vectors, we get

$$\lambda_{AB}^{\gamma\downarrow} = (\gamma, 1 - \gamma, 0, 0) \quad (6.42)$$

and

$$\lambda_S^{\gamma\downarrow} = \left(\left(\frac{\gamma}{2} - 1\right)^2, \frac{\gamma}{2} \left(1 - \frac{\gamma}{2}\right), \frac{\gamma}{2} \left(1 - \frac{\gamma}{2}\right), \frac{\gamma^2}{4} \right). \quad (6.43)$$

In order for relation (6.23) to be true, the following inequalities should be true

$$\begin{cases} \gamma \geq \left(\frac{\gamma}{2} - 1\right)^2 \\ 1 \geq \left(\frac{\gamma}{2} - 1\right)^2 + \frac{\gamma}{2} \left(1 - \frac{\gamma}{2}\right) \\ 1 \geq \left(\frac{\gamma}{2} - 1\right)^2 + \gamma \left(1 - \frac{\gamma}{2}\right) \end{cases} \quad (6.44)$$

After some calculation, we obtain

$$\begin{cases} \gamma \geq 4 - 2\sqrt{3} \\ \gamma \geq 0 \end{cases} \quad (6.45)$$

The two inequalities of (6.45) are always verified in the case where $\gamma \geq 2/3$ since we always have $\gamma \geq 0$.

Second case : $\gamma < 2/3$ In this case, we have $g(\gamma) = 1/3$. The eigenvalues of ρ_{AB}^γ and ρ_S^γ are therefore respectively given by

$$\lambda_{AB}^\gamma = \left(\frac{1}{3}, \frac{\gamma}{2} + \frac{1}{3}, \frac{1}{3} - \frac{\gamma}{2}, 0 \right), \quad (6.46)$$

and

$$\lambda_S^\gamma = \left(\frac{2}{9}, \frac{2}{9}, \frac{4}{9}, \frac{1}{9} \right). \quad (6.47)$$

If we order these vectors, we get

$$\lambda_{AB}^{\gamma\downarrow} = \left(\frac{\gamma}{2} + \frac{1}{3}, \frac{1}{3}, \frac{1}{3} - \frac{\gamma}{2}, 0 \right) \quad (6.48)$$

and

$$\lambda_S^{\gamma\downarrow} = \left(\frac{4}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9} \right). \quad (6.49)$$

In order for relation (6.23) to be true, the following inequalities should be true

$$\begin{cases} \frac{\gamma}{2} + \frac{1}{3} \geq \frac{4}{9} \\ \frac{\gamma}{2} + \frac{2}{3} \geq \frac{2}{3} \end{cases} \quad (6.50)$$

After some simple calculation, we obtain

$$\begin{cases} \gamma \geq \frac{2}{9} \\ \gamma \geq 0 \end{cases} \quad (6.51)$$

Since we have $\gamma \geq 0$, the two inequalities of (6.51) are always verified in the case where

$$\gamma \geq \frac{2}{9}. \quad (6.52)$$

Consequently, relation (6.23) is verified for state ρ_{AB}^γ only when $\gamma \geq 2/9$ (where $2/9 \simeq 0.2222..$ in order to fix ideas).

Let us test proposition 8 with a new state. Take the state defined by relation (5.1), that we rewrite here as

$$\rho_{AB}^x = \begin{pmatrix} \frac{1-x}{2} & 0 & 0 & 0 \\ 0 & x|a|^2 & xab^* & 0 \\ 0 & xa^*b & x|b|^2 & 0 \\ 0 & 0 & 0 & \frac{1-x}{2} \end{pmatrix}. \quad (6.53)$$

with $|a|^2 + |b|^2 = 1$. As we already showed, the state is separable if and only if (necessary and sufficient condition since we are in the 2×2 case)

$$x > \frac{1}{1 + 2|ab|}. \quad (6.54)$$

Suppose we take $a \in \mathbb{R}, a = [0, 1]$. We therefore have

$$|b|^2 = 1 - a^2. \quad (6.55)$$

If we take $b \in \mathbb{R}$ and $b \geq 0$, then we have

$$b = \sqrt{1 - a^2}. \quad (6.56)$$

In this case, the states are separable if and only if

$$x > \frac{1}{1 + 2ab}. \quad (6.57)$$

Again, we use numerical methods to test relation (6.23) with state ρ_{AB}^x which is now defined by

$$\rho_{AB}^x = \begin{pmatrix} \frac{1-x}{2} & 0 & 0 & 0 \\ 0 & xa^2 & xab & 0 \\ 0 & xab & xb^2 & 0 \\ 0 & 0 & 0 & \frac{1-x}{2} \end{pmatrix}. \quad (6.58)$$

since $a \in \mathbb{R}$ and $b \in \mathbb{R}$. Using the Matlab program `xab_majorization.m`, we in fact see that relation (6.23) is again not verified for some states. This is another example of a state for which proposition 8 is not verified.

In fact, it is possible to explain why proposition 8 is not verified in general. In order to do that, we can use property 3. According to it, if relation (6.23) is true, then relation

$$\phi(\lambda(\rho_{AB})) \geq \phi(\lambda(\rho_S)) \quad (6.59)$$

should also be verified *for all* Schur-convex functions ϕ . Therefore, if we find at least one Schur-convex function which does not verify equation (6.59), it means that relation (6.23) is not true in general. Remember that this relation was based on an entropy inequality, the subadditivity, which we rewrite here. The relation is as follows.

$$S(A, B) \leq S(A) + S(B). \quad (6.60)$$

Notice that it is equivalent to the relation

$$S(A, B) \leq S(S). \quad (6.61)$$

where $S(S)$ is the von Neumann entropy of the tensor product of systems A and B . We already explained that $S(S) = S(A) + S(B)$ when the state of system S is given by the simple tensor product of the states of systems A and B . The von Neumann entropy is Schur-concave. Now, we know that, as we stated before, all the Rényi entropies are Schur-concave functions (which are the negative of Schur-convex functions). Therefore, if relation (6.23) is true, relation (6.61) should be verified for the negative of all the Rényi entropies. However, the it is not verified for all the Rényi entropies, as we are going to show. Take for example the Rényi entropy of order 4. It is given by

$$I_4(X) = -\frac{1}{3} \log \left(\sum_k p_k^4 \right) \quad (6.62)$$

according to relation (3.29). Let us compute the entropy of order 4 of state ρ_{AB}^γ defined by equation (6.25). Since its vector of eigenvalues is provided by

$$\lambda_{AB}^\gamma = \left(1 - 2g(\gamma), g(\gamma) + \frac{\gamma}{2}, g(\gamma) - \frac{\gamma}{2}, 0 \right), \quad (6.63)$$

its Rényi entropy of order 4 is given by

$$I_4(\rho_{AB}^\gamma) = -\frac{1}{3} \log \left[(1 - 2g(\gamma))^4 + \left(g(\gamma) + \frac{\gamma}{2} \right)^4 + \left(g(\gamma) - \frac{\gamma}{2} \right)^4 \right]. \quad (6.64)$$

The vector of eigenvalues of state ρ_S^γ is given by

$$\lambda_S^\gamma = (g(\gamma)(1 - g(\gamma)), (g(\gamma))^2 - 2g(\gamma) + 1, (g(\gamma))^2, g(\gamma)(1 - g(\gamma))). \quad (6.65)$$

and its Rényi entropy of order 4 is given by

$$I_4(\rho_S^\gamma) = -\frac{1}{3} \log \left[2(g(\gamma)(1 - g(\gamma)))^4 + ((g(\gamma))^2 - 2g(\gamma) + 1)^4 + ((g(\gamma))^2)^4 \right]. \quad (6.66)$$

In order to verify relation (6.61) with Rényi's entropy of order 4, we plot the values of $I_4(\rho_{AB}^\gamma)$ and $I_4(\rho_S^\gamma)$ using the Matlab program `subadditivity_Renyi.m`. We obtain

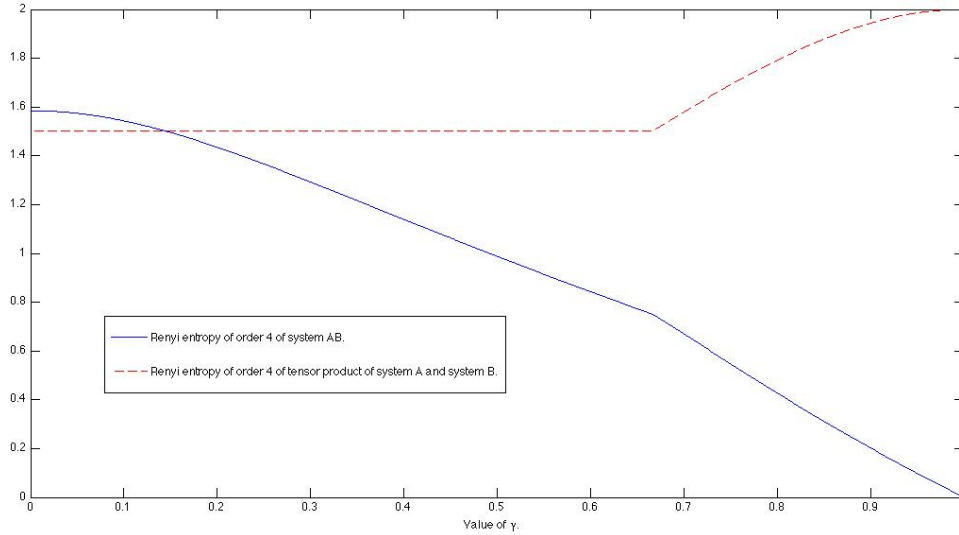


Figure 6.1: Rényi entropy of order 4 of system AB and of tensor product of systems A and B , in case of state ρ_{AB}^γ .

figure 6.1. We see that for some values of γ , $I_4(\rho_{AB}^\gamma)$ is greater than $I_4(\rho_S^\gamma)$. Therefore, we immediately conclude that relation (6.61) is not verified for the Rényi entropy of order 4. The latter being Schur-concave, we see that the relation

$$\Gamma(\lambda(\rho_{AB})) \leq \Gamma(\lambda(\rho_S)) \quad (6.67)$$

is not verified for a concave function Γ , which also means that relation

$$\phi(\lambda(\rho_{AB})) \geq \phi(\lambda(\rho_S)) \quad (6.68)$$

is not verified for a convex function ϕ . Therefore, relation (6.23) should not be true in general.

It is also possible to give an analytical proof by simply saying that the subadditivity relation is not verified for all the Rényi entropies (it is verified for the von Neumann entropy for example, as we already said, but for some others it is not). Indeed, in order to verify if relation (6.23) was true, we needed to verify that relation (6.61) was true for the all the Rényi entropies, as we said before. We also stated that relation (6.60) is equivalent to relation (6.61) for the von Neumann entropy. However, as we saw before, one of the properties of the Rényi entropy was the additivity, namely

$$I_\alpha(X, Y) = I_\alpha(X) + I_\alpha(Y), \quad (6.69)$$

for $\alpha \geq 0$, $\alpha \neq 1$. Therefore, if relation (6.61) was true for all the Rényi entropies, than it would also be the case for relation (6.60). Since the subadditivity is not a property of all the Rényi entropies, relation (6.23) is not verified in general.

6.4.2 Catalysis

What is really interesting to notice in what we did earlier is the fact that even when relation (6.23)

$$\lambda(\rho_{AB}) \succ \lambda(\rho_S) \quad (6.70)$$

is not verified, the inverse relation

$$\lambda(\rho_{AB}) \prec \lambda(\rho_S) \quad (6.71)$$

is not verified either. Take for example state ρ_{AB}^γ , for which relation (6.70) is not verified when $\gamma \geq 2/9$. In this case, in order for relation (6.71) to be verified, we should have $\gamma < 0$, which is never the case in the definition of state ρ_{AB}^γ . Therefore, we say that states ρ_{AB}^γ and ρ_S^γ are *incomparable* when $\gamma \geq 2/9$. Similarly, take state ρ_{AB}^x . Relation (6.70) is not verified (for the case in which $a \in \mathbb{R}$ and $b \in \mathbb{R}$) according to our program `xab_majorization.m`. Now, if we use the same program in order to verify relation (6.71) this time, we also find that it is not true. In fact, all the examples of states for which relation (6.70) is not verified also don't verify relation (6.71). This characteristic of relation (6.70) suggests the use of a new concept called *catalysis*.

The idea of catalysis is related to a property we gave when introducing an application of the theory of majorization. Theorem 1 states that a pure state $|\psi\rangle$ transforms to $|\phi\rangle$ using LOCC if and only if

$$\lambda_\psi \prec \lambda_\phi, \quad (6.72)$$

where λ_ψ and λ_ϕ are the vectors of eigenvalues of ρ_ψ and ρ_ϕ , respectively, with

$$\rho_\psi \equiv \text{Tr}_B (|\psi\rangle \langle\psi|) \quad (6.73)$$

and

$$\rho_\phi \equiv \text{Tr}_B (|\phi\rangle \langle\phi|). \quad (6.74)$$

Now the idea is that some of the states are incomparable, in the sense that neither relation (6.72), neither relation

$$\lambda_\psi \succ \lambda_\phi \quad (6.75)$$

is true. In this case, both transformations $|\psi\rangle \rightarrow |\phi\rangle$ and $|\phi\rangle \rightarrow |\psi\rangle$ are not possible. The idea of the catalysis is to use a third bipartite state, called catalysts, in order to make one of the transformations possible [13]. This is called entanglement-assisted local transformation. Suppose the catalyst is χ . Suppose furthermore we define states

$$\rho_\psi^\chi \equiv \text{Tr}_B (|\psi\rangle \otimes |\chi\rangle) (\langle\psi| \otimes \langle\chi|) \quad (6.76)$$

and

$$\rho_\phi^\chi \equiv \text{Tr}_B (|\phi\rangle \otimes |\chi\rangle) (\langle\phi| \otimes \langle\chi|) \quad (6.77)$$

along with their respective vectors of eigenvalues λ_ψ^χ and λ_ϕ^χ . If one of the two relations

$$\lambda_\psi^\chi \prec \lambda_\phi^\chi \quad (6.78)$$

and

$$\lambda_\psi^\chi \succ \lambda_\phi^\chi \quad (6.79)$$

is verified, then it is possible to perform LOCC using state $|\chi\rangle$, or entanglement-assisted local transformation. This application shows the practical importance of states called catalysts.

In this work, the idea is to try to use the concept of catalysis in order to see if relation (6.70) becomes verified in case we use catalyst states. However, this can only be done if it is possible to find catalysts for the states concerned. There is a property which gives some conditions required for the existence of catalysts. It is as follows [1].

