DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

ON SOME ENTROPY INEQUALITIES

by
Ayça İLERİ

August, 2022
İZMİR

ON SOME ENTROPY INEQUALITIES

A Thesis Submitted to the
Graduate School of Natural And Applied Sciences of Dokuz Eylül University
In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Mathematics

by
Ayça İLERİ

August, 2022
İZMİR

Ph.D. THESIS EXAMINATION RESULT FORM

We have read the thesis entitled "**ON SOME ENTROPY INEQUALITIES**" completed by **AYÇA İLERİ** under supervision of **PROF. DR. SELÇUK DEMİR** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy.

....

Prof. Dr. Selçuk DEMİR

Supervisor

....

Prof. Dr. Hasan KARABIYIK

Thesis Committee Member

....

Assist. Prof. Dr. Seçil GERGÜN

Thesis Committee Member

....

Prof. Dr. Müge KANUNİ ER

Examining Committee Member

....

Assoc. Prof. Dr. Ahmet YANTIR

Examining Committee Member

Prof. Dr. Okan FISTIKOĞLU
Director
Graduate School of Natural and Applied Sciences

ACKNOWLEDGEMENTS

Firstly, I would like to express my gratitude to Prof. Dr. Selçuk Demir, my Ph.D. supervisor, for his constant presence, generous guidance, and useful comments during my study. His counsel, unwavering support, never-ending motivation, and academic knowledge were vital in getting me through the difficult stages of my research. For my Ph.D. study, I could not have asked for a better supervisor.

In addition to my advisor, I would like to thank Prof. Dr. Noyan Fevzi ER and Prof. Dr. Halil ORUÇ for their critical reading and extremely useful ideas for my thesis and paper.

I would also want to thank TÜBİTAK (The Scientific and Technological Research Council of Turkey) for funding my Ph.D. thesis with a grant (BİDEB-2211).

Finally, I would like to thank my family for their endless patience and faith in me throughout my life.

Ayça İLERİ

ON SOME ENTROPY INEQUALITIES

ABSTRACT

In this thesis, we investigated some entropy-type inequalities in quantum information theory. We gave a proof of a conjecture from the paper Besenyei & Petz (2013) for a special case. This conjecture is a kind of partial subadditivity of quantum Tsallis entropy. Moreover, we obtained some operator inequalities and results related to this conjecture.

Keywords: Tsallis entropy, subadditivity, strong subadditivity, partial subadditivity, trace inequalities, operator inequalities.

BAZI ENTROPİ EŞİTSİZLİKLERİ ÜZERİNE

ÖZ

Bu tezde, kuantum bilgi teorisindeki bazı entropi tipi eşitsizlikleri araştırdık. Özel bir durum için Besenyei & Petz (2013) makalesindeki bir varsayımın ispatını verdik. Bu varsayım, kuantum Tsallis entropisinin bir tür kısmi alt eklenebilirliğidir. Ayrıca bu varsayımla ilgili bazı operatör eşitsizlikleri ve sonuçları elde ettik.

Anahtar kelimeler: Tsallis entropisi, alt toplamsallık, güçlü alt toplamsallık, kısmi alt toplamsallık, iz eşitsizlikleri, operatör eşitsizlikleri.

CONTENTS

LIST OF TABLES

Page

CHAPTER ONE
INTRODUCTION

1.1 Introduction And The Statement Of The Problem

Entropy is an important notion in both classical and quantum information theories. It is a measure of disorder or uncertainity in a system. From a classical point of view, the first attempt to measure information was made by Hartley in 1927 [McMahon (2007)]. Then, in his groundbreaking paper "The Mathematical Theory of Communication" [Shannon (1948)], Shannon presented a probabilistic way to measure the information content produced from an information source.

An information source is any physical devise sending messages (or signals) consisting of a string of letters from any alphabet. Suppose that a message is taken from an alphabet of n letters, say $a_1, \ldots, a_n$ where the probability of occurrence of the letter a_i is p_i. The important step taken by Shannon is that he quantified the information content of a message by taking the logarithm of the multiplicative inverse of the message's probability. It means that if the probability of a message is high, we will not get much information from it. A message with a low chance of occurrence, on the other hand, may include a considerable amount of information.

Shannon then defined information as the average of the probabilities of the messages, i.e. $H(X) = -\sum_{i=1}^{n} p_i \log_2 p_i$ where X is a discrete random variable with probability distribution $p_1, p_2, \ldots, p_n$ (for more detail see McMahon (2007), Petz (2007)).

One of the well-known properties of the Shannon entropy is the strong subadditivity [Petz (2007)]. Strong subadditivity relates the entropies of three subsystems to a bigger system where the probabilities of the subsystems are described by the marjinal distributions. Subadditivity is implied by the strong subadditivity.

The Tsallis entropy is a one-parameter extension of the Shannon entropy, defined by the formula $H_q(X) = -\sum_{i=1}^{n} p_i \log_q p_i$. Tsallis entropy has an important role in non-extensive statistics and statistical physics [Furuci (2006)]. Similar to the Shannon entropy, Tsallis entropy is strongly subadditive as well [Furuci (2006)] and hence is subadditive.

In [Besenyei & Petz (2013)], a new type of inequality which can be considered as "partial (strong) subadditivity" is introduced and proved for both Shannon and Tsallis entropies. The importance of partial (strong) subadditivity is that it implies (strong) subadditivity.

In this thesis, we are mostly intrested in quantum analogues of the above entropies. Every quantum mechanical system is associated with a complex Hilbert space that is called the state space. In the quantum world, instead of probability distributions, one uses density matrices, and density matrices are in one-to-one corresponce with states. A composite quantum mechanical system is described by the tensor product of the corresponding Hilbert spaces and marginal distributions by the partial traces of the density matrix [Petz (2007)].

The quantum analogue of the Shannon entropy is the von Neumann entropy, and it is defined by the formula $S(\rho) = -\operatorname{Tr}\rho\log\rho$ where ρ is a density operator. John von Neumann originally introduced this equality in 1932 in his groundbreaking paper "Mathematische Grundlagen der Quantenmechanik" [Von Neumann (2013)].

One of the fundamental properties of von Neumann entropy is its strong subadditivity. Lieb and Ruskai [Lieb & Ruskai (1973)] proved this fact in 1973. Strong subadditivity is used in coding theory, topological entanglement theory, conformal field theory and in some other research areas [Kim (2012)]. Due to the non-commutativity of the density matrices, proving strong subadditivity in the quantum case is more challenging than in the classical case. The strong subadditivity implies subadditivity in the quantum case as well.

A one-parameter extension of the von Neumann entropy is the quantum Tsallis entropy [Besenyei & Petz (2013), Petz & Virosztek (2014), Hiai & Petz (2014)].

Quantum Tsallis entropy is defined by the formula $S_q(\rho) = -\mathrm{Tr}\,\rho \log_q \rho$, where ρ is a density operator. As $q \to 1$, the von Neumann entropy is the limit of the Tsallis entropy [Petz (2007)]. The Tsallis entropy is known to be subadditive [Audenaert (2007)] but not strongly subadditive [Petz & Virosztek (2014)].

Almost at the same time as the paper Besenyei & Petz (2013) was published, Kim proved an operator extension of the strong subadditivity of von Neumann entropy, which is a kind of partial strong subadditivity [Kim (2012)]. In fact, this is the operator version of the Shannon entropy's partial strong subadditivity.

The following table summarizes all the cases:

Table 1.1 All Known Cases

Entropy	SA	PSA	SSA	PSSA
Shannon	Yes	Yes	Yes	Yes
Tsallis	Yes	Yes	Yes	Yes
Von Neumann	Yes	Yes	Yes	Partly
quantum Tsallis	Yes	?	No	No

SSA: strong subaddivity
SA: subadditivty
PSSA: partial strong subadditivity
PSA: partial subaddivity

In [Besenyei & Petz (2013)], an inequality related to the partial subadditivity of the quantum Tsallis entropy is conjectured when the real number q is greater than 1. This conjecture was proved in the same article for the product states, and it is also proved for 2×2 density matrices when $q = 2$. In this thesis, we investigate this conjecture and some related inequalities. The following paragraphs explain how this thesis is organized:

We begin with a quick overview of classical information theory in Chapter 2. We give definitions and some important facts related to the Shannon entropy and Tsallis entropy. Moreover, we give some simple results we obtained.

In Chapter 3, we give a brief explanation of quantum information theory's mathematical formalism. The postulates of quantum mechanics are covered in this chapter. Then we collect some important facts that will be useful for us.

In Chapter 4, we introduce the von Neumann entropy and collect some of its basic properties. Then we introduce the concepts “matrix monotonicity” and “matrix convexity”, which are essential tools for proving the fundamental theorems of quantum entropy. In addition, we give some important theorems of von Neumann entropy by making an analogy with the Shannon entropy.

In Chapter 5, we prove the inequality conjectured in the paper [Besenyei & Petz (2013)] for arbitrary m and n when $q = 2$. This completes the table above when $q = 2$. Then, motivated by this proof and some numerical examples, we conjecture an operator inequality which can be considered to be a kind of partial subadditivity of quantum Tsallis entropy. We also obtain some new results related to the operators in this inequality.