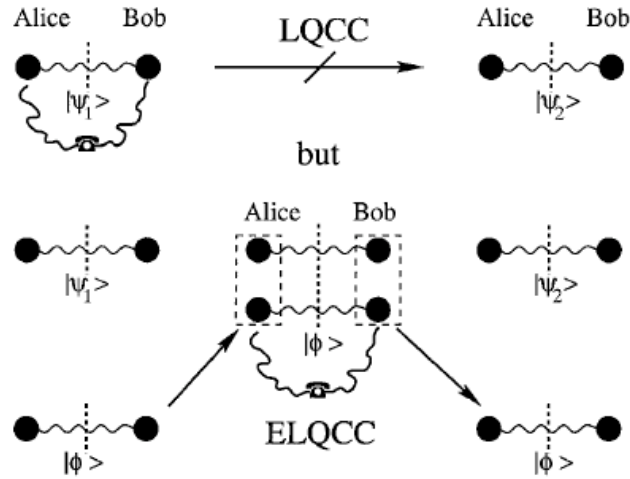


Figure 6.2: A diagram taken from Plenio's article [13], illustrating the concept of catalysis. LQCC stands for Local Quantum operations and classical communication.

Theorem 4. Suppose x and y are two vectors in \mathbb{R}^n which verify

$$x_i, y_i \geq 0 \quad i = 1, \dots, n \quad (6.80)$$

and

$$\sum_{i=1}^n x_i = 1 \quad \sum_{i=1}^n y_i = 1. \quad (6.81)$$

If there exists a vector z such that $x \otimes z \prec y \otimes z$, then

$$I_\alpha(x) \geq I_\alpha(y) \quad \forall \alpha \geq 0, \alpha \neq 1, \quad (6.82)$$

where I_α is the Rényi entropy of order α and

$$S(x) \geq S(y), \quad (6.83)$$

where S is the von Neumann entropy.

This property is usually related to more general norms, but we adapted it to the more restricted case of Rényi entropies. Our property mainly states that if we want to find catalysts states for the states that do not verify

$$\lambda(\rho_{AB}) \succ \lambda(\rho_S) \quad (6.84)$$

and

$$\lambda(\rho_{AB}) \prec \lambda(\rho_S), \quad (6.85)$$

then relation

$$I_\alpha(AB) \leq I_\alpha(S) \quad (6.86)$$

should be verified for all the Rényi entropies (for all the α). We already showed that this relation is not verified in general for all of them. However, we see from 6.1, for the entropy of order 4 for example, that for some state that don't verify both (6.84) and (6.85), relation (6.86) is still verified. Therefore, in order to find catalysts, we test relation (6.86) for all the Rényi entropies, for the incomparable states. Let us try to find

catalysts for state ρ_{AB}^γ . We saw that states ρ_{AB}^γ and ρ_S^γ are incomparable for $\gamma \geq 2/9$. The Matlab program `gamma_catalysis.m` allows us to verify this. In fact, we find that for all the incomparable states, it is not possible to find catalysts. We conclude from that fact that it is not always possible to find catalysts when relation (6.84) is not verified, which that it is not possible to find a more general property related to relation (6.84) using the concept of catalysis.

In order to conclude, what we can say about relation (6.84) is that when it is not verified, the state are incomparable, in the sense that relation (6.85) is never verified. However, it is not possible to find catalysts in general. In the next section, we are however going to see that the relation is satisfied for an important class of states, which therefore makes proposition 8 pretty interesting.

6.4.3 Bell diagonal states

A very interesting state which verifies relation (6.23) can be obtained using the class of the Bell states, introduced in section 4.2. It is the Bell diagonal state, which is, as its name suggest, diagonal in the Bell basis. It can be written as

$$\rho_{AB}^{Bell} = p_I |\phi^+\rangle \langle \phi^+| + p_x |\psi^+\rangle \langle \psi^+| + p_y |\psi^-\rangle \langle \psi^-| + p_z |\phi^-\rangle \langle \phi^-| \quad (6.87)$$

where state $|\phi^+\rangle$, $|\psi^+\rangle$, $|\psi^-\rangle$ and $|\phi^-\rangle$ are defined by equations (4.14), (4.16), (4.17) and (4.15), respectively. In matrix form, ρ_{AB}^{Bell} is given by

$$\rho_{AB}^{Bell} = \begin{pmatrix} p_I + p_z & 0 & 0 & p_I - p_z \\ 0 & p_x + p_y & p_x - p_y & 0 \\ 0 & p_x - p_y & p_x + p_y & 0 \\ p_I - p_z & 0 & 0 & p_I + p_z \end{pmatrix}. \quad (6.88)$$

Note that since ρ_{AB}^{Bell} is a state, its eigenvalues verify relation

$$p_I + p_x + p_y + p_z = 1. \quad (6.89)$$

Using this relation, we find that state ρ_A^{Bell} corresponding to subsystem A is given by

$$\rho_A^{Bell} = \text{Tr}_B [\rho_{AB}^{Bell}] = \frac{1}{2} (|0\rangle \langle 0| + |1\rangle \langle 1|). \quad (6.90)$$

Similarly, state ρ_B^{Bell} corresponding to subsystem B is given by

$$\rho_B^{Bell} = \text{Tr}_A [\rho_{AB}^{Bell}] = \frac{1}{2} (|0\rangle \langle 0| + |1\rangle \langle 1|). \quad (6.91)$$

If we compute state ρ_S^{Bell} given by the tensor product of ρ_A^{Bell} and ρ_B^{Bell} , we obtain

$$\rho_S^{Bell} = \rho_A^{Bell} \otimes \rho_B^{Bell} = \frac{1}{4} (|00\rangle \langle 00| + |01\rangle \langle 01| + |10\rangle \langle 10| + |11\rangle \langle 11|). \quad (6.92)$$

Its matrix can be represented by $\rho_S^{Bell} = (1/4)\mathcal{I}$. Therefore, whatever the values of p_I , p_x , p_y and p_z , ρ_S^{Bell} is given by the uniform probability distribution on 4 elements. Consequently, using proposition 1, we see that ρ_S^{Bell} is majorized by any $2 \otimes 2$ bipartite state. Relation (6.23) is thus verified for any Bell diagonal state. This is interesting since the Bell diagonal states are often encountered in the field of quantum information.

6.5 Conclusion

In this section, we have been studying majorization relations based on already existing entropic inequalities. We began by talking about Nielsen and Kempe's result in section 6.1 in order to show how they came to introduce it by basing themselves on an already existing separability criterion based on an entropic inequality. We then proposed in section 6.2 a majorization relation based on the positivity of the joint entropy of a separable state. The majorization relation turned out to be trivial, since the comparison with the vector of eigenvalues of a pure state is trivial in general. In section 6.3, we introduced one first adaptation of the subadditivity relation, which we finally rejected before exhibiting a second adaptation of it in terms of majorization relation in section 6.4. The last relation turned out to be the best adaptation for the subadditivity relation, but was still not verified in general, as we showed using examples. We then tried using the concept of catalysis in order to find further results, but realized that it was not possible. A very interesting result was however the fact that the majorization relation was verified for all the Bell diagonal states, a much used class of states in the field of quantum information. This could have interesting application when studying measures of disorder in Bell diagonal states.

In the next section, we are going to study majorization in the field of Gaussian quantum information. Gaussian states being usually the one used in practice when studying quantum information using quantum optics, studying them in terms of majorization should lead to some interesting results.

7 Gaussian Quantum Information

Until now, all the states we worked with are part of the field of discrete quantum information. Indeed, these states evolve in finite-dimensional Hilbert spaces. This is for example the case for the well-known quantum bit, or qubit, which we already introduced in the previous sections. The qubit is a quantum system with two distinguishable states, which can be for example denoted by the two numbers 0 and 1. Indeed, the two vectors $|0\rangle$ and $|1\rangle$ form a basis for the 2-dimensional Hilbert Space in which the qubit belongs. Any state ψ can be written as a superposition of these two vectors, namely

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle, \quad (7.1)$$

where α and β are two complex numbers which verify the condition

$$|\alpha|^2 + |\beta|^2 = 1. \quad (7.2)$$

On the other hand, the field of continuous quantum information treats systems which are modelled by states evolving in infinite-dimensional Hilbert spaces, described by observables with continuous spectra. [25] A very important example of a continuous-variable system is given by the modes of the electromagnetic field, which are bosonic modes, since the particle of light, the photon, is a boson. Each of these modes can be represented by a quantum harmonic oscillator. In the following section, we are going to introduce the mathematical model which will allow us to study these electromagnetic modes, as well as the states representing them.

7.1 Continuous Variables model

7.1.1 The quantum harmonic oscillator model

We already said that the mathematical model describing the modes of the electromagnetic field is based on the model of the quantum harmonic oscillator. Suppose we need to describe N bosonic modes. This means that we have N harmonic oscillators. The N modes are associated with a Hilbert space $\mathcal{H}^{\otimes N}$, which is the result of the tensor product of N Hilbert spaces, each one of them corresponding to one mode, that is

$$\mathcal{H}^{\otimes N} = \bigotimes_{k=1}^N \mathcal{H}_k \quad (7.3)$$

where each \mathcal{H}_k corresponds to a harmonic oscillator. To each one of them also corresponds a pair of bosonic field operators, \hat{a} , called annihilation operator, and \hat{a}^\dagger , called creation operator. We therefore have N pair of operators, which can be arranged in a vectorial operator

$$\hat{\mathbf{b}} := \left(\hat{a}_1, \hat{a}_1^\dagger, \dots, \hat{a}_N, \hat{a}_N^\dagger \right)^\top. \quad (7.4)$$

Every bosonic field operator must satisfy some bosonic commutation relations. These relations can be regrouped in the relation

$$\left[\hat{b}_i, \hat{b}_j \right] = \Omega_{ij}, \quad (7.5)$$

where

$$\Omega := \bigoplus_{k=1}^N \omega = \begin{pmatrix} \omega & & \\ & \ddots & \\ & & \omega \end{pmatrix}, \quad \omega := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (7.6)$$

Furthermore, each Hilbert space \mathcal{H}_k is spanned by a countable basis $\{|n\rangle\}_{n=0}^\infty$, called the Fock basis. The Fock states, or number states, $|n\rangle$ are in fact eigenstates of the number operator $\hat{n} := \hat{a}^\dagger \hat{a}$, which therefore satisfies the relation

$$\hat{n} |n\rangle = n |n\rangle, \quad (7.7)$$

where n (the eigenvalues) is a non-negative real number. The actions of the bosonic field operators over the number states can be found using the bosonic commutation relation, and are given by

$$\hat{a} |n\rangle = \sqrt{n} |n-1\rangle \quad n > 0, \quad (7.8)$$

and

$$\hat{a}^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle \quad n \geq 0. \quad (7.9)$$

The bosonic systems can be equivalently described by other field operators, which are called quadrature field operators $\{\hat{q}_k, \hat{p}_k\}_{k=1}^N$, and can be arranged in the vector

$$\hat{\mathbf{x}} := (\hat{q}_1, \hat{p}_1, \dots, \hat{q}_N, \hat{p}_N)^\top. \quad (7.10)$$

The quadrature field operators are related to the bosonic field operators through the relations

$$\begin{cases} \hat{q}_j = \hat{a}_j + \hat{a}_j^\dagger \\ \hat{p}_j = i(\hat{a}_j^\dagger - \hat{a}_j) \end{cases} \quad j = 1, \dots, N \quad (7.11)$$

and satisfy the canonical commutation relations

$$[\hat{x}_i, \hat{x}_j] = 2i\Omega_{ij}. \quad (7.12)$$

Note that here, the relations are given in natural units ($\hbar = 2$).

7.1.2 The Gaussian state

Like in the case of the field of discrete-variables, every N -mode bosonic system is represented by a quantum state ρ , which is still a trace-one positive operator. In the case of the field of continuous-variables, the density operator has an equivalent representation in terms of a quasiprobability distribution. In order to see this, let us introduce the Weyl operator defined as

$$D(\xi) := \exp(i\hat{\mathbf{x}}^T \boldsymbol{\Omega} \xi), \quad (7.13)$$

where we define the elements of $\xi \in \mathbb{R}^{2N}$ as

$$\xi := (\xi_1, \dots, \xi_{2N})^T. \quad (7.14)$$

Instead of using the density matrix ρ , we can use an equivalent representation given by the Wigner characteristic function

$$\chi(\xi) = \text{Tr}[\rho D(\xi)], \quad (7.15)$$

as well as a representation given by a Wigner function

$$W(\mathbf{x}) = \int_{\mathcal{R}^{2N}} \frac{d^{2N}\xi}{(2\pi)^{2N}} \exp(-i\mathbf{x}^T \boldsymbol{\Omega} \xi), \chi(\xi) \quad (7.16)$$

which is normalized to 1, and which is in fact related to the Wigner characteristic function by a Fourier transform. The Wigner function is nonpositive in general, that is why it is called quasiprobability distribution. Note that the elements of vector $\mathbf{x} \in \mathbb{R}^{2N}$ are the eigenvalues of quadrature operators $\hat{\mathbf{x}}$.

The most relevant quantities which characterize the Wigner representations (χ or W) are the statistical moments of the quantum state [25]. The first moment is the mean value, defined by

$$\bar{\mathbf{x}} := \langle \hat{\mathbf{x}} \rangle = \text{Tr}(\hat{\mathbf{x}}\rho), \quad (7.17)$$

and the second moment is the covariance matrix \mathbf{V} , whose elements are defined by

$$V_{ij} := \frac{1}{2} \langle \{\Delta \hat{x}_i, \Delta \hat{x}_j\} \rangle \quad (7.18)$$

where $\Delta \hat{x}_i := \hat{x}_i - \langle \hat{x}_i \rangle$. In fact, these two first moments are sufficient to characterize a particular class of states in which we are interested here, the Gaussian states. These are bosonic states whose Wigner representation (χ or W) is Gaussian, namely

$$\chi(\xi) = \exp\left[-\frac{1}{2}\xi^T (\boldsymbol{\Omega} \mathbf{V} \boldsymbol{\Omega}^T) \xi - i(\boldsymbol{\Omega} \bar{\mathbf{x}})^T \xi\right], \quad (7.19)$$

and

$$W(\mathbf{x}) = \frac{\exp\left[-(1/2)(\mathbf{x} - \bar{\mathbf{x}})^T \mathbf{V}^{-1}(\mathbf{x} - \bar{\mathbf{x}})\right]}{(2\pi)^N \sqrt{\det \mathbf{V}}}. \quad (7.20)$$

Gaussian states are of fundamental importance in the field of continuous-variable quantum information. They are indeed the primary tools for analysing continuous-variable quantum information processing, along with Gaussian transformations. Gaussian transformations are a type of quantum unitary transformations, which are themselves a type of quantum operations, also called quantum channels. A quantum operation is a transformation undergone by a quantum state [25]. It is a linear map $\mathcal{L} : \rho \rightarrow \mathcal{L}(\rho)$ which is completely positive and trace decreasing, namely

$$0 \leq \text{Tr} [\mathcal{L}(\rho)] \leq 1. \quad (7.21)$$

Some particular cases of quantum channels are the ones represented by unitary representations $U^{-1} = U^\dagger$. These transformations are reversible, and transform a state according to

$$\rho \rightarrow U\rho U^\dagger. \quad (7.22)$$

Now, a quantum operation is Gaussian when it transforms Gaussian states into Gaussian states. Gaussian states and Gaussian transformations are central in the field of continuous-variable information because they are of great practical importance. States used in quantum information experiments like the thermal state, for example, are usually Gaussian states. This is also the case for the operations used in these experiments, like most of the nonlinear operations, which can be well approximated by Gaussian transformations.

Consider a general state of one mode ($N = 1$) of the form

$$\rho = \sum_{k=0}^{\infty} c_k |k\rangle \langle k|. \quad (7.23)$$

This is a state which has a diagonal representation in the Fock basis. We will however see later that this state, for some particular values of c_k , can represent any Gaussian state of one mode. Like any density matrix, its trace should be equal to 1. The trace of ρ is given by

$$\text{Tr} [\rho] = \text{Tr} \left[\sum_{k=0}^{\infty} c_k |k\rangle \langle k| \right] = \sum_{n=0}^{\infty} \langle n| \left[\sum_{k=0}^{\infty} c_k |k\rangle \langle k| \right] |n\rangle = \sum_{k=0}^{\infty} c_k$$

This results in

$$\sum_{k=0}^{\infty} c_k = 1. \quad (7.24)$$

We will be using this condition in the following. Let us now compute the first two statistical moments of the quantum states. The first one, the mean value, is given by the vector

$$\bar{\mathbf{x}} := \langle \hat{\mathbf{x}} \rangle = \text{Tr}(\hat{\mathbf{x}}\rho), \quad (7.25)$$

which reduces to the vector

$$\bar{\mathbf{x}} = (\langle \hat{q} \rangle, \langle \hat{p} \rangle) \quad (7.26)$$

when $N = 1$. First, notice that the mean value of any operator \hat{A} is given by

$$\begin{aligned}\langle \hat{A} \rangle &= \text{Tr} [\hat{A} \rho] \\ &= \text{Tr} \left[\hat{A} \sum_{k=0}^{\infty} c_k |k\rangle \langle k| \right] \\ &= \sum_{n=0}^{\infty} \langle n| \left[\hat{A} \sum_{k=0}^{\infty} c_k |k\rangle \langle k| \right] |n\rangle \\ &= \sum_{k=0}^{\infty} c_k \langle k| \hat{A} |k\rangle\end{aligned}$$

The mean value of the quadrature operator \hat{q} is therefore

$$\langle \hat{q} \rangle = \sum_{k=0}^{\infty} c_k \langle k| \hat{q} |k\rangle$$

where the matrix element $\langle k| \hat{q} |k\rangle$ is equal to

$$\begin{aligned}\langle k| \hat{q} |k\rangle &= \langle k| (\hat{a} + \hat{a}^\dagger) |k\rangle \\ &= \langle k| \hat{a} |k\rangle + \langle k| \hat{a}^\dagger |k\rangle \\ &= \sqrt{k} \langle k| k-1\rangle + \sqrt{k+1} \langle k| k+1\rangle \\ &= 0\end{aligned}$$

resulting in

$$\langle \hat{q} \rangle = 0. \quad (7.27)$$

Similarly, the mean value of the quadrature operator \hat{p} is given by

$$\langle \hat{p} \rangle = 0. \quad (7.28)$$

The first statistical moment of the quantum state ρ is therefore

$$\bar{\mathbf{x}} = 0. \quad (7.29)$$

Let us now compute the second statistical moment of quantum state ρ , the covariance matrix, which arbitrary elements are

$$V_{ij} = \frac{1}{2} \langle \{\Delta \hat{x}_i, \Delta \hat{x}_j\} \rangle = \frac{1}{2} \langle \{\hat{x}_i - \langle \hat{x}_i \rangle, \hat{x}_j - \langle \hat{x}_j \rangle\} \rangle \quad (7.30)$$

The first element V_{11} is

$$\begin{aligned}V_{11} &= \frac{1}{2} \langle \{\hat{x}_1 - \langle \hat{x}_1 \rangle, \hat{x}_1 - \langle \hat{x}_1 \rangle\} \rangle \\ &= \langle (\hat{x}_1 - \langle \hat{x}_1 \rangle)^2 \rangle \\ &= \langle \hat{x}_1^2 \rangle - \langle \hat{x}_1 \rangle^2 \\ &= \langle \hat{q}^2 \rangle - \langle \hat{q} \rangle^2\end{aligned}$$

where

$$\langle \hat{q}^2 \rangle = \sum_{k=0}^{\infty} c_k \langle k| \hat{q}^2 |k\rangle. \quad (7.31)$$

If we compute the mean value of operator \hat{q}^2 , we obtain

$$\begin{aligned}\langle k | \hat{q}^2 | k \rangle &= \langle k | (\hat{a} + \hat{a}^\dagger)^2 | k \rangle \\ &= \langle k | \hat{a}^2 | k \rangle + \langle k | \hat{a} \hat{a}^\dagger | k \rangle + \langle k | \hat{a}^\dagger \hat{a} | k \rangle + \langle k | \hat{a}^{\dagger 2} | k \rangle\end{aligned}\quad (7.32)$$

Let us now calculate the diagonal matrix elements of operators \hat{a}^2 , $\hat{a} \hat{a}^\dagger$, $\hat{a}^\dagger \hat{a}$ and $\hat{a}^{\dagger 2}$. We respectively have

$$\langle k | \hat{a}^2 | k \rangle = \sqrt{k} \sqrt{k-1} \langle k | k-2 \rangle = 0 \quad (7.33)$$

and

$$\langle k | \hat{a}^{\dagger 2} | k \rangle = \sqrt{k+1} \sqrt{k+2} \langle k | k+2 \rangle = 0 \quad (7.34)$$

for \hat{a}^2 and $\hat{a}^{\dagger 2}$, but we obtain non-zero elements for the other two operators, given by

$$\langle k | \hat{a} \hat{a}^\dagger | k \rangle = \langle k | (1 + \hat{n}) | k \rangle = 1 + k \quad (7.35)$$

and

$$\langle k | \hat{a}^\dagger \hat{a} | k \rangle = \langle k | \hat{n} | k \rangle = k. \quad (7.36)$$

By plugging these results in equations (7.32) and (7.31), we obtain

$$\langle \hat{q}^2 \rangle = \sum_{k=0}^{\infty} c_k (1 + 2k). \quad (7.37)$$

Since the mean value of operator \hat{q} is null according to equation (7.27), the first element of the covariance matrix \mathbf{V} is given by

$$V_{11} = \sum_{k=0}^{\infty} c_k (1 + 2k) = \sum_{k=0}^{\infty} c_k + 2 \sum_{k=0}^{\infty} k c_k = 1 + 2 \sum_{k=0}^{\infty} k c_k$$

where we used equation (7.24). Similarly, the second diagonal element of covariance matrix \mathbf{V} is given by

$$V_{22} = \langle \hat{x}_2^2 \rangle - \langle \hat{x}_2 \rangle^2 = \langle \hat{p}^2 \rangle - \langle \hat{p} \rangle^2$$

where the mean value of operator \hat{p}^2 is

$$\langle \hat{p}^2 \rangle = \sum_{k=0}^{\infty} c_k \langle k | \hat{p}^2 | k \rangle. \quad (7.38)$$

The diagonal element $\langle k | \hat{p}^2 | k \rangle$ can be computed using again equations (7.35), (7.33), (7.36) and (7.34), resulting in

$$\begin{aligned}\langle k | \hat{p}^2 | k \rangle &= \langle k | (i [\hat{a}^\dagger - \hat{a}])^2 | k \rangle \\ &= -\langle k | \hat{a}^2 | k \rangle + \langle k | \hat{a} \hat{a}^\dagger | k \rangle + \langle k | \hat{a}^\dagger \hat{a} | k \rangle - \langle k | \hat{a}^{\dagger 2} | k \rangle \\ &= (1 + 2k)\end{aligned}$$

If we use condition (7.24), and knowing that the mean value of operator \hat{p} is null according to equation (7.28), we obtain for the second diagonal element of the covariance matrix \mathbf{V}

$$V_{22} = \sum_{k=0}^{\infty} c_k (1 + 2k) = 1 + 2 \sum_{k=0}^{\infty} k c_k.$$

Here, we see that $V_{11} = V_{22}$. We are now going to compute the nondiagonal elements of the covariance matrix \mathbf{V} . The first one is given by

$$\begin{aligned} V_{12} &= \frac{1}{2} \langle \{\hat{x}_1 - \langle \hat{x}_1 \rangle, \hat{x}_2 - \langle \hat{x}_2 \rangle\} \rangle \\ &= \frac{1}{2} \langle \{\hat{q} - \langle \hat{q} \rangle, \hat{p} - \langle \hat{p} \rangle\} \rangle \\ &= \frac{1}{2} [\langle \hat{q}\hat{p} \rangle + \langle \hat{p}\hat{q} \rangle - 2 \langle \hat{p} \rangle \langle \hat{q} \rangle] \end{aligned} \quad (7.39)$$

We begin by computing the mean value of operator $\hat{q}\hat{p}$

$$\langle \hat{q}\hat{p} \rangle = \sum_{k=0}^{\infty} c_k \langle k | \hat{q}\hat{p} | k \rangle. \quad (7.40)$$

The diagonal element $\langle k | \hat{q}\hat{p} | k \rangle$ is given by

$$\begin{aligned} \langle k | \hat{q}\hat{p} | k \rangle &= \langle k | (\hat{a} + \hat{a}^\dagger) i (\hat{a}^\dagger - \hat{a}) | k \rangle \\ &= i \langle k | (\hat{a}\hat{a}^\dagger - \hat{a}^2 + \hat{a}^{\dagger 2} - \hat{a}^\dagger\hat{a}) | k \rangle \\ &= i \end{aligned}$$

This results in

$$\langle \hat{q}\hat{p} \rangle = \sum_{k=0}^{\infty} i c_k = i. \quad (7.41)$$

In order to compute the mean value of operator $\hat{p}\hat{q}$, we use the canonical commutation relation (7.12), which reduces in our case to

$$\begin{aligned} [\hat{q}, \hat{p}] &= 2i \\ \Rightarrow \hat{q}\hat{p} - \hat{p}\hat{q} &= 2i \\ \Rightarrow \hat{p}\hat{q} &= \hat{q}\hat{p} - 2i \end{aligned}$$

resulting in

$$\langle k | \hat{p}\hat{q} | k \rangle = \langle k | (\hat{q}\hat{p} - 2i) | k \rangle = -i. \quad (7.42)$$

We consequently have

$$\langle \hat{p}\hat{q} \rangle = -i, \quad (7.43)$$

and since the mean values of \hat{p} and \hat{q} are null according to (7.28) and (7.27), we obtain, using equation (7.39),

$$V_{12} = 0. \quad (7.44)$$

Similarly, if we use relations (7.41), (7.43), and the fact that the mean values of \hat{p} and \hat{q} are null, we obtain the value of the second nondiagonal matrix element of covariance matrix \mathbf{V}

$$V_{21} = 0. \quad (7.45)$$

The covariance matrix of state ρ is then given by the diagonal matrix

$$\mathbf{V} = \begin{pmatrix} v & 0 \\ 0 & v \end{pmatrix}, \quad (7.46)$$

where

$$v = \left(1 + 2 \sum_{k=0}^{\infty} k c_k \right). \quad (7.47)$$

This expression will be very useful when computing the covariance matrix of a thermal state, which we are now going to introduce.

7.2 Thermal states

7.2.1 Definition and properties

In the previous section, we introduced the Gaussian state, which is of fundamental importance in the field of continuous-variable quantum information. Here, we introduce a particular state, called the thermal state, which, as we are going to point out, is enough to represent all the Gaussian states. Before showing this, let us introduce the thermal state, along with some of its properties. By definition, a thermal state is a bosonic state which maximizes the von Neumann entropy for fixed energy

$$\langle n \rangle = \text{Tr} [\rho \hat{a}^\dagger \hat{a}] \quad (7.48)$$

where $\langle n \rangle$ is the mean number of photons in the bosonic mode [25]. The thermal state is explicitly given by

$$\rho^{\text{th}}(\bar{n}) = \sum_{n=0}^{\infty} \frac{\bar{n}^n}{(\bar{n} + 1)^{n+1}} |n\rangle \langle n| \quad (7.49)$$

where \bar{n} is the mean number of photon in the mode,

$$\bar{n} = \langle n \rangle. \quad (7.50)$$

Let us begin by computing its trace, in order to make sure is it indeed equal to 1. We have

$$\begin{aligned} \text{Tr}(\rho^{\text{th}}) &= \text{Tr} \sum_{n=0}^{\infty} \frac{\bar{n}^n}{(\bar{n} + 1)^{n+1}} |n\rangle \langle n| \\ &= \sum_{k=0}^{\infty} \langle k| \left(\sum_{n=0}^{\infty} \frac{\bar{n}^n}{(\bar{n} + 1)^{n+1}} |n\rangle \langle n| \right) |k\rangle \\ &= \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \frac{\bar{n}^n}{(\bar{n} + 1)^{n+1}} \langle k| n\rangle \langle n| k\rangle \\ &= \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \frac{\bar{n}^n}{(\bar{n} + 1)^{n+1}} \\ &= \frac{1}{\bar{n} + 1} \sum_{n=0}^{\infty} \left(\frac{\bar{n}}{\bar{n} + 1} \right)^n \end{aligned}$$

When $\bar{n} = 0$, one easily sees that

$$\text{Tr}(\rho^{\text{th}}) = 1. \quad (7.51)$$

In order to see what happens when $\bar{n} \neq 1$, we first need to compute the following series:

$$\begin{aligned} \sum_{n=0}^{\infty} q^n &= \lim_{m \rightarrow \infty} \sum_{n=0}^m q^n && |q| < 1 \\ &= \lim_{m \rightarrow \infty} \frac{1 - q^{m+1}}{1 - q} && |q| < 1 \\ &= \frac{1}{1 - q} && |q| < 1 \end{aligned} \quad (7.52)$$

If $\bar{n} > 0$, we have

$$0 < \frac{\bar{n}}{\bar{n} + 1} < 1 \quad \Rightarrow \quad \left| \frac{\bar{n}}{\bar{n} + 1} \right| < 1$$

which leads to, using result (7.52),

$$\sum_{n=0}^{\infty} \left(\frac{\bar{n}}{\bar{n} + 1} \right)^n = \left(1 - \frac{\bar{n}}{\bar{n} + 1} \right)^{-1} = \bar{n} + 1$$

Consequently, we have, for $\bar{n} > 0$,

$$\text{Tr}(\rho^{\text{th}}) = \frac{1}{\bar{n} + 1}(\bar{n} + 1) = 1 \quad (7.53)$$

We therefore see that the trace over the thermal state is indeed equal to 1, which assures us that it is indeed a state.