CHAPTER TWO

ENTROPY AND INFORMATION

This chapter is devoted to the Shannon entropy and its one-parameter extension, the Tsallis entropy. In this chapter we will introduce some basic definitions, properties and some important theorems related to these entropies. We will also provide proof of some of these facts.

2.1 Shannon Entropy

Definition 2.1.1. *Let X be a discrete random variable with possible values $\{x_1, ..., x_n\}$ and probability distribution $p(x) = P(X = x)$. Then the Shannon entropy of X is defined as*

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i)$$

with $0 \log 0$ taken to be 0.

Note 2.1.2. *Throughout this text all random variables are discrete unless otherwise stated.*

For two random variables, one can define the joint entropy as follows:

Definition 2.1.3. *Let X be a random variable with possible values $\{x_1, ..., x_m\}$ and Y be another random variable with possible values $\{y_1, ..., y_n\}$. Then the joint entropy of X and Y is defined as*

$$H(X, Y) = -\sum_{i=1}^{m} \sum_{j=1}^{n} p(x_i, y_j) \log p(x_i, y_j)$$

where the joint distribution $p(x_i, y_j)$ is defined by $P(X = x_i, Y = y_j)$.

In the following we collect some important and useful facts related to the Shannon entropy:

Proposition 2.1.4. *Let $p(x_i)$ represents the probability distribution of the random variable X with possible values $\{x_1,...,x_n\}$. Then $0 \leq H(x) \leq \log n$ with equality if and only if X has a uniform distribution, that is $p(x_i) = \frac{1}{n}$ for all i.*

One of the main characteristics of the Shannon entropy is its additivity.

Proposition 2.1.5. *If the random variables X and Y are independent (that is $p(x,y) = p(x)p(y)$), then*

$$H(X \times Y) = H(X) + H(Y). \tag{2.1}$$

Suppose we have two random variables, X and Y, and we know what the value of Y is. Then we know the information content $H(Y)$ of Y and in this case, we could make the following definition :

Definition 2.1.6. ***(Conditional Entropy)*** *The entropy of the random variable X with respect to another random variable Y is defined by*

$$H(X|Y) = -\sum_y p(y) \sum_x p(x|y) \log p(x|y).$$

(The conditional probability is denoted by the notation $p(x|y)$, which is defined by the formula $p(x|y) = \frac{p(x,y)}{p(y)}$.)

The following chain rule holds for the Shannon entropy :

Proposition 2.1.7. *Let X and Y be two random variables. Then*

$$H(X,Y) = H(X|Y) + H(Y). \tag{2.2}$$

The chain rule is important in the way that it relates the conditional entropy to the joint entropy.

Proposition 2.1.8. *For each* X *and* Y*, we have* $H(X|Y) \geq 0$ *and hence* $H(X,Y) \geq H(Y)$*. Similarly,* $H(Y|X) \geq 0$*, so* $H(X,Y) \geq H(X)$*.*

The nonnegativity of conditional entropy is simple to prove. For more details, see Nielsen & Chuang (2010). The other results of the proposition 2.1.8 follows immediately from the identity (2.2).

2.2 Basic Inequalities Related To The Shannon Entropy

One of the main results of standard entropy is the subadditivity:

Proposition 2.2.1. ***(Subadditivity of Shannon entropy)*** *If* X, Y *are two random variables, then*

$$H(X,Y) \leq H(X) + H(Y) \tag{2.3}$$

with equality if and only if X *and* Y *are independent random variables, that is,* $p(x,y) = p(x)p(y)$*.*

Inequality (2.3) is called the subadditivity of Shannon entropy.

Remark 2.2.2. *From Proposition 2.1.7 above, one can obtain the following inequality:*

$$\max\{H(X), H(Y)\} \leq H(X,Y) \leq H(X) + H(Y).$$

Subadditivity is a special case of a general result known as strong subadditivity:

Theorem 2.2.3. ***(Strong Subadditivity of Shannon entropy)*** *Let* X, Y, Z *be three ran dom variables with possible values* $\{x_1, ..., x_m\}$*,* $\{y_1, ..., y_n\}$ *and* $\{z_1, ..., z_r\}$*. Then*

$$H(X,Y,Z) + H(Y) \leq H(X,Y) + H(Y,Z) \tag{2.4}$$

with equality if and only if $X \longrightarrow Y \longrightarrow Z$ forms a Markov chain, that is $p(x,y,z) = p(x)p(y|x)p(z|y)$.

Inequality (2.4) is called the strong subadditivity of Shannon entropy.

Remark 2.2.4. *Consider the joint distribution*

$$p_{ijk} = p(x_i, y_j, z_k), \qquad 1 \le i \le m, \quad 1 \le j \le n, \quad 1 \le k \le r$$

and the marginal distributions

$$p_{ij-} = \sum_k p_{ijk}, \quad p_{-j-} = \sum_{i,k} p_{ijk}, \quad p_{-jk} = \sum_i p_{ijk}.$$

Then, by Definition 2.1.1, ineaquality (2.4) is equivalent to

$$\sum_{i,j,k} p_{ijk}(\log p_{ijk} + \log p_{-j-} - \log p_{ij-} - \log p_{-jk}) \ge 0. \tag{2.5}$$

Strong subadditivity is a much deeper result than the subadditivity. Because the subadditivity is implied by the strong subadditivity. The following remark is about this fact:

Remark 2.2.5. *If the random variable Y takes only one value, that is if $n = 1$ in (2.4), then we have two random variables and the strong subadditivity reduces to the subadditivity*

$$\sum_{i,k} p_{ik}(\log p_{ik} - \log p_{i-} - \log p_{-j}) \ge 0.$$

.

In most of this thesis, we will be interested in a new concept called "partial subadditivity", which is introduced in the paper [Besenyei & Petz (2013)]. One of the main observations of this paper is the following theorem :

Theorem 2.2.6. ***(Partial Strong Subadditivity of Shannon entropy)*** *Suppose that* X, Y, Z *are three random variables with possible values* $\{x_1, ..., x_m\}$, $\{y_1, ..., y_n\}$, *and* $\{z_1, ..., z_r\}$. *Then*

$$\sum_i p_{ijk}(\log p_{ijk} + \log p_{-j-} - \log p_{ij-} - \log p_{-jk}) \geq 0. \tag{2.6}$$

This inequality is introduced and proved in [Besenyei & Petz (2013)]. In (2.6), summing over j, k, one gets the SSA of Shannon entropy. Therefore the inequality (2.6) can be safely called partial strong subadditivity (PSSA).

Moreover, if the random variable Y takes only one value then we have two random variables and the partial strong subadditivity, which is the inequality (2.6), reduces to the partial subadditivity , that is

$$\sum_i p_{ik}(\log p_{ik} - \log p_{i-} - \log p_{-j}) \geq 0.$$

We close this section with one of the main concetps of the information theory: the relative entropy. It is a kind of distance used to measure the closeness of two probability distributions.

Definition 2.2.7. *Let* $p(x)$ *and* $r(x)$ *be two probability distributions on the same set. The relative entropy of* p *to* r *is defined by*

$$S(p(x) \| r(x)) = \sum_x p(x)(\log p(x) - \log r(x))$$

with $0 \log 0 = 0$ *and* $-p \log 0 = \infty$ *when* $p \neq 0$

Remark 2.2.8. *The relative entropy is nonnegative. The subadditivity of Shannon entropy can be proved by using this fact: Let* X *and* Y *be two probability* *distributions. Then the relative entropy of $p(x,y)$ to $p(x)p(y)$ is*

$$\begin{aligned} S(p(x,y)\|p(x)p(y)) &= \sum_{x,y} p(x,y)(\log p(x,y) - \log p(x)p(y)) \\ &= \sum_{x,y} p(x,y)\log p(x,y) - \sum_{x,y} p(x,y)\log p(x) - \sum_{x,y} p(x,y)\log p(y) \\ &= -H(X,Y) + H(X) + H(Y) \geq 0 \end{aligned}$$

which proves the subadditivity. Equality occurs if and only if $p(x,y) = p(x)p(y)$, that is X and Y are independent random variables.

2.3 Tsallis Entropy

The Tsallis entropy is a q-extention (or one-parameter extension) of the Shannon entropy. Before giving the definition of the Tsallis entropy we define a function:

Definition 2.3.1. *For $q \in \mathbb{R}$ the q-logarithm function $\log_q : \mathbb{R}^+ \longrightarrow \mathbb{R}$ is defined by*

$$\log_q x = \frac{x^{q-1} - 1}{q-1} \qquad (q \neq 1).$$

Taking the limit when $q \to 1$ we obtain the natural logarithm function.

Related to the q-logarithm function one can define the q-entropy in the following way:

Definition 2.3.2. *Let X be a random variable with possible values $\{x_1, ..., x_n\}$ and probability distribution $p(x) = P(X = x)$. Then the Tsallis entropy of X is defined by*

$$H_q(X) = -\sum_{i=1}^{n} p(x_i)\log_q p(x_i) = \frac{1}{1-q}\sum_{i=1}^{n}(p(x_i)^q - p(x_i)).$$

Note that the Shannon entropy is the limiting case of the Tsallis entropy when $q \to 1$, that is,

$$\lim_{q\to 1} H_q(X) = H(X).$$

Definition 2.3.3. ***(Tsallis Joint Entropy)*** *Let X and Y be two random variables with possible values $\{x_1, ..., x_m\}$ and $\{y_1, ..., y_n\}$. Then the Tsallis joint entropy [Furuci (2006)] of the pair (X, Y) with joint distribution $p(x_i, y_j)$ is*

$$H_q(X, Y) = -\sum_{i=1}^{m}\sum_{j=1}^{n} p(x_i, y_j)\log_q p(x_i, y_j).$$

In the followings we collect some important results of the Tsallis entropy:

Proposition 2.3.4. *Tsallis entropy is nonnegative and it attains its maximum value $-\log_q \frac{1}{n}$ when X is uniformly distributed over n outcomes, that is $p(x_i) = \frac{1}{n}$ for all i.*

Recall from the previous section that the Shannon entropy holds the additivity property for two independent random variables. For the Tsallis entropy we have a slightly different situation:

Proposition 2.3.5. *Let X, Y be two independent random variables. Then*

$$H_q(X\times Y) = H_q(X) + H_q(Y) + (1-q)H_q(X)H_q(Y), \tag{2.7}$$

which is called the pseduo-additivity of the Tsallis entropy [Furuci (2006)]. The identity (2.7) follows from the fact that $\log_q xy = \log_q x + \log_q y + (q-1)\log_q x \log_q y$.

Definition 2.3.6. ***(Tsallis Conditional Entropy)*** *Let X and Y be two random variables. Then, the Tsallis conditional entropy of X with respect to Y is*

$$H_q(X|Y) = -\sum_{y} p(y)^q \sum_{x} p(x|y)\log_q p(x|y). \tag{2.8}$$

The following proposition is the q-analogue of the chain rule of Shannon entropy:

Proposition 2.3.7. ***(Chain Rule For The Tsallis Entropy)*** *Let X and Y be two different random variables. Then*

$$H_q(X|Y) + H_q(Y) = H_q(X, Y). \tag{2.9}$$

There is a proof of (2.9) in Furuci (2006). In the following, we use a different approach to prove this identity.

Proof. Let us denote the joint distribution of X, Y by $p(x, y)$ and the marjinal distributions of X and Y by $p(x) = \sum_y p(x, y)$ and $p(y) = \sum_x p(x, y)$ respectively. Then we have

$$\begin{aligned} H_q(X, Y) - H_q(Y) &= -\sum_{x,y} p(x, y) \log_q p(x, y) + \sum_y p(y) \log_q p(y) \\ &= \sum_{x,y} p(x, y) \big(\log_q p(y) - \log_q p(x, y) \big) \end{aligned}$$

using the formula $\log_q x - \log_q y = -\log_q \left(\frac{y}{x}\right) x^{q-1}$, we have

$$\begin{aligned} \sum_{x,y} p(x, y) \big(\log_q p(y) - \log_q p(x, y) \big) &= -\sum_{x,y} p(x, y) \log_q \left(\frac{p(x, y)}{p(y)} \right) p(y)^{q-1} \\ &= -\sum_{x,y} p(x|y) p(y) \log_q p(x|y) p(y)^{q-1} \\ &= -\sum_{x,y} p(y)^q p(x|y) \log_q p(x|y) \\ &= -\sum_y p(y)^q \left(\sum_x p(x|y) \log_q p(x|y) \right) \\ &= H_q(X|Y). \end{aligned}$$

□

2.4 Basic Inequalities Related To The Tsallis Entropy

As for the Shannon entropy, Tsallis entropy is also subadditive:

Theorem 2.4.1. ***(Subadditivity of Tsallis entropy)*** *If X, Y are two random variables, then*

$$H_q(X,Y) \leq H_q(X) + H_q(Y) \tag{2.10}$$

for $q \geq 1$.

The inequality (2.10) is a special case of the following strong subadditivity:

Theorem 2.4.2. ***(Strong subadditivity of Tsallis entropy)*** *Let X, Y and Z be three random variables. Then*

$$H_q(X,Y,Z) + H_q(Y) \leq H_q(X,Y) + H_q(Y,Z) \tag{2.11}$$

for $q \geq 1$.

Proof. See Furuci (2006), Petz & Virosztek (2014). □

Note that the inequalities (2.10) and (2.11) are the q-analogues of the inequalities (2.3) and (2.4) respectively.

If the random variable Y takes only one value then we have two random variables and the strong subadditivity reduces to subadditivity, that is

$$H_q(X,Z) \leq H_q(X) + H_q(Z).$$

In the previous section, we introduced a new inequality which is called the partial strong subadditivity, and we saw that it holds true for the Shannon entropy. The following theorem is the q-extension of this inequality.

Theorem 2.4.3. ***(Partial strong subaddivity of Tsallis entropy)*** *Suppose that X,Y,Z are three random variables with possible values $\{x_1,...,x_m\}$, $\{y_1,...,y_n\}$, and $\{z_1,...,z_r\}$. Then*

$$\sum_i p_{ijk}(\log_q p_{ijk} + \log_q p_{-j-} - \log_q p_{-jk} - \log_q p_{ij-}) \geq 0 \tag{2.12}$$

for $q > 1$.

In [Besenyei & Petz (2013)], there is a nice proof of the above theorem. In the following, we prove the theorem above by using a different approach:

Proof. In the followings we use a similar methods used in the papers Furuci (2006), Petz & Virosztek (2014). The inequality (2.12) can be written as

$$-\sum_i p_{ijk}(\log_q p_{ijk} - \log_q p_{ij-}) \leq -\sum_i p_{ijk}(\log_q p_{-jk} - \log_q p_{-j-}) \tag{2.13}$$

using the following property

$$\log_q x - \log_q y = -\log_q \left(\frac{y}{x}\right) x^{q-1}$$

(2.13) can be written as

$$-\sum_i p_{ijk}\left(-\log_q \left(\frac{p_{ij-}}{p_{ijk}}\right) p_{ijk}^{q-1}\right) \leq -\sum_i p_{ijk}\left(-\log_q \left(\frac{p_{-j-}}{p_{-jk}}\right) p_{-jk}^{q-1}\right)$$

or equivalently

$$\sum_i p_{ijk}^q \log_q \left(\frac{p_{ij-}}{p_{ijk}}\right) \leq p_{-jk}^q \log_q \left(\frac{p_{-j-}}{p_{-jk}}\right). \tag{2.14}$$

We have to show the inequality (2.14). At this step we introduce the following function:

$$\mathrm{Log}_q x = -x \log_q x$$

observe that $\mathrm{Log}_q x = x^q \log_q \frac{1}{x}$. By using this identity, the left hand side of (2.14) can be written as

$$\sum_i p_{ijk}^q \log_q\left(\frac{p_{ij-}}{p_{ijk}}\right) = \sum_i p_{ij-}^q \left(\frac{p_{ijk}}{p_{ij-}}\right)^q \log_q\left(\frac{p_{ij-}}{p_{ijk}}\right)$$
$$= \sum_i p_{ij-}^q \operatorname{Log}_q\left(\frac{p_{ijk}}{p_{ij-}}\right)$$
$$= \sum_i p_{-j-}^q \left(\frac{p_{ij-}}{p_{-j-}}\right)^q \operatorname{Log}_q\left(\frac{p_{ijk}}{p_{ij-}}\right)$$

since $\left(\frac{p_{ij-}}{p_{-j-}}\right)^q \leq \left(\frac{p_{ij-}}{p_{-j-}}\right)$, for $q > 1$, we have

$$\sum_i p_{-j-}^q \left(\frac{p_{ij-}}{p_{-j-}}\right)^q \operatorname{Log}_q\left(\frac{p_{ijk}}{p_{ij-}}\right) \leq \sum_i p_{-j-}^q \frac{p_{ij-}}{p_{-j-}} \operatorname{Log}_q\left(\frac{p_{ijk}}{p_{ij-}}\right)$$
$$= p_{-j-}^q \left(\sum_i \frac{p_{ij-}}{p_{-j-}} \operatorname{Log}_q\left(\frac{p_{ijk}}{p_{ij-}}\right)\right)$$

since Log_q is concave we have

$$p_{-j-}^q \left(\sum_i \frac{p_{ij-}}{p_{-j-}} \operatorname{Log}_q\left(\frac{p_{ijk}}{p_{ij-}}\right)\right) \leq p_{-j-}^q \operatorname{Log}_q\left(\sum_i \frac{p_{ij-}}{p_{-j-}} \frac{p_{ijk}}{p_{ij-}}\right)$$
$$= p_{-j-}^q \operatorname{Log}_q\left(\frac{p_{-jk}}{p_{-j-}}\right).$$

Now we have the following inequality

$$\sum_i p_{ijk}^q \log_q\left(\frac{p_{ij-}}{p_{ijk}}\right) \leq p_{-j-}^q \operatorname{Log}_q\left(\frac{p_{-jk}}{p_{-j-}}\right).$$

Here we use the identity $\operatorname{Log}_q x = x^q \log_q \frac{1}{x}$ again. Then we have

$$\sum_i p_{ijk}^q \log_q\left(\frac{p_{ij-}}{p_{ijk}}\right) \leq p_{-j-}^q \operatorname{Log}_q\left(\frac{p_{-jk}}{p_{-j-}}\right)$$
$$= p_{-j-}^q \left(\frac{p_{-jk}}{p_{-j-}}\right)^q \log_q\left(\frac{p_{-j-}}{p_{-jk}}\right)$$
$$= p_{-jk}^q \log_q\left(\frac{p_{-j-}}{p_{-jk}}\right)$$

which proves the inequality (2.14). □

If the random variable Y takes only one value then we have two random variables and the inequality (2.12) reduces to the partial subadditivity inequality, that is

$$\sum_i p_{ik}(\log_q p_{ik} - \log_q p_k - \log_q p_i) \geq 0.$$

We close this chapter with a diagram that represent the relations between the inequalities we have seen so far.