Let us now compute the covariance matrix of the thermal state. Using equations (7.46) and (7.47), we see that it is given by

$$\mathbf{V}^{\text{th}} = \begin{pmatrix} v^{\text{th}} & 0 \\ 0 & v^{\text{th}} \end{pmatrix} \quad (7.54)$$

where

$$v^{\text{th}} = 1 + 2 \sum_{k=0}^{\infty} k \frac{\bar{n}^k}{(\bar{n} + 1)^{k+1}}$$

$$v^{\text{th}} = 1 + \frac{2}{\bar{n} + 1} \sum_{k=0}^{\infty} k \frac{\bar{n}^k}{(\bar{n} + 1)^k}$$

In order to compute it, we need the following mathematical result:

$$\sum_{n=0}^{\infty} nq^n = \frac{q}{(q - 1)^2} \quad |q| < 1 \quad (7.55)$$

This leads easily to

$$v^{\text{th}} = 1 + 2\bar{n}. \quad (7.56)$$

The covariance matrix \mathbf{V}^{th} of the thermal state is therefore given by the diagonal matrix

$$\mathbf{V}^{\text{th}} = (1 + 2\bar{n})\mathcal{I}. \quad (7.57)$$

We are now going to compute the von Neumann entropy of the thermal state. Using the expression of the von Neumann entropy, we obtain

$$\begin{aligned} S(\rho^{\text{th}}) &= -\text{Tr}(\rho^{\text{th}} \log \rho^{\text{th}}) \\ &= -\text{Tr} \left(\left[\sum_{n=0}^{\infty} \frac{\bar{n}^n}{(\bar{n} + 1)^{n+1}} |n\rangle \langle n| \right] \log \left[\sum_{n=0}^{\infty} \frac{\bar{n}^n}{(\bar{n} + 1)^{n+1}} |n\rangle \langle n| \right] \right) \\ &= -\sum_{k=0}^{\infty} \langle k| \left(\left[\sum_{n=0}^{\infty} \frac{\bar{n}^n}{(\bar{n} + 1)^{n+1}} |n\rangle \langle n| \right] \log \left[\sum_{n=0}^{\infty} \frac{\bar{n}^n}{(\bar{n} + 1)^{n+1}} |n\rangle \langle n| \right] \right) |k\rangle \\ &= -\sum_{n=0}^{\infty} \frac{\bar{n}^n}{(\bar{n} + 1)^{n+1}} \log \left(\frac{\bar{n}^n}{(\bar{n} + 1)^{n+1}} \right) \end{aligned}$$

If we set $q = \frac{\bar{n}}{\bar{n}+1}$, we can rewrite the last expression as

$$\begin{aligned}
 S(\rho^{\text{th}}) &= - \sum_{n=0}^{\infty} \frac{q^n}{\bar{n}+1} \log \left(\frac{q^n}{\bar{n}+1} \right) \\
 &= - \frac{1}{\bar{n}+1} \sum_{n=0}^{\infty} q^n [\log(q^n) - \log(\bar{n}+1)] \\
 &= - \frac{1}{\bar{n}+1} \sum_{n=0}^{\infty} q^n \log(q^n) + \frac{1}{\bar{n}+1} \sum_{n=0}^{\infty} q^n \log(\bar{n}+1) \\
 &= - \frac{1}{\bar{n}+1} \sum_{n=0}^{\infty} q^n n \log(q) + \frac{1}{\bar{n}+1} \sum_{n=0}^{\infty} q^n \log(\bar{n}+1) \\
 &= - \frac{\log(q)}{\bar{n}+1} \sum_{n=0}^{\infty} n q^n + \frac{\log(\bar{n}+1)}{\bar{n}+1} \sum_{n=0}^{\infty} q^n
 \end{aligned}$$

If we use results (7.52) and (7.55) again, we obtain

$$S(\rho^{\text{th}}) = - \frac{\log(q)}{\bar{n}+1} \frac{q}{(q-1)^2} + \frac{\log(\bar{n}+1)}{\bar{n}+1} \frac{1}{1-q}$$

This leads to the expression of the von Neumann entropy of thermal state ρ^{th} with mean number of photon \bar{n}

$$S(\rho^{\text{th}}) = (\bar{n}+1) \log(\bar{n}+1) - \bar{n} \log \bar{n}. \quad (7.58)$$

This expression will be useful when stating some property of the thermal state related to its entropy.

7.2.2 Thermal decomposition of Gaussian state

A very powerful characteristic of the thermal state is the fact that it can be used to represent any Gaussian states of N modes using some unitary transformations, as we are going to see. In order to give this representation, we first need to give the definition of a symplectic transformation. A symplectic transformation is characterized by a symplectic matrix \mathbf{S} , namely a matrix \mathbf{S} which satisfies the relation [25]

$$\mathbf{S} \mathbf{\Omega} \mathbf{S}^T = \mathbf{\Omega}, \quad (7.59)$$

where $\mathbf{\Omega}$ is given by relation (7.6). Now suppose we have an arbitrary N -mode covariance matrix \mathbf{V} , then there exists a symplectic matrix \mathbf{S} which transforms \mathbf{V} in the phase space in the way [25]

$$\mathbf{V} = \mathbf{S} \mathbf{V}^{\oplus} \mathbf{S}^T \quad (7.60)$$

where \mathbf{V}^{\oplus} is a diagonal matrix defined by

$$\mathbf{V}^{\oplus} = \bigoplus_{k=1}^N \nu_k \mathcal{I} \quad (7.61)$$

where the N positive quantities ν_k are called the symplectic eigenvalues of \mathbf{V} . Using this transformation, one can write the general expression of a Gaussian state in terms of thermal states. Indeed, the following proposition is true [25].

Proposition 9. Any Gaussian state $\rho(\bar{\mathbf{x}}, \mathbf{V})$, where $\bar{\mathbf{x}}$ and \mathbf{V} are its first and second statistical moments, respectively, can be written in its thermal decomposition

$$\rho(\bar{\mathbf{x}}, \mathbf{V}) = D(\bar{\mathbf{x}})U_{\mathbf{S}} [\rho(\mathbf{0}, \mathbf{V}^{\oplus})] U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \quad (7.62)$$

with

$$\rho(\mathbf{0}, \mathbf{V}^{\oplus}) = \bigotimes_{k=1}^N \rho^{th} \left(\frac{\nu_k - 1}{2} \right) \quad (7.63)$$

where the thermal state $\rho^{th}(\bar{n})$ is defined as

$$\rho^{th}(\bar{n}) = \sum_{n=0}^{\infty} \frac{\bar{n}^n}{(\bar{n} + 1)^{n+1}} |n\rangle \langle n|, \quad (7.64)$$

the displacement operator $D(x)$ is defined as

$$D(x) = \exp(x\hat{a}^{\dagger} - x^*\hat{a}) \quad (7.65)$$

and $U_{\mathbf{S}}$ is some unitary matrix.

We notice that transformation (7.60) of the covariance matrix \mathbf{V} in the phase space corresponds to a certain transformation (7.62) of the state $\rho(\bar{\mathbf{x}}, \mathbf{V})$ in the corresponding Hilbert space. We consequently from proposition 9 that any Gaussian state can be studied using a tensor product of thermal states, if we know its symplectic values.

Let us compute the von Neumann entropy of such a Gaussian state, using its thermal decomposition. We have

$$\begin{aligned} S[\rho(\bar{\mathbf{x}}, \mathbf{V})] &= -\text{Tr} [\rho(\bar{\mathbf{x}}, \mathbf{V}) \log \rho(\bar{\mathbf{x}}, \mathbf{V})] \\ &= -\text{Tr} \left[D(\bar{\mathbf{x}})U_{\mathbf{S}} [\rho(\mathbf{0}, \mathbf{V}^{\oplus})] U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \log \left\{ D(\bar{\mathbf{x}})U_{\mathbf{S}} [\rho(\mathbf{0}, \mathbf{V}^{\oplus})] U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right\} \right] \end{aligned}$$

We can make a Taylor decomposition of the logarithm function of the form

$$\log(\rho) = -\sum_{n=1}^{\infty} \frac{(\mathcal{I} - \rho)^n}{n} \quad |\rho| < 1 \quad (7.66)$$

This results in

$$S[\rho(\bar{\mathbf{x}}, \mathbf{V})] = \text{Tr} \left[D(\bar{\mathbf{x}})U_{\mathbf{S}} [\rho(\mathbf{0}, \mathbf{V}^{\oplus})] U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \sum_{n=1}^{\infty} \frac{\left(\mathcal{I} - D(\bar{\mathbf{x}})U_{\mathbf{S}} [\rho(\mathbf{0}, \mathbf{V}^{\oplus})] U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right)^n}{n} \right]$$

In the following, we will simply write ρ^{\oplus} instead of $\rho(\mathbf{0}, \mathbf{V}^{\oplus})$ for simplicity. We therefore have

$$\begin{aligned} S[\rho(\bar{\mathbf{x}}, \mathbf{V})] &= \text{Tr} \left[D(\bar{\mathbf{x}})U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \sum_{n=1}^{\infty} \frac{\left(\mathcal{I} - D(\bar{\mathbf{x}})U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right)^n}{n} \right] \\ &= \sum_{n=1}^{\infty} \frac{1}{n} \text{Tr} \left[D(\bar{\mathbf{x}})U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \left(\mathcal{I} - D(\bar{\mathbf{x}})U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right)^n \right] \end{aligned}$$

After expanding the sum, we obtain

$$\begin{aligned}
 S[\rho(\bar{\mathbf{x}}, \mathbf{V})] &= \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \left(\mathcal{I} - D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right) \right] \\
 &\quad + \frac{1}{2} \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \left(\mathcal{I} - D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right)^2 \right] \\
 &\quad + \frac{1}{3} \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \left(\mathcal{I} - D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right)^3 \right] \\
 &\quad + \dots
 \end{aligned}$$

Using the Binomial theorem

$$(1-x)^n = \sum_{k=0}^n (-1)^k \binom{n}{k} x^k, \quad (7.67)$$

we end up with

$$\begin{aligned}
 S[\rho(\bar{\mathbf{x}}, \mathbf{V})] &= \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \left(\mathcal{I} - D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right) \right] \\
 &\quad + \frac{1}{2} \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \sum_{k=0}^2 (-1)^k \binom{2}{k} \left(D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right)^k \right] \\
 &\quad + \frac{1}{3} \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \sum_{k=0}^3 (-1)^k \binom{3}{k} \left(D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right)^k \right] \\
 &\quad + \dots \\
 S[\rho(\bar{\mathbf{x}}, \mathbf{V})] &= \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} - \left(D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right)^2 \right] \\
 &\quad + \frac{1}{2} \text{Tr} \left[\sum_{k=0}^2 (-1)^k \binom{2}{k} \left(D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right)^{k+1} \right] \\
 &\quad + \frac{1}{3} \text{Tr} \left[\sum_{k=0}^3 (-1)^k \binom{3}{k} \left(D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right)^{k+1} \right] \\
 &\quad + \dots
 \end{aligned}$$

Moreover, it is easily seen that

$$\left(D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right)^k = \prod_{i=1}^k D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} = D(\bar{\mathbf{x}}) U_{\mathbf{S}} [\rho^{\oplus}]^k U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \quad (7.68)$$

We consequently have

$$\begin{aligned}
 S[\rho(\bar{\mathbf{x}}, \mathbf{V})] &= \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} - D(\bar{\mathbf{x}}) U_{\mathbf{S}} [\rho^{\oplus}]^2 U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right] \\
 &\quad + \frac{1}{2} \text{Tr} \left[\sum_{k=0}^2 (-1)^k \binom{2}{k} \left(D(\bar{\mathbf{x}}) U_{\mathbf{S}} [\rho^{\oplus}]^{k+1} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right) \right] \\
 &\quad + \frac{1}{3} \text{Tr} \left[\sum_{k=0}^3 (-1)^k \binom{3}{k} \left(D(\bar{\mathbf{x}}) U_{\mathbf{S}} [\rho^{\oplus}]^{k+1} U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right) \right] \\
 &\quad + \dots
 \end{aligned}$$

$$\begin{aligned}
S[\rho(\bar{\mathbf{x}}, \mathbf{V})] &= \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \left(\rho^{\oplus} - [\rho^{\oplus}]^2 \right) U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right] \\
&\quad + \frac{1}{2} \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \left(\sum_{k=0}^2 (-1)^k \binom{2}{k} [\rho^{\oplus}]^{k+1} \right) U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right] \\
&\quad + \frac{1}{3} \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \left(\sum_{k=0}^3 (-1)^k \binom{3}{k} [\rho^{\oplus}]^{k+1} \right) U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right] \\
&\quad + \dots \\
S[\rho(\bar{\mathbf{x}}, \mathbf{V})] &= \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} (\mathcal{I} - \rho^{\oplus}) U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right] \\
&\quad + \frac{1}{2} \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} \left(\sum_{k=0}^2 (-1)^k \binom{2}{k} [\rho^{\oplus}]^k \right) U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right] \\
&\quad + \frac{1}{3} \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} \left(\sum_{k=0}^3 (-1)^k \binom{3}{k} [\rho^{\oplus}]^k \right) U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right] \\
&\quad + \dots
\end{aligned}$$