Remark 2.4.4. *The following diagram holds for both Shannon entropy and Tsallis entropy*

$$\begin{array}{ccc} PSSA & \Longrightarrow & SSA \\ \Downarrow & & \Downarrow \\ PSA & \Longrightarrow & SA \end{array}$$

CHAPTER THREE
BASICS OF QUANTUM INFORMATION THEORY

In this chapter we shall give a brief explanation of the mathematical formalism of quantum information theory. Most of this chapter is based on the books Petz (2007) and Nielsen & Chuang (2010).

3.1 Quantum States And Density Matrices

Every quantum mechanical system is associated to a complex Hilbert space that is called the state space. The system is identified by its state vector which is a unit vector in the Hilbert space. In this thesis we will be mostly concerned with the finite quantum systems whose associated Hilbert space is $\mathbb{C}^n$.

In quantum mechanics a vector x is denoted by the notation $|x\rangle$ which is called a 'ket'. The inner product between the vectors $|x\rangle$ and $|y\rangle$ is denoted by the notation $\langle x|y\rangle$ and it is defined by

$$\langle x|y\rangle = \sum_{j=1}^{n} \overline{x}_j y_j.$$

In this notation $\langle x|$ is used for the dual vector of $|x\rangle$ and the inner product $\langle x|y\rangle$ is called the 'bra-ket' notation. Furthermore, the operator $|x\rangle\langle y|$ is linear and it is defined as

$$(|x\rangle\langle y|)(|z\rangle) = |x\rangle\langle y|z\rangle = \langle y|z\rangle|x\rangle.$$

Hence

$$|x\rangle\langle y| = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \left[\overline{y}_1, \overline{y}_2, \cdots, \overline{y}_n\right].$$

Example 3.1.1. *The most widely used quantum system in quantum information theory is the qubit. The state space of a qubit is described by two dimensional Hilbert space* $\mathbb{C}^2$. *The standard basis vectors* $(1,0)$ *and* $(0,1)$ *of* $\mathbb{C}^2$ *are denoted by* $|0\rangle$, $|1\rangle$ *respectively. A state vector* $|\phi\rangle$ *can be written as*

$$|\phi\rangle = a_0|0\rangle + a_1|1\rangle$$

where $a_0, a_1 \in \mathbb{C}$.
Similarly, any element $|\psi\rangle \in \mathbb{C}^n$ *can be written as*

$$|\psi\rangle = a_0|0\rangle + a_1|1\rangle + \ldots + a_{n-1}|n-1\rangle$$

where $a_0, a_1, \ldots, a_{n-1} \in \mathbb{C}$ *and* $\{|0\rangle, |1\rangle, \ldots, |n-1\rangle\}$ *is the standard basis of* $\mathbb{C}^n$.

The set of $n \times n$ matrices with complex entries, denoted by $\mathbb{M}_n(\mathbb{C})$, becomes a Hilbert space with respect to the Hilbert Schmidt inner product $\langle A, B\rangle = \operatorname{Tr} A^*B$. Furthermore the set of self adjoint matrices in $\mathbb{M}_n(\mathbb{C})$ is a real vector space.

The states of finite quantum mechanical systems are in one-to-one correspondence with density matrices. A density matrix $\rho \in \mathbb{M}_n(\mathbb{C})$ is a positive (definite) matrix (denoted by $\rho \geq 0$) with $\operatorname{Tr}\rho = 1$ and it is defined on the Hilbert space of the quantum system. When a positive definite matrix ρ is invertible we use the term strictly positive (definite) and it is denoted by the notation $\rho > 0$.

If we know the state of the system, say $|x\rangle$, then the system is in a pure state and the density matrix of the system is $|x\rangle\langle x|$. Otherwise the system is in a mixed state.

Mixed states are represented by the statistical mixture of pure states. More precisely if the quantum system is prepared in the state $|x_i\rangle$ with probability p_i then the density operator for the system is $\rho = \sum_{i=1}^{n} p_i |x_i\rangle\langle x_i|$ where $\sum_i p_i = 1$.

A self adjoint matrix in the vector space $\mathbb{M}_2(\mathbb{C})$ can be represented by the Pauli matrices $\sigma_0, \sigma_1, \sigma_2, \sigma_3$. That is, if $B \in \mathbb{M}_2(\mathbb{C})$ is a self adjoint matrix then B can be written as

$$B = x_0\sigma_0 + x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3$$

where

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

and $x_0, x_1, x_2, x_3 \in \mathbb{R}$.

Let $\mathcal{S}_n(\mathbb{C})$ be the state space of a quantum system. Namely,

$$\mathcal{S}_n(\mathbb{C}) = \{\rho \in \mathbb{M}_n(\mathbb{C}) : \quad \rho \geq 0, \quad \operatorname{Tr}\rho = 1\}.$$

The set $\mathcal{S}_n(\mathbb{C})$ is convex. It is also closed and bounded. Convexity of $\mathcal{S}_n(\mathbb{C})$ is obvious, so we show the boundedness and closedness of $\mathcal{S}_n(\mathbb{C})$:

Proof. (of boundednes of $\mathcal{S}_n(\mathbb{C})$) :
Let ρ be a density matrix. Then $\operatorname{sp}(\rho) \subset [0,1]$ and $\operatorname{Tr}\rho = 1$. The sup norm (or the maximum norm) of ρ is

$$\begin{aligned} \|\rho\| &= \max\Big\{|\lambda| : \quad \lambda \in \operatorname{sp}(\rho)\Big\} \\ &= 1 \end{aligned}$$

Hence $0 \leq \|\rho\| \leq 1$ and $\mathcal{S}_n(\mathbb{C})$ is bounded with respect to the sup norm. All norms are equivalent in finite dimensional Hilbert spaces, hence $\mathcal{S}_n(\mathbb{C})$ is bounded. □

Proof. *(of closedness of $\mathcal{S}_n(\mathbb{C})$)* :

Let $\mathcal{A}$ be the set of positive definite matrices in $\mathbb{M}_n(\mathbb{C})$ and h be a function defined by

$$h: \quad \mathcal{A} \longrightarrow \mathbb{R}^+,$$
$$B \longmapsto \operatorname{Tr} B.$$

Consider the set:

$$M = h^{-1}(1) = \Big\{B \in \mathcal{A} : \quad h(B) = 1\Big\}.$$

Since h is continuous, the set M(which is the set of density matrices) is closed. □

A density matrix $\rho \in \mathcal{S}_2(\mathbb{C})$ can be represented by

$$\rho = \frac{1}{2}(\sigma_0 + x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3) \tag{3.1}$$

where $\sigma_0, \sigma_1, \sigma_2, \sigma_3$ are Pauli matrices.

Formula (3.1) is equivalent to

$$\rho = \frac{1}{2}\begin{bmatrix} 1+x_3 & x_1 - ix_2 \\ x_1 + ix_2 & 1 - x_3, \end{bmatrix}$$

where $x_1, x_2, x_3 \in \mathbb{R}$. The matrix ρ is a density matrix if and only if $x_1^2 + x_2^2 + x_3^2 \leq 1$. Hence the state space of the qubit system can be described by the unit ball in $\mathbb{R}^3$ which is called the Bloch ball. The pure states are the points that satisfy $x_1^2 + x_2^2 + x_3^2 = 1$, which describes the unit sphere in $\mathbb{R}^3$, and this unit sphere is called the Bloch sphere. Any mixed state is a convex combination of pure states.

Let $\rho \in \mathbb{M}_2(\mathbb{C})$ be a density matrix. Then it has a Schmidt decomposition

$$\rho = \sum_j \mu_j |x_j\rangle\langle x_j| \tag{3.2}$$

where $0 \leq \mu_j \leq 1$, $\sum_j \mu_j = 1$ and x_j are unit vectors. Since ρ is a self-adjoint matrix, (3.2) is obtained from the spectral decomposition of ρ. Hence x_j's may be chosen as the eigenvectors of ρ and μ_j's are the corresponding eigenvalues. If all μ_j's are different, then the Schmidt decomposition is unique.

3.2 Composite Systems

Suppose that we have n physical systems with corresponding Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2, .., \mathcal{H}_n$ respectively. Then the Hibert space $\mathcal{H}$ of the composite system is defined by

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_n.$$

If the dimension of the subsystem $\mathcal{H}_i$ is N_i then the dimension of the composite system $\mathcal{H}$ is the product of the dimensions of the subsystems, that is

$$\dim\mathcal{H} = \prod_{i=1}^{n} N_i.$$

For $n = 2$ we have two subsystems and the composite system $\mathcal{H}_1 \otimes \mathcal{H}_2$ is called a bipartite system. If $\{u_i\}_{i=1}^{m}$ is a basis of $\mathcal{H}_1$ and $\{v_j\}_{j=1}^{n}$ is a basis of $\mathcal{H}_2$, then $\{u_i \otimes v_j\}$ is a basis of $\mathcal{H}_1 \otimes \mathcal{H}_2$, where $i = 1, \ldots, m$, $j = 1, \ldots, n$. The dimension of $\mathcal{H}_1 \otimes \mathcal{H}_2$ is mn.

Let $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ be a bipartite state. Then $|\psi\rangle$ can be written as a linear combination of the vectors from the component systems. That is

$$|\psi\rangle = \sum_{j=1}^{n} a_j |x_j\rangle \otimes |y_j\rangle.$$

Let $|x\rangle \in \mathcal{H}_1$, $|y\rangle \in \mathcal{H}_2$ and $A \in B(\mathcal{H}_1)$, $B \in B(\mathcal{H}_2)$, then the operator $A \otimes B$ acting on a vector $|x\rangle \otimes |y\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ is defined by

$$(A \otimes B)(|x\rangle \otimes |y\rangle) = A|x\rangle \otimes B|y\rangle.$$

This definition can be extended to all elements of $\mathcal{H}_1 \otimes \mathcal{H}_2$ in the following manner

$$\big(A \otimes B\big)\Big(\sum_j \alpha_j |x_j\rangle \otimes |y_j\rangle\Big) = \sum_j \alpha_j A|x_j\rangle \otimes B|y_j\rangle.$$

3.3 Entanglement

Tensor product of quantum systems give rise to define one of the most interesting and striking idea of quantum mechanics: The entanglement. To understand the entanglement let us look at an example:

Example 3.3.1. *Let us look at the following two-qubit state*

$$|\psi\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}$$

where $|00\rangle = |0\rangle \otimes |0\rangle$ *and* $|11\rangle = |1\rangle \otimes |1\rangle$. *The state* $|\psi\rangle$ *can not be written as a product of two states* $|x\rangle, |y\rangle \in \mathbb{C}^2$. *To prove this let us assume that* $|\psi\rangle = |x\rangle \otimes |y\rangle$, *where*

$$|x\rangle = a_1|0\rangle + a_2|1\rangle$$
$$|y\rangle = b_1|0\rangle + b_2|1\rangle$$

then

$$\begin{aligned}\frac{|00\rangle - |11\rangle}{\sqrt{2}} &= (a_1|0\rangle + a_2|1\rangle) \otimes (b_1|0\rangle + b_2|1\rangle)\\ &= a_1 b_1|00\rangle + a_1 b_2|01\rangle + a_2 b_1|10\rangle + a_2 b_2|11\rangle\end{aligned}$$

where $a_1 b_1 = \frac{1}{\sqrt{2}}, \quad a_1 b_2 = 0, \quad a_2 b_1 = 0, \quad a_2 b_2 = \frac{-1}{\sqrt{2}}.$

If $a_1 b_2 = 0$, *then* $a_1 = 0$ *or* $b_2 = 0$ *which contradicts the fact that* $a_1 b_1 = \frac{1}{\sqrt{2}}$ *or*

$a_2b_2 = \frac{-1}{\sqrt{2}}$. *Therefore* $|\psi\rangle$ *can not be expressed as the product of two single states.*

These type of states which can not be written as a product of states of its subsystems like given in Example 3.3.1 are called entangled states. The other states are called separable states.

One can formulate entanglement using of the density operator language. Let $\mathbb{M}_m(\mathbb{C})$, $\mathbb{M}_n(\mathbb{C})$ be the matrix algebras defined on the Hilbert spaces $\mathbb{C}^m$ and $\mathbb{C}^n$. Then the matrix algebra of the composite system $\mathbb{C}^m \otimes \mathbb{C}^n$ is

$$\mathbb{M}_{mn}(\mathbb{C}) = \mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C}).$$

There are two types of positive matrices in $\mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$. One consists of the elements written in the following form

$$\sum_j M_j \otimes N_j$$

where $M_j \in \mathbb{M}_m(\mathbb{C})$, $N_j \in \mathbb{M}_n(\mathbb{C})$ are positive matrices. These matrices are called separable positive matrices. Not every positive matrix is separable. That is, there are positive matrices in $\mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$ whose components are not positive.

A state of a quantum system is called separable if its density matrix is separable otherwise it is called entangled. A pure state is separable if and only if it is a product state.

Lemma 3.3.2. *Suppose that* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *are two Hilbert spaces and* $\{u_j\}_{j=1}^m$ *and* $\{v_i\}_{i=1}^n$ *are bases of* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *respectively. Let* $\psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ *be a unit vector with the following expansion*

$$\psi = \sum_{i,j} x_{ij} u_j \otimes v_i$$

and X *be the matrix determined by the entries* x_{ij}. *Then* $(X^*X)^T$ *is a density matrix* *and*

$$\langle \psi, (A \otimes I_2)\psi \rangle = \operatorname{Tr} A(X^*X)^T$$

for any $A \in B(\mathcal{H}_1)$.

The above lemma enables us to define the reduced density matrix in the following way:

Definition 3.3.3. *Let ρ be a density matrix in $\mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$. Then the reduced densities $\rho_1 \in \mathbb{M}_m(\mathbb{C})$ and $\rho_2 \in \mathbb{M}_n(\mathbb{C})$ of ρ are defined by*

$$\operatorname{Tr}(A \otimes I_2)\rho = \operatorname{Tr}(A\rho_1), \qquad \operatorname{Tr}(I_1 \otimes B)\rho = \operatorname{Tr}(B\rho_2)$$

for $A \in \mathbb{M}_m(\mathbb{C})$, $B \in \mathbb{M}_n(\mathbb{C})$. In the above formulas, $I_1 \in \mathbb{M}_m(\mathbb{C})$ and $I_2 \in \mathbb{M}_n(\mathbb{C})$ are the identity matrices.

One should note that the reduced densities are the quantum analogue of the marginal distributions.

Remark 3.3.4. *In Lemma 3.3.2, one can see that the reduced density of the pure state $|\psi\rangle\langle\psi|$ on the first subsystem is $(X^*X)^T$ and the reduced density of $|\psi\rangle\langle\psi|$ on the second subsystem is XX^*. The lemma shows that if the total system is in a pure state then the reduced densities have the same nonzero eigenvalues.*

A matrix in the tensor product space $\mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$ can be written as a block matrix with respect to the product basis. Let $\rho \in \mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$ be a density matrix. Then ρ can be written as

$$\rho = \sum_{i,j=1}^{m} E_{ij} \otimes A_{ij}$$

where $E_{ij} \in \mathbb{M}_m(\mathbb{C})$, $A_{ij} \in \mathbb{M}_n(\mathbb{C})$, and E_{ij} are called matrix units. For $m = n = 2$ we have

$$\rho = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{bmatrix}.$$

Then by Definition 3.3.3, the reduced densities of ρ can be written by the following formulas

$$\rho_1 = \begin{bmatrix} \operatorname{Tr} A_{11} & \operatorname{Tr} A_{12} \\ \operatorname{Tr} A_{12}^* & \operatorname{Tr} A_{22} \end{bmatrix} \qquad \text{and} \qquad \rho_2 = A_{11} + A_{22}.$$

It is possible to generalize these formulas for the density matrices in $\mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$:

$$\rho_{12} = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1m} \\ A_{12}^* & A_{22} & \dots & A_{2m} \\ \vdots & & & \vdots \\ A_{1m}^* & A_{2m}^* & \dots & A_{mm} \end{bmatrix}, \qquad (A_{ij} \in \mathbb{M}_n(\mathbb{C}),\ i, j = 1, 2, \dots, m).$$

Then the reduced densities $\rho_1 \in \mathbb{M}_m(\mathbb{C})$ and $\rho_2 \in \mathbb{M}_n(\mathbb{C})$ are defined as

$$\rho_1 = \begin{bmatrix} \mathrm{Tr}A_{11} & \mathrm{Tr}A_{12} & \dots & \mathrm{Tr}A_{1m} \\ \mathrm{Tr}A_{12}^* & \mathrm{Tr}A_{22} & \dots & \mathrm{Tr}A_{2m} \\ \vdots & & & \vdots \\ \mathrm{Tr}A_{1m}^* & \mathrm{Tr}A_{2m}^* & \dots & \mathrm{Tr}A_{mm} \end{bmatrix} \qquad \text{and} \qquad \rho_2 = A_{11} + A_{22} + \dots + A_{mm}.$$

Definition 3.3.5. *The linear mappings* $\mathrm{Tr}_1 : \mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C}) \longrightarrow \mathbb{M}_n(\mathbb{C})$ *and* $Tr_2 : \mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C}) \longrightarrow \mathbb{M}_m(\mathbb{C})$ *represented by the formulas*

$$Tr_1(C \otimes D) = (TrC)D \qquad and \qquad Tr_2(C \otimes D) = (TrD)C$$

on elementary tensors are called partial traces.