If we use relation (7.67) again, we obtain

$$\begin{aligned}
S[\rho(\bar{\mathbf{x}}, \mathbf{V})] &= \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} (\mathcal{I} - \rho^{\oplus}) U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right] \\
&\quad + \frac{1}{2} \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} (\mathcal{I} - \rho^{\oplus})^2 U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right] \\
&\quad + \frac{1}{3} \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} (\mathcal{I} - \rho^{\oplus})^3 U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right] \\
&\quad + \dots \\
S[\rho(\bar{\mathbf{x}}, \mathbf{V})] &= \sum_{n=1}^{\infty} \frac{1}{n} \text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} (\mathcal{I} - \rho^{\oplus})^n U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right] \tag{7.69}
\end{aligned}$$

Furthermore, using the cyclic property of the trace, we have

$$\begin{aligned}
\text{Tr} \left[D(\bar{\mathbf{x}}) U_{\mathbf{S}} \rho^{\oplus} (\mathcal{I} - \rho^{\oplus})^n U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} \right] &= \text{Tr} \left[\rho^{\oplus} (\mathcal{I} - \rho^{\oplus})^n U_{\mathbf{S}}^{\dagger} D(\bar{\mathbf{x}})^{\dagger} D(\bar{\mathbf{x}}) U_{\mathbf{S}} \right] \\
&= \text{Tr} \left[\rho^{\oplus} (\mathcal{I} - \rho^{\oplus})^n U_{\mathbf{S}}^{\dagger} U_{\mathbf{S}} \right] \\
&= \text{Tr} \left[\rho^{\oplus} (\mathcal{I} - \rho^{\oplus})^n \right] \tag{7.70}
\end{aligned}$$

This leads to

$$\begin{aligned}
S[\rho(\bar{\mathbf{x}}, \mathbf{V})] &= \sum_{n=1}^{\infty} \frac{1}{n} \text{Tr} \left[\rho^{\oplus} (\mathcal{I} - \rho^{\oplus})^n \right] \\
S[\rho(\bar{\mathbf{x}}, \mathbf{V})] &= \text{Tr} \left[\rho^{\oplus} \sum_{n=1}^{\infty} \frac{1}{n} (\mathcal{I} - \rho^{\oplus})^n \right]
\end{aligned}$$

Using relation (7.66) again, we obtain

$$S[\rho(\bar{\mathbf{x}}, \mathbf{V})] = \text{Tr} [\rho^{\oplus} \log \rho^{\oplus}] = \text{Tr} [\rho(\mathbf{0}, \mathbf{V}^{\oplus}) \log \rho(\mathbf{0}, \mathbf{V}^{\oplus})]. \tag{7.71}$$

We consequently see that the von Neumann entropy of state $\rho(\bar{\mathbf{x}}, \mathbf{V})$ is the same as the von Neumann entropy of state $\rho(\mathbf{0}, \mathbf{V}^{\oplus})$, namely

$$S[\rho(\bar{\mathbf{x}}, \mathbf{V})] = S[\rho(\mathbf{0}, \mathbf{V}^{\oplus})]. \tag{7.72}$$

Let us compute the von Neumann entropy of state $\rho(\mathbf{0}, \mathbf{V}^\oplus)$. The latter is given by expression (7.63). We see that it is in fact a product state, being the tensor product of N thermal states. For such a state, the expression of the von Neumann entropy can be simplified. Indeed, as we are going to prove later in section 6.4, the von Neumann entropy $S(\rho_\otimes)$ of a product state of the form (4.20)

$$\rho_\otimes = \bigotimes_{k=1}^n \rho_k \quad (7.73)$$

can be simplified into

$$S(\rho_\otimes) = \sum_{k=1}^n S(\rho_k). \quad (7.74)$$

The entropy of state $\rho(\mathbf{0}, \mathbf{V}^\oplus)$ can consequently be simplified into

$$S(\rho[\mathbf{0}, \mathbf{V}^\oplus]) = \sum_{k=1}^N S\left[\rho^{th}\left(\frac{\nu_k - 1}{2}\right)\right] \quad (7.75)$$

Furthermore, we already computed the von Neumann entropy of a thermal state. It is indeed given by expression (7.58), which leads to

$$S\left[\rho^{th}\left(\frac{\nu_k - 1}{2}\right)\right] = \left(\frac{\nu_k + 1}{2}\right) \log\left(\frac{\nu_k + 1}{2}\right) - \left(\frac{\nu_k - 1}{2}\right) \log\left(\frac{\nu_k - 1}{2}\right). \quad (7.76)$$

Finally, after plugging expression (7.76) into (7.75) and using relation (7.72), we get the von Neumann entropy of state $\rho(\bar{\mathbf{x}}, \mathbf{V})$

$$S[\rho(\bar{\mathbf{x}}, \mathbf{V})] = \sum_{k=1}^N g(\nu_k) \quad (7.77)$$

where fonction $g : \mathbb{R} \rightarrow \mathbb{R}$ is defined as

$$g(x) = \left(\frac{x + 1}{2}\right) \log\left(\frac{x + 1}{2}\right) - \left(\frac{x - 1}{2}\right) \log\left(\frac{x - 1}{2}\right). \quad (7.78)$$

This expression generalizes the one we find in the 1-mode thermal case, and will be very useful when studying N -mode Gaussian states in terms of entropic inequalities, before adapting them to majorization relations.

7.3 Majorization in a beam splitter

Here we introduce a brief application of the theory of majorization in the field of quantum optics. This application concerns the evolution of quantum entanglement in a beam splitter. We will not be explaining the operating mode of a beam splitter, since it is not part of the study subject of the present report. We are however going to introduce an interesting results regarding the use of the theory of majorization in the beam splitter. Let us just say that a beam splitter is an apparatus which, as its name suggests, splits a ray of light into two. Let $|\psi^{(k)}(\theta)\rangle$ be the output state of a beam splitter if the input state is $|k, 0\rangle$, where θ is a characteristic of the transformation resulting from the beam splitter. The parameters k and 0 represent the number of photons at the entrance of each

of the arms of the beam splitter. When comparing the output states corresponding to two different inputs $|k, 0\rangle$ and $|k + 1, 0\rangle$, the following majorization relation is satisfied[8]

$$\lambda^{(k+1)} \prec \lambda^{(k)}, \quad (7.79)$$

where $\lambda^{(k)}$ is the vector of eigenvalues of the reduced density matrix

$$\rho_A^{(k)} = \text{Tr}_B |\psi^{(k)}(\theta)\rangle \langle \psi^{(k)}(\theta)|. \quad (7.80)$$

This implies that when we increase the number of photons at the entrance of the beam splitter, the entanglement of the 2-mode output state can only increase, and the 1-mode reduced states are getting more disordered, since for all measures of entanglement μ , the latter increases. This is a consequence of the following theorem [14].

Theorem 5. *Define the vector of eigenvalues of the reduced density matrix*

$$\rho_A^\psi = \text{Tr} |\psi\rangle \langle \psi| \quad (7.81)$$

by λ_ψ and the vector of eigenvalues of the reduced density matrix

$$\rho_A^\phi = \text{Tr} |\phi\rangle \langle \phi| \quad (7.82)$$

by λ_ϕ . *The relation $\lambda_\psi \prec \lambda_\phi$ is satisfied if and only if $\mu(\psi) \geq \mu(\phi)$ for all measures of entanglement μ .*

With this application to beam splitters, we see that the theory of majorization is also useful in the field of quantum optics. In the next section, we are going to look for majorization relation which can be applied to the field of quantum optics, and quantum Gaussian information in general.

8 Investigating the possible interconversion between Gaussian states

In this section, we are going to study majorization in the field of continuous-variable quantum information. We will in fact be mainly studying Gaussian states since they are, as we already said, of great practical importance in the field of continuous-variable. Moreover, we will be working with thermal states practically all the time since, as we showed when introducing proposition 9, we can use thermal decompositions in order to represent any Gaussian state.

We are first going to study the one-mode case, introducing a majorization relation in the case of thermal states, before generalizing it to the Gaussian states. We will then work in the N -mode case, using the one-mode case in order to prove some majorization relations satisfied by the tensor product of thermal states, before generalizing it to the N -mode Gaussian states, like in the one-mode case. Finally, we will test the majorization relations we found with some examples of 2-mode Gaussian states.

8.1 One-mode Gaussian states

We are first going to study one-mode Gaussian states. Take a thermal state $\rho^{th}(\bar{n})$, where \bar{n} is the mean number of photons in the state. We already defined it as

$$\rho^{th}(\bar{n}) = \sum_{n=0}^{\infty} \frac{\bar{n}^n}{(\bar{n}+1)^{n+1}} |n\rangle \langle n|. \quad (8.1)$$

First notice that according to relation (7.58), the von Neumann entropy of a thermal state can be written as

$$S[\rho^{th}(\bar{n})] = \log \left[\frac{(\bar{n}+1)^{\bar{n}+1}}{\bar{n}^{\bar{n}}} \right] \quad (8.2)$$

Let us see how the entropy of a thermal state evolves in terms of the mean photon number. Define the function $f: \mathbb{R} \rightarrow \mathbb{R}$ as

$$f(x) = \frac{(x+1)^{x+1}}{x^x}. \quad (8.3)$$

Using the relation

$$\frac{\partial}{\partial x} (U(x)^{U(x)}) = U(x)^{U(x)} U'(x) (\log[U(x)] + 1), \quad (8.4)$$

one finds that

$$\begin{aligned} \frac{\partial}{\partial x} f(x) &= \frac{1}{x^{2x}} [(x+1)^{x+1} (\log[x+1] + 1) x^x - x^x (\log[x] + 1) (x+1)^{x+1}] \\ &= \frac{1}{x^{2x}} (x+1)^{x+1} x^x \log \left[\frac{x+1}{x} \right] \end{aligned} \quad (8.5)$$

We therefore have

$$\frac{\partial}{\partial \bar{n}} \left[\frac{(\bar{n}+1)^{\bar{n}+1}}{\bar{n}^{\bar{n}}} \right] = \frac{1}{\bar{n}^{2\bar{n}}} (\bar{n}+1)^{\bar{n}+1} \bar{n}^{\bar{n}} \log \left[\frac{\bar{n}+1}{\bar{n}} \right] \quad (8.6)$$

Since $\bar{n} \geq 0$,

$$\frac{\bar{n}+1}{\bar{n}} \geq 1, \quad (8.7)$$

and

$$\frac{\partial}{\partial \bar{n}} \left[\frac{(\bar{n}+1)^{\bar{n}+1}}{\bar{n}^{\bar{n}}} \right] \geq 0. \quad (8.8)$$

Consequently, when the mean number of photon \bar{n} increases, the quantity

$$\frac{(\bar{n}+1)^{\bar{n}+1}}{\bar{n}^{\bar{n}}} \quad (8.9)$$

increases, as well as the entropy $S[\rho^{th}(\bar{n})]$. Therefore the following proposition is true.

Proposition 10. *For two thermal states $\rho^{th}(\bar{n}_1)$ and $\rho^{th}(\bar{n}_2)$ of respective mean numbers of photons \bar{n}_1 and \bar{n}_2 , the inequality $\bar{n}_1 \geq \bar{n}_2$ is equivalent to the inequality*

$$S[\rho^{th}(\bar{n}_1)] \geq S[\rho^{th}(\bar{n}_2)]. \quad (8.10)$$

The idea now is to find a similar inequality in terms of majorization relation. The proposition we introduce is as follows.

Proposition 11. *For two thermal states $\rho^{th}(\bar{n}_1)$ and $\rho^{th}(\bar{n}_2)$ of respective mean numbers of photons \bar{n}_1 and \bar{n}_2 , the inequality $\bar{n}_1 \geq \bar{n}_2$ is equivalent to the majorization relation*

$$\lambda_1 \prec \lambda_2, \quad (8.11)$$

where λ_1 and λ_2 are the vectors of eigenvalues of $\rho^{th}(\bar{n}_1)$ and $\rho^{th}(\bar{n}_2)$, respectively.

We are now going to show that this proposition is true. First, notice that the expression (8.1) of a thermal state is given in its diagonal representation. We therefore immediately conclude that the eigenvalues of a thermal state $\rho^{th}(\bar{n})$ are given by

$$\frac{\bar{n}^n}{(\bar{n} + 1)^{n+1}}, \quad (8.12)$$

for different values of n , and can be grouped in the vector

$$\lambda = \left(\frac{1}{\bar{n} + 1}, \frac{\bar{n}}{(\bar{n} + 1)^2}, \frac{\bar{n}^2}{(\bar{n} + 1)^3}, \dots \right). \quad (8.13)$$

In order to study proposition 11, we need to sort the elements of vector λ in decreasing order. Again, define the function $h_m : \mathbb{R} \rightarrow \mathbb{R}$ as

$$h_m(x) = \frac{m^x}{(m + 1)^{x+1}}. \quad (8.14)$$

The derivative of h_m in terms of x is given by

$$\frac{\partial}{\partial x} h_m(x) = m^x (m + 1)^{-x-1} \log \left[\frac{m}{m + 1} \right]. \quad (8.15)$$

When m is positive, the derivative of $h_m(x)$ is negative, which means that when n increases, the eigenvalues (8.12) of $\rho^{th}(\bar{n})$ decrease. The vector λ represented as (8.13) is therefore already in decreasing order, which means that we have

$$\lambda^\downarrow = \left(\frac{1}{\bar{n} + 1}, \frac{\bar{n}}{(\bar{n} + 1)^2}, \dots, 0 \right). \quad (8.16)$$

Now suppose we have two thermal states $\rho^{th}(\bar{n}_1)$ and $\rho^{th}(\bar{n}_2)$. Their vectors of ordered eigenvalues are respectively given by

$$\lambda_1^\downarrow = \left(\frac{1}{\bar{n}_1 + 1}, \frac{\bar{n}_1}{(\bar{n}_1 + 1)^2}, \dots, 0 \right) \quad (8.17)$$

and

$$\lambda_2^\downarrow = \left(\frac{1}{\bar{n}_2 + 1}, \frac{\bar{n}_2}{(\bar{n}_2 + 1)^2}, \dots, 0 \right). \quad (8.18)$$

Let us prove proposition 11. Given these two states and using the first definition of majorization, $\lambda_1 \prec \lambda_2$ if and only if

$$\begin{cases} \sum_{i=1}^k \lambda_{1i}^\downarrow \leq \sum_{i=1}^k \lambda_{2i}^\downarrow, & \forall k \\ \sum_{i=1}^{\infty} \lambda_{1i}^\downarrow = \sum_{i=1}^{\infty} \lambda_{2i}^\downarrow \end{cases} \quad (8.19)$$