Remark 3.3.6. *Partial traces can be extended linearly to any matrix T in the tensor product space $\mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$. That is, for any matrix $T \in \mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$, partial* *traces of* T *are* $\mathrm{Tr}_1 T = T_2$ *and* $\mathrm{Tr}_2 T = T_1$.

In the followings we collect some useful facts of the partial trace:

1. For $A \in \mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$ and $B \in \mathbb{M}_m(\mathbb{C})$, $C \in \mathbb{M}_n(\mathbb{C})$ we have

$$\mathrm{Tr}_1(A(I \otimes C)) = (\mathrm{Tr}_1 A)C, \qquad \mathrm{Tr}_1((I \otimes C)A) = C(\mathrm{Tr}_1 A),$$

   similarly

$$\mathrm{Tr}_2(A(B \otimes I)) = (\mathrm{Tr}_2 A)B, \quad , \quad \mathrm{Tr}_2((B \otimes I)A) = B(\mathrm{Tr}_2 A).$$

   One can observe from the above relations that

$$\mathrm{Tr}_1(A(I \otimes C)) \neq \mathrm{Tr}_1((I \otimes C)A),$$
$$\mathrm{Tr}_2(A(B \otimes I)) \neq \mathrm{Tr}_2((B \otimes I)A).$$

2. For $A \in \mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$, $B \in \mathbb{M}_m(\mathbb{C})$ and $C \in \mathbb{M}_n(\mathbb{C})$ the following relations hold

$$\mathrm{Tr}_2(A(I \otimes B)) = \mathrm{Tr}_2((I \otimes B)A),$$
$$\mathrm{Tr}_1(A(C \otimes I)) = \mathrm{Tr}_1((C \otimes I)A).$$

Remark 3.3.7. *The reduced densities* $\rho_1 \in \mathbb{M}_m(\mathbb{C})$ *and* $\rho_2 \in \mathbb{M}_n(\mathbb{C})$ *of a density matrix* $\rho \in \mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$ *can be defined by using the partial trace as following*

$$\mathrm{Tr}_2 \rho = \rho_1, \qquad \mathrm{Tr}_1 \rho = \rho_2.$$

We close this chapter by defining the partial trace by using the operator language:

Definition 3.3.8. *Let* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *be finite dimensional Hilbert spaces with orthonormal bases* $\{e_i\}_{i=1}^m$, $\{f_j\}_{j=1}^n$ *respectively. For* $T \in B(\mathcal{H}_1 \otimes \mathcal{H}_2)$ *its partial trace* $T_1 = \mathrm{Tr}_{\mathcal{H}_2} T =$

$\mathrm{Tr}_2\, T$ *is an operator on* $\mathcal{H}_1$ *defined by the following equality*

$$\langle x, T_1 y\rangle = \sum_{j=1}^{n} \langle (x \otimes f_j), T(y \otimes f_j)\rangle$$

for all $x, y \in \mathcal{H}_1$.

Similarly one can define the partial trace $T_2 = \mathrm{Tr}_{\mathcal{H}_1} T = \mathrm{Tr}_1 T$ *by the following formula*

$$\langle x, T_2 y\rangle = \sum_{i=1}^{m} \langle (e_i \otimes x), T(e_i \otimes y)\rangle$$

for all $x, y \in \mathcal{H}_2$.

CHAPTER FOUR
ENTROPY IN QUANTUM INFORMATION THEORY

In Chapter 2, we investigated the entropies in the probability theory. By doing so, we used the classical probability vector, that is $(p_1, p_2, ..., p_n)$ of $p_i \geq 0$ with $\sum p_i = 1$. The quantum analogue of a probability vector is the density matrix. Recall from Chapter 3 that a density matrix $\rho \in \mathbb{M}_n(\mathbb{C})$ is a positive definite matrix with $\mathrm{Tr}\,\rho = 1$ and it is defined on the Hilbert space of the quantum system. This means that the vector $(\mu_1, \mu_2, ..., \mu_n)$ consisting of the eigenvalues of ρ forms a probability vector. This fact allows us to generalize the classical entropy to the density operators.

4.1 Definition And Some Basic Properties Of Von Neumann Entropy

Definition 4.1.1. *Let $\rho \in \mathbb{M}_n(\mathbb{C})$ be a density matrix. Then the von Neumann entropy of ρ is defined by*

$$S(\rho) = -\mathrm{Tr}\,\rho \log \rho \tag{4.1}$$

The self adjoint matrix $\rho \log \rho \in \mathbb{M}_n(\mathbb{C})$ is defined by using the spectral theorem.

Note that the Shannon entropy is a special case of the von Neumann entropy where the density matrices in the formula (4.1) are diagonal.

In the followings we collect some important and useful facts related to the von Neumann entropy:

Proposition 4.1.2. *The von Neumann entropy is basis independent, which means that if we choose the basis consisting of the eigenvectors of the density matrix ρ, the formula (4.1) is equivalent to*

$$S(\rho) = -\sum_{i=1}^{n} \lambda_i \log \lambda_i, \qquad \lambda_i \in \mathrm{sp}(\rho) \tag{4.2}$$

with $0\log 0 = 0$.

As with the classical entropy, the quantum entropy is nonnegative as well:

Proposition 4.1.3. *Let $\rho \in \mathbb{M}_n(\mathbb{C})$ be a density matrix. Then $0 \leq S(\rho) \leq \log n$. There is equality on the left hand side if and only if ρ is the density matrix of a pure state. And $S(\rho) = \log n$ if and only if $\rho = (\frac{1}{n})I$, that is ρ is the density matrix of a completely mixed state.*

Remember from Remark 3.3.4 that the reduced densities of a pure state have the same nonzero eigenvalues. The following proposition is related to this fact:

Proposition 4.1.4. *If the composite system is in a pure state ρ_{12} with reduced densities ρ_1 and ρ_2, (Note that reduced densities are defined in Section 3.3.) then $S(\rho_1) = S(\rho_2)$.*

We have seen in Section 2.1 that the Shannon entropy is additive for two independent random variables. Analoguosly, von Neumann entropy is additive for product states:

Proposition 4.1.5. *For a product state $\rho_1 \otimes \rho_2 \in \mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$, von Neumann entropy is additive, that is*

$$S(\rho_1 \otimes \rho_2) = S(\rho_1) + S(\rho_2). \tag{4.3}$$

In quantum information theory we use the functions of density matrices. Hence it seems to be useful to define matrix monotonicity and matrix convexity.

Definition 4.1.6. *Let (c,d) be an open interval of the set of real numbers and $g : (c,d) \longrightarrow \mathbb{R}$ be a real-valued function. The function g is matrix monotone (or operator monotone) if for every $n \in \mathbb{N}$ and every self adjoint matrix $C, D \in \mathbb{M}_n(\mathbb{C})$ with the spectrum* $\mathrm{sp}(C), \mathrm{sp}(D) \subset (c,d)$

$$C \leq D \Longrightarrow g(C) \leq g(D).$$

When $-g$ is matrix monotone, then g is matrix monotone decreasing.

Also, the function g is said to be matrix convex (or operator convex) if for every $n \in \mathbb{N}$ and every self adjoint matrix $C, D \in \mathbb{M}_n(\mathbb{C})$ with $\mathrm{sp}(C), \mathrm{sp}(D) \subset (c, d)$

$$g(tC + (1-t)D) \leq tg(C) + (1-t)g(D), \quad 0 \leq t \leq 1.$$

When $-g$ is matrix convex, then g is matrix concave.

Examples:

1. (Nielsen & Petz (2004)) The function $g(x) = 1/x$ is matrix(operator) monotone decreasing and matrix(operator) convex on $(0, \infty)$.

   Let $X \leq Y$ and X, Y be two strictly positive matrices. To prove the function $g(x) = 1/x$ is matrix monotone decreasing one can start with a special case where $X = I$. Since Y and I commute they are simiultaneously diagonalizable. So the result $Y^{-1} \leq I$ follows from the monotonically decreasing property of the real function $f(x) = 1/x$. The general result follows by taking $Y = X^{-1/2} Y X^{-1/2}$. The operator convexity of the function $1/x$ can be proved in a similar way.