Since $\rho^{th}(\bar{n}_1)$ and $\rho^{th}(\bar{n}_2)$ are density matrices of quantum states, the second relation of system (8.19) is always verified. On the other hand, the first relation is true if and only if

$$\sum_{i=1}^k \frac{\bar{n}_1^i}{(\bar{n}_1 + 1)^{i+1}} \leq \sum_{i=1}^k \frac{\bar{n}_2^i}{(\bar{n}_2 + 1)^{i+1}}, \quad \forall k \quad (8.20)$$

Since, as we already mentioned,

$$\sum_{i=1}^{\infty} \lambda_{1i} = \sum_{i=1}^{\infty} \lambda_{i2} = 1, \quad (8.21)$$

equation (8.20) is true if and only if

$$\begin{aligned} 1 - \sum_{i=k+1}^{\infty} \frac{\bar{n}_1^i}{(\bar{n}_1 + 1)^{i+1}} &\leq 1 - \sum_{i=k+1}^{\infty} \frac{\bar{n}_2^i}{(\bar{n}_2 + 1)^{i+1}}, \quad \forall k \\ \Leftrightarrow \sum_{i=k+1}^{\infty} \frac{\bar{n}_1^i}{(\bar{n}_1 + 1)^{i+1}} &\geq \sum_{i=k+1}^{\infty} \frac{\bar{n}_2^i}{(\bar{n}_2 + 1)^{i+1}}, \quad \forall k \\ \Leftrightarrow \frac{1}{\bar{n}_1 + 1} \sum_{i=k+1}^{\infty} \left(\frac{\bar{n}_1}{\bar{n}_1 + 1} \right)^i &\geq \frac{1}{\bar{n}_2 + 1} \sum_{i=k+1}^{\infty} \left(\frac{\bar{n}_2}{\bar{n}_2 + 1} \right)^i, \quad \forall k \end{aligned} \quad (8.22)$$

Moreover, when $|q| < 1$, we have

$$\sum_{i=m}^{\infty} q^i = \sum_{i=0}^{\infty} q^i - \sum_{i=0}^{m-1} q^i = \frac{1}{1-q} - \frac{1-q^m}{1-q} = \frac{q^m}{1-q} \quad (8.23)$$

where we used relation (7.52) and the formula of the sum of a geometric progression. Consequently, we see after some calculations that relation (8.22) is verified if and only if

$$\left(\frac{\bar{n}_1}{\bar{n}_1 + 1} \right)^{k+1} \geq \left(\frac{\bar{n}_2}{\bar{n}_2 + 1} \right)^{k+1}, \quad \forall k \quad (8.24)$$

$$\Leftrightarrow \frac{\bar{n}_1}{\bar{n}_1 + 1} \geq \frac{\bar{n}_2}{\bar{n}_2 + 1} \quad (8.25)$$

As we already did before, define function $t : \mathbb{R} \rightarrow \mathbb{R}$ as

$$t(x) = \frac{x}{x+1}. \quad (8.26)$$

The derivative of t in terms of x is given by

$$t'(x) = \frac{1}{(x+1)^2}, \quad (8.27)$$

which is always positive. We therefore deduce that (8.25) will be verified if and only if

$$\bar{n}_1 \geq \bar{n}_2. \quad (8.28)$$

This proves proposition 11.

In order to see what consequence proposition 11 has on Gaussian states in general, let us show that unitary transformations do not change the spectrum of a state. Any state ρ can always be written in its diagonal representation

$$\rho = \sum_i \sigma_i |i\rangle \langle i|, \quad (8.29)$$

where the σ_i represent the spectrum of ρ . Suppose a unitary transformation represented by an operator U verifying

$$U^\dagger = U^{-1}. \quad (8.30)$$

The transformation of state ρ according to U is given by

$$U\rho U^\dagger = U \sum_i \sigma_i |i\rangle \langle i| U^\dagger = \sum_i \sigma_i U|i\rangle \langle i| U^\dagger \quad (8.31)$$

If we define

$$|x_i\rangle = U|i\rangle, \quad (8.32)$$

then we have

$$U\rho U^\dagger = \sum_i \sigma_i |x_i\rangle \langle x_i|. \quad (8.33)$$

We consequently see that a unitary transformation does not change the spectrum, but only changes the eigenvectors according to relation (8.32). We saw that an arbitrary one-mode Gaussian state can be written, according to proposition 9, as

$$\rho(\bar{\mathbf{x}}, \mathbf{V}) = D(\bar{\mathbf{x}}) U_{\mathbf{S}} [\rho(\mathbf{0}, \mathbf{V}^\oplus)] U_{\mathbf{S}}^\dagger D(\bar{\mathbf{x}})^\dagger \quad (8.34)$$

with

$$\rho(\mathbf{0}, \mathbf{V}^\oplus) = \rho^{th} \left(\frac{\nu - 1}{2} \right) \quad (8.35)$$

where ν are the symplectic eigenvalues of \mathbf{V} , and $D(\bar{\mathbf{x}})$ and $U_{\mathbf{S}}$ are some unitary matrices. Therefore, the spectra of $\rho(\bar{\mathbf{x}}, \mathbf{V})$ and $\rho(\mathbf{0}, \mathbf{V}^\oplus)$ are the same. This leads to the following proposition.

Proposition 12. *Let $\rho(\bar{\mathbf{x}}_1, \mathbf{V}_1)$ and $\rho(\bar{\mathbf{x}}_2, \mathbf{V}_2)$ be two Gaussian states, then inequality $\nu_1 \geq \nu_2$ is equivalent to the majorization relation*

$$\lambda_1 \prec \lambda_2, \quad (8.36)$$

where ν_1 and ν_2 are the symplectic eigenvalues of \mathbf{V}_1 and \mathbf{V}_2 , respectively, and λ_1 and λ_2 are the vectors of eigenvalues of $\rho(\bar{\mathbf{x}}_1, \mathbf{V}_1)$ and $\rho(\bar{\mathbf{x}}_2, \mathbf{V}_2)$, respectively.

Before giving the motivation for the the results we found in the 1-mode case, we will generalize them to the N -mode case in the next section.

8.2 N -mode Gaussian states

The results we got in the one-mode case can be used in order to find more general results in the N -mode case. Suppose we have two N -mode Gaussian states $\rho(\bar{\mathbf{x}}_A, \mathbf{V}_A)$ and $\rho(\bar{\mathbf{x}}_B, \mathbf{V}_B)$. Suppose furthermore that the N symplectic eigenvalues of \mathbf{V}_A and \mathbf{V}_B are given by

$$\nu_k^A, \quad k = 1, \dots, N, \quad (8.37)$$

and

$$\nu_k^B, \quad k = 1, \dots, N, \quad (8.38)$$

respectively, and that the states can be written, according to proposition 9, as

$$\rho(\bar{\mathbf{x}}_A, \mathbf{V}_A) = D(\bar{\mathbf{x}}_A) U_{\mathbf{S}}^A [\rho(\mathbf{0}, \mathbf{V}_A^\oplus)] U_{\mathbf{S}}^{A\dagger} D_A(\bar{\mathbf{x}})^\dagger \quad (8.39)$$

with

$$\rho(\mathbf{0}, \mathbf{V}_A^\oplus) = \bigotimes_{k=1}^N \rho^{th} \left(\frac{\nu_k^A - 1}{2} \right) \quad (8.40)$$

and

$$\rho(\bar{\mathbf{x}}_B, \mathbf{V}_B) = D(\bar{\mathbf{x}}_B) U_{\mathbf{S}}^B [\rho(\mathbf{0}, \mathbf{V}_B^\oplus)] U_{\mathbf{S}}^{B\dagger} D_B(\bar{\mathbf{x}})^\dagger \quad (8.41)$$

with

$$\rho(\mathbf{0}, \mathbf{V}_B^\oplus) = \bigotimes_{k=1}^N \rho^{th} \left(\frac{\nu_k^B - 1}{2} \right) \quad (8.42)$$

where $D(\bar{\mathbf{x}}_A)$, $U_{\mathbf{S}}^A$, $D(\bar{\mathbf{x}}_B)$ and $U_{\mathbf{S}}^B$ are unitary matrices. We introduce the following proposition.

Proposition 13. *Let λ^A and λ^B be the vectors of eigenvalues of $\rho(\mathbf{0}, \mathbf{V}_A^\oplus)$ and $\rho(\mathbf{0}, \mathbf{V}_B^\oplus)$, respectively. Vector λ^A is majorized by λ^B , i.e. $\lambda^A \prec \lambda^B$, if*

$$\nu_k^A \geq \nu_k^B, \quad \forall k = 1, \dots, N, \quad (8.43)$$

where the ν_k^A are defined by relation

$$\mathbf{V}_A^\oplus = \bigoplus_{k=1}^N \nu_k^A \mathcal{I}, \quad (8.44)$$

and the ν_k^B are defined by relation

$$\mathbf{V}_B^\oplus = \bigoplus_{k=1}^N \nu_k^B \mathcal{I}. \quad (8.45)$$

Let us now prove proposition 13. Suppose we have

$$\nu_k^A \geq \nu_k^B, \quad \forall k = 1, \dots, N. \quad (8.46)$$

Obviously, it means that

$$\frac{\nu_k^A - 1}{2} \geq \frac{\nu_k^B - 1}{2}, \quad \forall k = 1, \dots, N. \quad (8.47)$$

Consequently, according to proposition 11,

$$\lambda_A^{(k)} \prec \lambda_B^{(k)}, \quad \forall k = 1, \dots, N, \quad (8.48)$$

where $\lambda_A^{(k)}$ and $\lambda_B^{(k)}$ are the vectors of eigenvalues of $\rho^{th} \left(\frac{\nu_k^A - 1}{2} \right)$ and $\rho^{th} \left(\frac{\nu_k^B - 1}{2} \right)$ respectively, for all k , which define states (8.40) and (8.42). Using property 1, one can say that

$$\lambda_A^{(k)} = D^{(k)} \lambda_B^{(k)}, \quad \forall k = 1, \dots, N \quad (8.49)$$

where $D^{(k)}$ is a doubly stochastic matrix for all k . Since $\lambda_A^{(k)}$ and $\lambda_B^{(k)}$ are the vectors of eigenvalues of $\rho^{th} \left(\frac{\nu_k^A - 1}{2} \right)$ and $\rho^{th} \left(\frac{\nu_k^B - 1}{2} \right)$, we can write them in their diagonal decompositions as

$$\rho^{th} \left(\frac{\nu_k^A - 1}{2} \right) = \sum_{p_k} \lambda_{A,p_k}^{(k)} |p_{kA}\rangle \langle p_{kA}|, \quad \forall k = 1, \dots, N \quad (8.50)$$

where $|p_{kA}\rangle$ are orthogonal states, for all k , and

$$\rho^{th} \left(\frac{\nu_k^B - 1}{2} \right) = \sum_{q_k} \lambda_{B,q_k}^{(k)} |q_{kB}\rangle \langle q_{kB}|, \quad \forall k = 1, \dots, N \quad (8.51)$$

where $|q_{kB}\rangle$ are orthogonal states, for all k . We therefore have

$$\begin{aligned} \rho(\mathbf{0}, \mathbf{V}_A^\oplus) &= \bigotimes_{k=1}^N \rho^{th} \left(\frac{\nu_k^A - 1}{2} \right) \\ &= \bigotimes_{k=1}^N \left(\sum_{p_k} \lambda_{A,p_k} |p_{kA}\rangle \langle p_{kA}| \right) \\ &= \sum_{p_1, p_2, \dots, p_N} \lambda_{A,p_1} \lambda_{A,p_2} \dots \lambda_{A,p_N} |p_{1A}, p_{2A}, \dots, p_{NA}\rangle \langle p_{1A}, p_{2A}, \dots, p_{NA}| \end{aligned} \quad (8.52)$$

Relation (8.52) gives a diagonal representation of $\rho(\mathbf{0}, \mathbf{V}_A^\oplus)$, where the vector of eigenvalues of the latter is given by the tensor product

$$\lambda^A = \bigotimes_{k=1}^N \lambda_A^{(k)}. \quad (8.53)$$

Similarly, the vector of eigenvalues of $\rho(\mathbf{0}, \mathbf{V}_B^\oplus)$ is given by

$$\lambda^B = \bigotimes_{k=1}^N \lambda_B^{(k)}. \quad (8.54)$$

Using property 2.16 of the tensor product and equation (8.49), we have

$$\lambda^A = \bigotimes_{k=1}^N D^{(k)} \lambda_B^{(k)} = \left(\bigotimes_{k=1}^N D^{(k)} \right) \left(\bigotimes_{k=1}^N \lambda_B^{(k)} \right) = \left(\bigotimes_{k=1}^N D^{(k)} \right) \lambda^B \quad (8.55)$$

Let us now show that the tensor product of two bistochastic is also bistochastic. Suppose A and B are two bistochastic matrices of dimension n , then their tensor product is given by

$$C = \begin{pmatrix} A_{11}B & \cdots & A_{1n}B \\ \vdots & \ddots & \vdots \\ A_{n1}B & \cdots & A_{nn}B \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} & \cdots & A_{11}B_{1n} & \cdots & A_{1n}B_{1n} \\ \vdots & & & & \\ A_{11}B_{n1} & & \ddots & & \vdots \\ \vdots & & & & \\ A_{n1}B_{n1} & \cdots & A_{n1}B_{nn} & \cdots & A_{nn}B_{nn} \end{pmatrix}. \quad (8.56)$$

We therefore have

$$\sum_i C_{ij} = \sum_k A_{kj} \left(\sum_l B_{lj} \right) = \sum_k A_{kj} = 1, \quad (8.57)$$

and

$$\sum_i C_{ji} = \sum_k A_{jk} \left(\sum_l B_{jl} \right) = \sum_k A_{jk} = 1. \quad (8.58)$$

We consequently see that C is also a bistochastic matrix. Similarly, one can say that the tensor product

$$D = \bigotimes_{k=1}^N D^{(k)} \quad (8.59)$$

is a bistochastic matrix. We therefore have

$$\lambda^A = D\lambda^B \quad (8.60)$$

where D is a bistochastic matrix. This leads to the condition $\lambda^A \prec \lambda^B$. Thus, we have proven proposition 13.

Like in the one-mode case, one can use the fact that the spectra of $\rho(\bar{\mathbf{x}}, \mathbf{V})$ and $\rho(\mathbf{0}, \mathbf{V}^\oplus)$, defined by (7.62) and (7.63), respectively, are the same, and introduce the following proposition.

Proposition 14. *Let λ^A and λ^B be the vectors of eigenvalues of $\rho(\bar{\mathbf{x}}_A, \mathbf{V}_A)$ and $\rho(\bar{\mathbf{x}}_B, \mathbf{V}_B)$, respectively. Vector λ^A is majorized by vector λ^B , i.e. $\lambda^A \prec \lambda^B$, if*

$$\nu_k^A \geq \nu_k^B, \quad \forall k = 1, \dots, N, \quad (8.61)$$

where the ν_k^A are the symplectic eigenvalues of \mathbf{V}_A for all k and the ν_k^B are the symplectic eigenvalues of \mathbf{V}_B for all k .

Notice that propositions 13 and 14 are only implications, since it is always possible to find states $\rho(\bar{\mathbf{x}}_A, \mathbf{V}_A)$ and $\rho(\bar{\mathbf{x}}_B, \mathbf{V}_B)$ that do not verify

$$\nu_k^A \geq \nu_k^B \quad (8.62)$$

for some values of k between 1 and N , but still verify the condition

$$\lambda^{G,A} \prec \lambda^{G,B}. \quad (8.63)$$

In order to illustrate proposition 14, let us take the case of the two-mode Gaussian states. Take a two-mode Gaussian state $\rho(\bar{\mathbf{x}}, \mathbf{V})$. Its covariance matrix can be written as [25]

$$\mathbf{V} = \begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{B} \end{pmatrix}, \quad (8.64)$$

where $\mathbf{A} = \mathbf{A}^T$, $\mathbf{B} = \mathbf{B}^T$, and \mathbf{C} are 2×2 real matrices. Note that for the uncertainty principle to be verified, the following conditions should be verified

$$\mathbf{V} > 0, \quad \det \mathbf{V} \geq 1 \quad \text{and} \quad \Delta \leq 1 + \det \mathbf{V}. \quad (8.65)$$

In this case, we have

$$\mathbf{V}^\oplus = (\nu_- \mathcal{I}) \oplus (\nu_+ \mathcal{I}) \quad (8.66)$$

where the symplectic spectrum $\{\nu_-, \nu_+\}$ is given by

$$\nu_\pm = \sqrt{\frac{\Delta \pm \sqrt{\Delta^2 - 4\det \mathbf{V}}}{2}}, \quad (8.67)$$

with $\Delta = \det \mathbf{A} + \det \mathbf{B} + 2\det \mathbf{C}$ and \det is the determinant. Let us take a particular case of the two-mode Gaussian states, which is in fact of importance. It has a covariance matrix of the form

$$\mathbf{V} = \begin{pmatrix} a\mathcal{I} & \mathbf{C} \\ \mathbf{C} & b\mathcal{I} \end{pmatrix}, \quad (8.68)$$

with

$$\mathbf{C} = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix}, \quad (8.69)$$

where a, b, c_1 and c_2 are real numbers. When $c_1 = -c_2 := c \geq 0$, conditions (8.65) are equivalent to

$$a + b \pm \sqrt{(a+b)^2 - 4(ab - c^2)} > 0, \quad (8.70)$$

$$(ab - c^2)(ab + c^2) \geq 1 \quad (8.71)$$

and

$$a^2 + b^2 - 2c^2 - 1 - (ab)^2 + c^4 \leq 0. \quad (8.72)$$

Moreover, the symplectic eigenvalues are given by

$$\nu_\pm = \frac{\sqrt{y} \pm (b-a)}{2} \quad (8.73)$$

where $y := (a+b)^2 - 4c^2$. Let us take a known state which verifies the last conditions. The EPR state [25]

$$\rho^{EPR}(r) = |r\rangle \langle r|_{EPR}, \quad (8.74)$$

where

$$|r\rangle_{EPR} = \sqrt{1 - \lambda^2} \sum_{n=0}^{\infty} (-\lambda)^n |n\rangle_a |n\rangle_b, \quad (8.75)$$

with $\lambda = \tanh r \in [0, 1]$, is a Gaussian state with zero mean and covariance matrix

$$\mathbf{V}_{EPR}(\eta) = \begin{pmatrix} \eta\mathcal{I} & \sqrt{\eta^2 - 1}\mathbf{Z} \\ \sqrt{\eta^2 - 1}\mathbf{Z} & \eta\mathcal{I} \end{pmatrix}, \quad (8.76)$$

where $\eta = \cosh 2r$ and

$$\mathbf{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (8.77)$$

According to relation (8.73), the symplectic eigenvalues of $\mathbf{V}_{EPR}(\eta)$ are given by

$$\nu_\pm^{EPR} = 1. \quad (8.78)$$

Now take a two-mode Gaussian state given by the tensor product of two thermal states

$$\rho(\mathbf{0}, \mathbf{V}^\oplus) = \rho^{th} \left(\frac{\nu_1 - 1}{2} \right) \otimes \rho^{th} \left(\frac{\nu_2 - 1}{2} \right). \quad (8.79)$$

The mean number of photon of a thermal state is obviously always non-negative. Therefore, we have

$$\begin{aligned} \frac{\nu_i - 1}{2} &\geq 0 & i = 1, 2 \\ \Rightarrow \nu_i &\geq 1 & i = 1, 2 \end{aligned} \quad (8.80)$$

This leads to

$$\begin{cases} \nu_1 \geq \nu_+^{EPR} \\ \nu_2 \geq \nu_-^{EPR} \end{cases} \quad (8.81)$$

Consequently, according to proposition 14, we should have

$$\lambda^\otimes \prec \lambda^{EPR} \quad (8.82)$$

where λ^\otimes and λ^{EPR} are the vectors of eigenvalues of $\rho(\mathbf{0}, \mathbf{V}^\oplus)$ and $\rho^{EPR}(r)$, respectively. This is in fact indeed the case, since $\rho^{EPR}(r)$ is a pure state. This example of a comparison between the EPR state and a 2-mode thermal state us in the idea that proposition 14 is true.

The utility of proposition 14 resides in the fact that it can be used to study the possibility of interconversion between different Gaussian states. Theorem 3 allows us to know if it is possible to transform a bipartite state $|\psi\rangle$ into another bipartite state $|\phi\rangle$ using LOCC. There is another theorem which studies the possibility to transform any state (bipartite or not) ρ into another state σ . It is in fact based on Uhlmann's theorem 3, and is as follows [8].

Theorem 6. *State ρ can be obtained from state σ by applying a mixture of unitaries if and only if*

$$\lambda(\rho) \prec \lambda(\sigma), \quad (8.83)$$

where $\lambda(\rho)$ and $\lambda(\sigma)$ are the vectors of eigenvalues of ρ and σ , respectively.

We see from theorem 6 that by simply comparing the eigenvalues of two states ρ and σ using the theory of majorization, it is possible to know if we can transform one state into the other. Theorem 6 therefore does not only give the motivation for proposition 14, but also for the propositions we gave in section 6.

8.3 Conclusion

In this section, we have been looking for majorization relations in the field of Gaussian quantum information. We began by investigating the 1-mode case, for which we first proposed a majorization relation between two thermal states of different mean number of photons. We then generalized the result to the case of 1-mode Gaussian states. Afterwards, we looked for majorization relations in the N -mode case, using the results we obtained in the 1-mode case. Again, the first relation we found concerns thermal states, giving a comparison between two tensor products of order N of thermal states with different mean number of photons. This result allowed us to introduce our most general relation applied to the field of Gaussian quantum information. This relation mainly compares N -mode Gaussian states using the theory of majorization. We then applied it to an example, in which we compared the EPR Gaussian state with a 2-mode thermal state. Theorem 6 allowed us to introduce the main motivation for the results we obtained. Our conclusion was that our relations are interesting since they permit us to investigate the possibility of converting one Gaussian state into another.

Conclusion

The theory of majorization and the notion of entropic measure of disorder are closely related. Based on this fact, the aim of this report was to look for majorization relations similar to already existing entropic inequalities. This was interesting to do for two main reasons. The first one is the fact that majorization relations are usually stronger than entropic inequalities, in the sense that they imply these entropic inequalities, but that the converse is not true. The second reason was the fact that when we dispose of majorization relations between two different quantum states, we know that we can transform one of the states into the other using some unitary transformation. The concept of entropy alone would not allow us to prove such a property.

For the purpose of doing this, we began by introducing the theory of majorization in the first section without using any concept of quantum mechanics. We gave some useful properties of majorization and explained its relation with the concept of disorder. The fundamental notions of the quantum theory we needed were introduced later in section 2. Thereby, we presented the basic concept of density matrix, along with operations of linear algebra we would need later. In section 3, we talked about the notion of entropy which was crucial in order to introduce our majorization relations. We gave both definitions of the Shannon entropy in classical theory and of the von Neumann entropy in quantum theory. Another important concept we introduced in section 4 was the fascinating resource of quantum entanglement, which we illustrated with the concept of quantum teleportation. Entanglement was related to the concepts of entropy and majorization through the different separability criteria we gave in section 5. Section 6 was dedicated to the introduction of new majorization relations, along with some examples and the limitations of these relations. The main results were linked to the entropic inequality of subadditivity and the interesting Bell diagonal state. We then introduced the field of Gaussian quantum information in section 7, in which we presented the notion of Gaussian state and the particular case of the thermal state, which we showed was of fundamental importance through the concept of thermal decomposition. We used these notions in order to introduce new majorization relation in section 8, in which we began by investigating the 1-mode case, before looking into the N -mode case, which we illustrated through an example.

The main relations we proved in this work exhibit the power of the theory of majorization. It has not been long since the concept of majorization was first applied to the quantum theory. This prompts us to say that the potential of the theory of majorization has not been entirely exploited, and that many mathematical properties have yet to be discovered using majorization.

References

- [1] Guillaume Aubrun and Ion Nechita. Catalytic majorization and l_p norms. *Communications in Mathematical Physics*, 278:133–144, May 2008.
- [2] Jürgen Audretsch. *Entangled Systems*. Wiley-Vch, 2007.
- [3] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Physical Review Letters*, 70(13):1895–1899, March 1993.
- [4] N. J. Cerf, C. Adami, and R. M. Gingrich. Reduction criterion for separability. *Physical Review A*, 60(2):898–909, 1999.
- [5] Nicolas J. Cerf and Christoph Adami. Negative entropy and information in quantum mechanics. *Physical Review Letters*, 79(26):5194–5197, December 1997.
- [6] Nicolas J. Cerf and Christoph Adami. Quantum extension of conditional probability. *Physical Review A*, 60(2):893–897, August 1999.
- [7] Claude Cohen-Tannoudji, Bernard Diu, and Frank Laloe. *Quantum Mechanics*, volume 1. Hermann, 1978.
- [8] C. N. Gagatsos, O. Oreshkov, and N. J. Cerf. Majorization relations and entanglement generation in a beam splitter.
- [9] N. Gisin. Hidden quantum nonlocality revealed by local filters. *Physics Letters A*, 210:151–156, January 1996.
- [10] Michal Horodecki, Pawel Horodecki, and Ryszard Horodecki. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A*, 223:1–8, November 1996.
- [11] Ryszard Horodecki, Pawel Horodecki, and Michal Horodecki. Quantum α -entropy inequalities: independent condition for local realism? *Physics Letters A*, 210:377–381, January 1996.
- [12] Ryszard Horodecki, Pawel Horodecki, Michal Horodecki, and Karol Horodecki. Quantum entanglement. *Reviews of Modern Physics*, 81:865–942, April-June 2009.
- [13] Danial Jonathan and Martin B. Plenio. Entanglement-assisted local manipulation of pure quantum states. *Physical Review Letters*, 83(17):3566–3569, October 1999.
- [14] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: Theory of Majorization and Its Applications*. Springer, 2011.
- [15] W. J. Munro, D. F. V. James, A. G. White, and P. G. Kwiat. Maximizing the entanglement of two mixed qubits. *Physical Review A*, 64(030302):1–4, 2001.
- [16] Michael A. Nielsen. Conditions for a class of entanglement transformations. *Physical Review Letters*, 83(2):436–439, July 1999.

- [17] Michael A. Nielsen. *An introduction to majorization and its applications to quantum mechanics*. 2002.
- [18] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [19] Michael A. Nielsen and Julia Kempe. Separable states are more disordered globally than locally. *Physical Review Letters*, 86(22):5184–5187, May 2001.
- [20] Michael A. Nielsen and Guifré Vidal. Majorization and the interconversion of bipartite states. *Quantum Information and Computation*, 1(1):76–93, 2001.
- [21] Asher Peres. Separability criterion for density matrices. *Physical Review Letters*, 77(8):1413–1415, August 1996.
- [22] José C. Principe. *INFORMATION THEORETIC LEARNING Renyi's Entropy and Kernel Perspectives*. Springer, 2010.
- [23] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.
- [24] Guifré Vidal. Entanglement monotones. *Journal of Modern Optics*, 47(2/3):355–376, 2000.
- [25] Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J. Cerf, Timothy C. Ralph, Jeffrey H. Shapiro, and Seth Lloyd. Gaussian quantum information. *Reviews of Modern Physics*, 84(2):621–662, April-June 2012.