2. (Hiai & Petz (2014)) The function $x \longmapsto \log x$ is matrix monotone and matrix concave on $(0, \infty)$.

   The proof of concavity follows from the well known formula

$$\log x = \int_0^\infty \Big(\frac{1}{1+t} - \frac{1}{x+t}\Big) dt \tag{4.4}$$

   and the operator convexity of the function $x \longmapsto 1/x$.

3. The function $x \longmapsto x \log x$ is matrix convex on $(0, \infty)$.

   Proof. By (4.4) we have

$$x \log x = \int_0^\infty \Big(\frac{x}{1+t} - \frac{x}{x+t}\Big) dt \tag{4.5}$$

or equivalently

$$x\log x = \int_0^\infty \Big(\frac{x}{1+t} + \frac{t}{x+t} - 1\Big)dt \tag{4.6}$$

Let A be a positive definite matrix. By the spectral theorem we have

$$A\log A = \int_0^\infty \Big((1+t)^{-1}A + t(A+tI)^{-1} - I\Big)dt \tag{4.7}$$

since the integrand in (4.7) is operator convex then the integral is also operator convex. □

4. (Hiai & Petz (2014), Carlen (2010)) The function $x \longmapsto \sqrt{x}$ is operator monotone and operator concave on $(0,\infty)$.

5. (Hiai (2017)) A power function $f(x) = x^m$ defined on $(0,\infty)$ is

   - operator monotone and operator concave when $m \in [0,1]$,
   - operator convex when $m \in [1,2]$,
   - operator convex and operator monotone decreasing when $m \in [-1,0]$,
   - when $\in (-\infty,-1) \cup (2,\infty)$, x^m is convex but not operator convex.

Related to the above definitions it is useful to give the following theorem:

Theorem 4.1.7. *(Peierls Inequality) Let $A \in \mathbb{M}_n(\mathbb{C})$ be a self adjoint matrix and f be a convex function on $\mathbb{R}$. The following inequality holds for any orthonormal basis $\{v_1, v_2, ..., v_n\}$ of $\mathbb{C}^n$*

$$\sum_{i=1}^{n} f(\langle v_i, Av_i\rangle) \leq \operatorname{Tr} f(A)$$

There is equality if each v_i is an eigenvector of A.

Proof. See, Carlen (2010), Theorem 2.9. □

An immediate consequence of the above theorem is the following proposition:

Proposition 4.1.8. *Let $f : (a,b) \longmapsto \mathbb{R}$ be a continuous convex function and $A \in \mathbb{M}_n(\mathbb{C})$ be a self adjoint matrix with $\mathrm{sp}(A) \subset (a,b)$. Then the function $A \longmapsto \mathrm{Tr}\, f(A)$ is operator convex on the set of self adjoint matrices.*

Proof. Since the set of self adjoint matrices is convex, A can be written as a convex combination of two self adjoint matrices A_1 and A_2, that is

$$A = \alpha A_1 + (1-\alpha)A_2, \qquad 0 \le \alpha \le 1.$$

We will show that

$$\mathrm{Tr}\, f(\alpha A_1 + (1-\alpha)A_2) \le \alpha\, \mathrm{Tr}\, f(A_1) + (1-\alpha)\, \mathrm{Tr}\, f(A_2). \tag{4.8}$$

Towards this end consider the orthonormal basis $\{v_1, v_2, ..., v_n\}$ of $\mathbb{C}^n$ consisting of the eigenvectors of A. By the above theorem we have

$$\begin{aligned}
\mathrm{Tr}\, f(A) &= \sum_{i=1}^{n} f(\langle v_i, Av_i\rangle) \\
&= \sum_{i=1}^{n} f\Big(\langle v_i, (\alpha A_1 + (1-\alpha)A_2)v_i\rangle\Big) \\
&= \sum_{i=1}^{n} f\Big(\alpha\langle v_i, A_1 v_i\rangle + (1-\alpha)\langle v_i, A_2 v_i\rangle\Big)
\end{aligned}$$

by the convexity of f we obtain

$$\begin{aligned}
\sum_{i=1}^{n} f\Big(\alpha\langle v_i, A_1 v_i\rangle + (1-\alpha)\langle v_i, A_2 v_i\rangle\Big) &\le \sum_{i=1}^{n} \alpha f(\langle v_i, A_1 v_i\rangle) + (1-\alpha)\langle v_i, A_2 v_i\rangle \\
&= \alpha \sum_{i=1}^{n} f(\langle v_i, A_1 v_i\rangle) + (1-\alpha) \sum_{i=1}^{n} f(\langle v_i, A_2 v_i\rangle).
\end{aligned}$$

Using the above theorem again we obtain

$$\alpha \sum_{i=1}^{n} f(\langle v_i, A_1 v_i\rangle) + (1-\alpha) \sum_{i=1}^{n} f(\langle v_i, A_2 v_i\rangle) \le \alpha\, \mathrm{Tr}\, f(A_1) + (1-\alpha)\, \mathrm{Tr}\, f(A_2)$$

which proves (4.8). □

The above proposition can be generalized as a theorem :

Theorem 4.1.9. *Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be continuous and let $n \in \mathbb{N}$. Then if $x \longmapsto f(x)$ is monotone then $A \longmapsto Tr(f(A))$ is operator monotone on the set of self adjoint matrices in $\mathbb{M}_n(\mathbb{C})$. Similarly, if $x \longmapsto f(x)$ is convex, then $A \longmapsto Tr(f(A))$ is operator convex on the set of self adjoint matrices.*

Proof. See, Carlen (2010), Theorem 2.10. □

It is clear that the von Neumann entropy is a concave function of the eigenvalues of the density matrix ρ. But more is true:

Proposition 4.1.10. *Von Neumann entropy is operator concave, that is*

$$S(\lambda\rho_1 + (1-\lambda)\rho_2) \geq \lambda S(\rho_1) + (1-\lambda)S(\rho_2), \qquad 0 \leq \lambda \leq 1 \tag{4.9}$$

where ρ_1, ρ_2 are density matrices in $\mathbb{M}_n(\mathbb{C})$.

Since the function $x \longmapsto -x \log x$ is concave, it is easy to prove the inequality (4.9) by using the Theorem 4.1.9 .

4.2 Quantum Relative Entropy

In Section 2.2 we defined relative entropy for the classical probability distributions. Now we will define it for density matrices:

Definition 4.2.1. *Let ρ, $\sigma \in \mathbb{M}_n(\mathbb{C})$ be density matrices. The relative entropy of ρ to σ is defined by*

$$S(\rho\|\sigma) = \begin{cases} \operatorname{Tr}\rho(\log\rho - \log\sigma) & \textit{if} \quad \operatorname{supp}(\rho) \leq \operatorname{supp}(\sigma) \\ \infty & \text{otherwise.} \end{cases}$$

Let us show that if $\text{supp}(\rho) \leq \text{supp}(\sigma)$ then $S(\rho\|\sigma) < \infty$: Let the Schmidt decomposition of ρ and σ be given by

$$\rho = \sum_{i=1}^{n} p_i |\psi_i\rangle\langle\psi_i|, \qquad \rho = \sum_{j=1}^{n} q_j |\phi_j\rangle\langle\phi_j|$$

then we have

$$\mathrm{Tr}\,\rho \log \rho = \sum_{i=1}^{n} p_i \log p_i \,, \qquad \mathrm{Tr}\,\rho \log \sigma = \sum_{i,j} p_i \log q_j |\langle\psi_i|\phi_j\rangle|^2$$

and

$$S(\rho\|\sigma) = \sum_{i} p_i \log p_i - \sum_{i,j} p_i \log q_j |\langle\psi_i|\phi_j\rangle|^2 \tag{4.10}$$

suppose that $|\psi_i\rangle \in \text{supp}(\rho)$ then $|\psi_i\rangle \in \text{supp}(\sigma)$ and we have

$$\begin{aligned} S(\rho\|\sigma) &= \sum_{i} p_i \log p_i - \sum_{i} p_i \log q_i |\langle\psi_i|\psi_i\rangle|^2 \\ &= \sum_{i} p_i \log p_i - p_i \log q_i \end{aligned} \tag{4.11}$$

since $|\psi_i\rangle$ in $\text{supp}(\rho) \cap \text{supp}(\sigma)$, the eigenvalues p_i of ρ and q_i of σ will be nonzero which makes the sum (4.11) finite.