A MATLAB codes

A.1 Matlab code `majorization_gamma.m`

```

1 clear
2 clc
3
4 format long
5
6 gamma = 0:0.01:1;
7
8 for i=1:max(size(gamma))
9
10     if (gamma(i) >= 2/3)
11         g = gamma(i)/2;
12     else
13         g = 1/3;
14     end
15
16     lambda_AB = [1-2*g g+gamma(i)/2 g-gamma(i)/2 0];
17     lambda_S = [g*(1-g) g*(1-g) g^2-2*g+1 g^2];
18     lambda_A = [1-g g];
19
20     major = Major(lambda_AB,lambda_S);
21
22 end

```

A.2 Matlab code `xab_majorization.m`

```

1 % We take a real and a=[0,1]
2 % We also take b real and b>=0
3 % We then have b = sqrt{1-a^2}
4
5 clear
6 clc
7
8 format long
9
10 x = 0:0.1:1;
11 a = 0:0.1:1;
12
13 M = []; %Majorization not verified, incomparable states
14
15 N = []; %Majorization verified
16
17 T = []; %Majorization not verified, inverse relation verified
18
19 for i=1:max(size(x))
20
21     for j=1:max(size(a))
22
23         b = sqrt(1-(a(j))^2);

```

```

24
25     rho_AB = zeros(4,4);
26
27     rho_AB(1,1) = (1-x(i))/2;
28     rho_AB(4,4) = (1-x(i))/2;
29     rho_AB(2,2) = x(i)*abs(a(j))^2;
30     rho_AB(3,3) = x(i)*(abs(b))^2;
31     rho_AB(2,3) = x(i)*a(j)*conj(b);
32     rho_AB(3,2) = x(i)*conj(a(j))*b;
33
34     rho_AB;
35
36     rho_A = TrX23(rho_AB, 2, [2 2]);
37     rho_B = TrX23(rho_AB, 1, [2 2]);
38     rho_S = tensor(rho_A,rho_B);
39
40     lambda_AB = eigs(rho_AB, length(rho_AB));
41     lambda_S = eigs(rho_S, length(rho_S));
42
43     error1 = Major(lambda_AB',lambda_S');
44     error2 = Major(lambda_S',lambda_AB');
45
46     sep = 0;
47
48     if x(i) <= 1/(1 + 2*a(j)*b)
49         sep = 1; % sep=1 means that the state is not separable
50     end
51
52     if error1 == 1
53
54         if error2 == 1
55
56             M = [M; x(i) a(j) b sep];
57
58         else
59
60             T = [T; x(i) a(j) b sep];
61
62         end
63
64     else
65
66         N = [N; x(i) a(j) b sep];
67
68     end
69
70     end
71
72 end
73
74 format short
75
76 M
77
78 N
79
80 T

```

A.3 Matlab code `subadditivity_Renyi.m`

```

1 clear;
2 clc;
3 close all
4
5 alpha = 4;
6
7 gamma = 0:0.001:1;
8 g = zeros(1,max(size(gamma)));
9 for i=1:max(size(gamma))
10     if (gamma(i) >= 2/3)
11         g(i) = gamma(i)/2;
12     else
13         g(i) = 1/3;
14     end
15 end
16
17 I_AB = 1/(1-alpha)*log2((1-2.*g).^alpha + (g+gamma/2).^alpha + ...
    (g-gamma/2).^alpha);
18 I_S = 1/(1-alpha)*log2(2*(g.*(1-g)).^alpha + (g.^2-2.*g+1).^alpha + ...
    (g.^2).^alpha);
19
20 plot(gamma, I_AB);
21 hold on;
22 plot(gamma, I_S, 'r');
23 xlabel('Value of \gamma.')
```

```

24 legend('Renyi entropy of order 4 of system AB.', 'Renyi entropy of ...
    order 4 of tensor product of system A and system B.');
```

A.4 Matlab code `gamma_catalysis.m`

```

1 clear
2 clc
3
4 format long
5
6 n = 1000; %Maximum Order of the Renyi Entropy
7
8 gamma = 0:0.01:1;
9
10 for i=1:max(size(gamma))
11
12     if (gamma(i) >= 2/3)
13         g = gamma(i)/2;
14     else
15         g = 1/3;
16     end
17
18     lambda_AB = [1-2*g; g+gamma(i)/2; g-gamma(i)/2; 0];
19     lambda_A = [1-g; g];
20     lambda_B = [1-g; g];
21
```

```

22     S_AB = alpha_entropy(lambda_AB,1);
23     S_A = alpha_entropy(lambda_A,1);
24     S_B = alpha_entropy(lambda_B,1);
25
26     if (S_AB > S_A + S_B)
27         error('No. ');
28     end
29
30     for alpha=1:n
31
32         [H_AB,lim_AB] = alpha_entropy(lambda_AB,alpha,0);
33         [H_A, lim_A] = alpha_entropy(lambda_A,alpha,lim_AB);
34         [H_B, lim_B] = alpha_entropy(lambda_B,alpha,lim_AB);
35
36         diff = H_AB-H_A-H_B;
37
38         if (H_AB > H_A + H_B)
39             %             alpha
40             %             H_AB
41             %             H_A
42             %             H_B
43             display('No. ');
44             gamma_i = gamma(i)
45             alpha
46             pause;
47             break;
48         end
49
50         if lim_AB==1
51             break;
52         end
53
54     end
55
56 end

```

A.5 Matlab code `alpha_entropy.m`

```

1 function [H,lim2] = alpha_entropy(lambda,alpha,lim1)
2
3 lim2=0;
4 H=0;
5
6 if alpha ==1
7
8     for i = 1:max(size(lambda))
9         if lambda(i,1) ~= 0
10            H = H - lambda(i,1)*log2(lambda(i,1));
11        end
12    end
13
14 else
15
16    for i = 1:max(size(lambda))
17        H = H + (lambda(i,1))^alpha;

```

```

18     end
19     H = 1/(1-alpha)*log2(H);
20     if ((H == Inf) || (H == -Inf) || lim1==1)
21         H = min(-log2(lambda));
22         lim2=1;
23     end
24
25 end

```

A.6 Matlab code `Major.m`

```

1 function [error] = Major(lambda_max,lambda_min)
2
3 % error will be equal to 0 if lambda_max majorizes lambda_min, and ...
4   it will
5 % be equal to 1 if not (in this case, it doesn't necessarily mean that
6 % lambda_min majorizes lambda_max).
7
8 error = 0;
9
10 dim_max = size(lambda_max,2);
11 dim_min = size(lambda_min,2);
12
13 dim = max(dim_max,dim_min);
14
15 lambda_max0 = zeros(1,dim);
16 lambda_min0 = zeros(1,dim);
17
18 lambda_max0(1,1:dim_max) = lambda_max;
19 lambda_min0(1,1:dim_min) = lambda_min;
20
21 lambda_max_maj = sort(lambda_max0,2,'descend');
22 lambda_min_maj = sort(lambda_min0,2,'descend');
23
24 sum_max = cumsum(lambda_max_maj);
25 sum_min = cumsum(lambda_min_maj);
26
27 condition = (sum_max - sum_min) > -1e-10;
28 last_condition = abs(sum_max(end) - sum_min(end)) < 1e-10;
29
30 if ((~all(condition)) || (~all(last_condition)))
31     sum_max;
32     sum_min;
33     %disp('No. ');
34     error = 1;
35 else
36     %disp('Ok. ');
37 end
38
39 end

```

A.7 Matlab code `tensor.m`

Code taken from website <http://www.dr-qubit.org/matlab.php>.

```

1 function M = tensor(varargin)
2
3 % TENSOR Tensor product
4 % author: Toby Cubitt
5 % requires: none
6 % license: GPL2
7 %
8 %   m = TENSOR(a,b,c,...) returns the kronecker product of its ...
   arguments.
9 %
10 %   Each argument should either be a matrix, or a cell array ...
   containing a
11 %   matrix and an integer. In the latter case, the integer specifies the
12 %   repeat count for the matrix, e.g. TENSOR(a,{b,3},c) = ...
   TENSOR(a,b,b,b,c).
13
14
15 %% Copyright (C) 2004–2009 Toby Cubitt
16 %%
17 %% This program is free software; you can redistribute it and/or
18 %% modify it under the terms of the GNU General Public License
19 %% as published by the Free Software Foundation; either version 2
20 %% of the License, or (at your option) any later version.
21 %%
22 %% This program is distributed in the hope that it will be useful,
23 %% but WITHOUT ANY WARRANTY; without even the implied warranty of
24 %% MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
25 %% GNU General Public License for more details.
26 %%
27 %% You should have received a copy of the GNU General Public License
28 %% along with this program; if not, write to the Free Software
29 %% Foundation, Inc., 51 Franklin Street, Fifth Floor, Boston,
30 %% MA 02110–1301, USA.
31
32
33 M = 1;
34 for j = 1:nargin
35     if iscell(varargin{j})
36         for k = 1:varargin{j}{2}
37             M = kron(M,varargin{j}{1});
38         end
39     else
40         M = kron(M,varargin{j});
41     end
42 end

```

A.8 Matlab code `TrX.m`

Code taken from website <http://www.dr-qubit.org/matlab.php>.

```

1 function x = TrX(p, sys, dim)

```

```

2
3 % TRX   Partial trace
4 % requires: nothing
5 % author: Toby Cubitt
6 % license: GPL2
7 %
8 %     RHO = TrX(PSI,SYS,DIM) traces out the subsystems specified in
9 %     vector SYS of state PSI (a state vector or density matrix) whose
10 %     subsystem dimensions are specified by the vector DIM.
11
12
13 %% Copyright (C) 2004–2009 Toby Cubitt
14 %%
15 %% This program is free software; you can redistribute it and/or
16 %% modify it under the terms of the GNU General Public License
17 %% as published by the Free Software Foundation; either version 2
18 %% of the License, or (at your option) any later version.
19 %%
20 %% This program is distributed in the hope that it will be useful,
21 %% but WITHOUT ANY WARRANTY; without even the implied warranty of
22 %% MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.  See the
23 %% GNU General Public License for more details.
24 %%
25 %% You should have received a copy of the GNU General Public License
26 %% along with this program; if not, write to the Free Software
27 %% Foundation, Inc., 51 Franklin Street, Fifth Floor, Boston,
28 %% MA 02110–1301, USA.
29
30
31 % check arguments
32 if any(sys > length(dim)) || any(sys < 0)
33     error('Invalid subsystem in SYS')
34 end
35 if (length(dim) == 1 && mod(length(p)/dim,1) ~= 0)...
36     || length(p) ~= prod(dim)
37     error('Size of state PSI inconsistent with DIM');
38 end
39
40
41 % remove singleton dimensions
42 if exist('setdiff')
43     % matlab
44     sys = setdiff(sys,find(dim == 1));
45 else
46     % octave
47     sys = complement(find(dim == 1),sys);
48 end
49 dim = dim(find(dim ~= 1));
50
51
52 % calculate systems, dimensions, etc.
53 n = length(dim);
54 rdim = dim(end:-1:1);
55 keep = [1:n];
56 keep(sys) = [];
57 dimtrace = prod(dim(sys));
58 dimkeep = length(p)/dimtrace;
59

```

```

60
61 if any(size(p) == 1)
62     % state vector
63     if size(p,1) == 1
64         p = p';
65     end
66     % reshape state vector to "reverse" ket on traced subsystems into ...
67     % a bra,
68     % then take outer product
69     perm = n+1-[keep(end:-1:1), sys];
70     x = reshape(permute(reshape(p, rdim), perm), [dimkeep, dimtrace]);
71     x = x*x';
72
73 else
74     % density matrix
75
76     % reshape density matrix into tensor with one row and one column index
77     % for each subsystem, permute traced subsystem indices to the end,
78     % reshape again so that first two indices are row and column
79     % multi-indices for kept subsystems and third index is a flattened ...
80     % index
81     % for traced subsystems, then sum third index over "diagonal" entries
82     perm = n+1-[keep(end:-1:1), keep(end:-1:1)-n, sys, sys-n];
83     x = reshape(permute(reshape(p, [rdim, rdim]), perm), ...
84                 [dimkeep, dimkeep, dimtrace^2]);
85     x = sum(x(:, :, [1:dimtrace+1:dimtrace^2]), 3);
86 end

```

A.9 Matlab code `TrX23.m`

Code taken from website <http://www.dr-qubit.org/matlab.php>.

```

1 function x = TrX23(p, sys, dim);
2
3 % TRX23 Partial trace of bi/tri-partite systems
4 % requires: nothing
5 % author: Toby Cubitt
6 % license: GPL2
7 %
8 % X = TrX23(P, SYS, DIM) traces out system SYS of state P (a state
9 % vector or a density matrix) with subsystem dimensions specified
10 % by DIM.
11 %
12 % If only one dimension is specified, i.e. DIM=[dim1], a
13 % dim1 x length(p)/dim1 system is assumed.
14 %
15 % If two are specified, i.e. DIM=[dim1, dim2], a dim1 x dim2
16 % system is assumed.
17 %
18 % DIM=[dim1, dim2, dim3] specifies a dim1 x dim2 x dim3 system
19 % (duh!)
20
21
22 %% Copyright (C) 2004-2009 Toby Cubitt

```



```
23 %%
24 %% This program is free software; you can redistribute it and/or
25 %% modify it under the terms of the GNU General Public License
26 %% as published by the Free Software Foundation; either version 2
27 %% of the License, or (at your option) any later version.
28 %%
29 %% This program is distributed in the hope that it will be useful,
30 %% but WITHOUT ANY WARRANTY; without even the implied warranty of
31 %% MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
32 %% GNU General Public License for more details.
33 %%
34 %% You should have received a copy of the GNU General Public License
35 %% along with this program; if not, write to the Free Software
36 %% Foundation, Inc., 51 Franklin Street, Fifth Floor, Boston,
37 %% MA 02110-1301, USA.
38
39
40 % check arguments
41 if sys > 2 && length(dim) < 3
42     error('SYS greater than number of subsystems')
43 end
44 if (length(dim) == 1 && mod(length(p)/dim,1) ~= 0)...
45     || length(p) ~= prod(dim)
46     error('Size of P inconsistent with DIM');
47 end
48
49
50 % sort out sys and dim arguments
51 switch length(dim)
52 % case 0
53 % dim1 = 2;
54 % dim2 = 1;
55 % dim3 = 2;
56 % if (sys == 2) sys = 3; end
57 case 1
58     dim1 = dim(1);
59     dim2 = 1;
60     dim3 = length(p)/dim1;
61     if (sys == 2) sys = 3; end
62 case 2
63     dim1 = dim(1);
64     dim2 = 1;
65     dim3 = dim(2);
66     if (sys == 2) sys = 3; end
67 case 3
68     dim1 = dim(1);
69     dim2 = dim(2);
70     dim3 = dim(3);
71 end
72
73
74 % calculate partial trace
75 switch any(size(p)==1)
76 case 1
77     % state vector
78     if size(p,1) == 1
79         p = p';
80     end
```

```
81
82  switch sys
83  case 1
84    x = reshape(p, dim2*dim3, dim1);
85  case 2
86    x = ...
      reshape(permute(reshape(p, dim3, dim2, dim1), [1, 3, 2]), dim1*dim3, dim2);
87  case 3
88    x = reshape(p, dim3, dim1*dim2).';
89  end
90  x = x*x';
91
92
93  case 0
94    % density matrix
95    switch sys
96    case 1
97      x=zeros(dim2*dim3);
98      indx=(1:dim2*dim3);
99      for k=0:dim1-1
100       x=x+p(indx+k*dim2*dim3,indx+k*dim2*dim3);
101     end
102
103    case 2
104      x=zeros(dim1*dim3);
105      indx=kron(ones(1, dim1), [1:dim3]) + ...
106        kron(dim2*dim3*[0:dim1-1], ones(1, dim3));
107      for k=0:dim2-1
108       x=x+p(indx+k*dim3,indx+k*dim3);
109     end
110
111    case 3
112      x=zeros(dim1*dim2);
113      indx=dim3*(0:dim1*dim2-1);
114      for k=1:dim3
115       x=x+p(indx+k,indx+k);
116     end
117  end
118 end
```