As with the classical relative entropy, the quantum relative entropy is also nonnegative. One can show this fact by using Klein's inequality:

Theorem 4.2.2. ***(Klein's inequality)*** *For all self adjoint matrices $X, Y \in \mathbb{M}_n(\mathbb{C})$ and all differentiable convex functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ we have*

$$\mathrm{Tr}[f(X) - f(Y) - (X - Y)f'(Y)] \geq 0.$$

Proof. See, Carlen (2010), Theorem 2.11. □

Note that replacing f by $t \log t$ in the above theorem, one obtains the following inequality

$$\operatorname{Tr} X(\log X - \log Y) - \operatorname{Tr}(X - Y) \geq 0.$$

If X and Y are density matrices, say $X = \rho$, $Y = \sigma$ then $\operatorname{Tr}(\rho - \sigma) = 0$ and we have

$$\operatorname{Tr} \rho(\log \rho - \log \sigma) \geq 0, \tag{4.12}$$

which shows the nonnegativity of the relative entropy:

Remark 4.2.3. *Quantum relative entropy is nonnegative, that is*

$$S(\rho \| \sigma) \geq 0. \tag{4.13}$$

4.3 Some Important Inequalities Of Von Neumann Entropy

Theorem 4.3.1. ***(Subadditivity of Von Neumann entropy)*** *Let ρ_{12} be a density matrix in $\mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$ with reduced densites $\rho_1 \in \mathbb{M}_m(\mathbb{C})$ and $\rho_2 \in \mathbb{M}_n(\mathbb{C})$. Then*

$$S(\rho_{12}) \leq S(\rho_1) + S(\rho_2) \tag{4.14}$$

or equivalently

$$\operatorname{Tr} \rho_{12}(\log \rho_{12} - \log \rho_1 - \log \rho_2) \geq 0.$$

Equality holds in (4.14) if and only if ρ_{12} is a product state, that is $\rho_{12} = \rho_1 \otimes \rho_2$.

Proof. See Carlen (2010), Bhatia (2009), Petz (2008). □

An extended version of quantum relative entropy is the ‘relative quasi-entropy’. This concept was first introduced by Dénes Petz (Petz, 1986), in the following sense:

Definition 4.3.2. *Let $\rho, \sigma \in \mathbb{M}_n(\mathbb{C})$ be two invertible density matrices, $K \in \mathbb{M}_n(\mathbb{C})$*

and $f : (0,\infty) \longrightarrow \mathbb{R}$ be a real function. Then the relative quasi-entropy (or f-divergence) is defined by

$$S_f^K(\rho\|\sigma) = \langle K\rho^{1/2}, f(\Delta_{\sigma,\rho})K\rho^{1/2}\rangle \tag{4.15}$$
$$= \operatorname{Tr} K^* f(\Delta_{\sigma,\rho})K\rho$$

where $\langle X, Y\rangle = \operatorname{Tr} X^*Y$ is the Hilbert Schmidt inner product and $\Delta_{\sigma,\rho} : \mathbb{M}_n(\mathbb{C}) \longrightarrow \mathbb{M}_n(\mathbb{C})$ is the relative modular operator defined by Araki (1976) as follows

$$\Delta_{\sigma,\rho}(K) = L_\sigma R_\rho^{-1}(K) = \sigma K\rho^{-1}.$$

The operators L and R in the above formula are called the superoperators and they commute, that is, $LR = RL$.

Note that setting $K = I$ and $f(x) = -\log x$ in (4.15) we obtain the quantum relative entropy, namely $S(\rho\|\sigma) = \operatorname{Tr}\rho(\log\rho - \log\sigma)$. Thus the relative entropy is a special case of the relative quasi-entropy.

The relative entropy is monotone in the following sense:

Theorem 4.3.3. ***(Monotonicity of the Relative Entropy)*** *Let $\rho_{12}, \sigma_{12} \in \mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$ be density matrices with reduced densities $\rho_1, \sigma_1 \in \mathbb{M}_m(\mathbb{C})$ respectively. Then*

$$S(\rho_{12}\|\sigma_{12}) \geq S(\rho_1\|\sigma_1). \tag{4.16}$$

Proof. A delicate proof of the monotonicity can be found in Nielsen & Petz (2004). □

The monotonicity inequality (4.16) holds true not only for the function $-\log x$ but also for all operator convex functions, that is, if f is an operator convex function then

$$S_f(\rho_{12}\|\sigma_{12}) \geq S_f(\rho_1\|\sigma_1). \tag{4.17}$$

Moreover in Petz & Virosztek (2014) there is an extension of the inequality (4.17) in the following way

$$S_f^{T\otimes I}(\rho_{12}\|\sigma_{12}) \geq S_f^{T}(\rho_1\|\sigma_1) \tag{4.18}$$

where T is any matrix in $\mathbb{M}_m(\mathbb{C})$ and I is the identity matrix in $\mathbb{M}_n(\mathbb{C})$.

The operator $T\otimes I$ in (4.18) can be replaced by the operator $T\otimes V_2$, where V_2 is a unitary matrix:

Lemma 4.3.4. *Let $P_{12}, Q_{12} \in \mathbb{M}_m(\mathbb{C})\otimes\mathbb{M}_n(\mathbb{C})$ be two strictly positive matrices and $V_2 \in \mathbb{M}_n(\mathbb{C})$ be a unitary matrix. For any operator convex function f and any matrix $T \in \mathbb{M}_m(\mathbb{C})$ the following inequality holds*

$$S_f^{T\otimes V_2}(P_{12}\|Q_{12}) \geq S_f^{T}(P_1\|Q_1). \tag{4.19}$$

There is a proof of the above lemma in Jenčová & Ruskai (2010). In the following we give another proof which is based on a technique due to Petz & Virosztek (2014).

Proof. Let $U : \mathbb{M}_m(\mathbb{C}) \longrightarrow \mathbb{M}_m(\mathbb{C})\otimes\mathbb{M}_n(\mathbb{C})$ be a linear map defined by the formula

$$U(X) = (XP_1^{-1/2}\otimes V_2)P_{12}^{1/2}. \tag{4.20}$$

Claim: U is an isometry:
Let X, Y in $\mathbb{M}_m(\mathbb{C})$. Then

$$\begin{aligned}
\langle U(X), U(Y)\rangle &= \langle (XP_1^{-1/2}\otimes V_2)P_{12}^{1/2}, (YP_1^{-1/2}\otimes V_2)P_{12}^{1/2}\rangle \\
&= \mathrm{Tr}(P_1^{-1/2}X^*YP_1^{-1/2}\otimes V_2^*V_2)P_{12} \\
&= \mathrm{Tr}(P_1^{-1/2}X^*YP_1^{-1/2}\otimes I_2)P_{12} \\
&= \mathrm{Tr}\, P_1^{-1/2}X^*YP_1^{-1/2}P_1 \\
&= \mathrm{Tr}\, X^*Y = \langle X, Y\rangle.
\end{aligned}$$

Now we will find the operator U^* which is the adjoint of U:

$$\begin{aligned}
\langle Y, U(X)\rangle &= \operatorname{Tr}(Y^*(XP_1^{-1/2}\otimes V_2)P_{12}^{1/2}) \\
&= \operatorname{Tr}(YP_{12}^{1/2})^*(X\otimes I)(P_1^{-1/2}\otimes V_2) \\
&= \operatorname{Tr}(P_1^{-1/2}\otimes V_2)(YP_{12}^{1/2})^*(X\otimes I) \\
&= \operatorname{Tr}\left[\left((YP_{12}^{1/2})(P_1^{-1/2}\otimes V_2^*)\right)^*(X\otimes I)\right] \\
&= \operatorname{Tr}\left[\operatorname{Tr}_2\left(YP_{12}^{1/2}(P_1^{-1/2}\otimes V_2^*)\right)^* X\right] \\
&= \langle \operatorname{Tr}_2\left(YP_{12}^{1/2}(P_1^{-1/2}\otimes V_2^*)\right), X\rangle
\end{aligned}$$

hence $U^*(Y) = \operatorname{Tr}_2\left(YP_{12}^{1/2}(P_1^{-1/2}\otimes V_2^*)\right)$.

Moreover for any matrix X in $\mathbb{M}_m(\mathbb{C})$, U satisfies the following identity:

$$\begin{aligned}
U^*\Delta_{Q_{12},P_{12}}U(X) &= U^*(Q_{12}(XP_1^{-1/2}\otimes V_2)P_{12}^{-1/2}) \\
&= \operatorname{Tr}_2\left(Q_{12}(XP_1^{-1/2}\otimes V_2)P_{12}^{-1/2}P_{12}^{1/2}(P_1^{-1/2}\otimes V_2^*)\right) \\
&= \operatorname{Tr}_2\left(Q_{12}(XP_1^{-1}\otimes V_2V_2^*)\right) \\
&= \operatorname{Tr}_2\left(Q_{12}(XP_1^{-1}\otimes I)\right) \\
&= Q_1XP_1^{-1} = \Delta_{Q_1,P_1}(X).
\end{aligned}$$

By the formula (4.15)

$$\begin{aligned}
S_f^T(P_1\|Q_1) &= \langle TP_1^{1/2}, f(\Delta_{Q_1,P_1})TP_1^{1/2}\rangle \\
&= \langle TP_1^{1/2}, f(U^*\Delta_{Q_{12},P_{12}}U)TP_1^{1/2}\rangle.
\end{aligned}$$

Since f is operator convex and U is an isometry (Lemma 2 of Nielsen & Petz (2004)) we have

$$f(U^*\Delta_{Q_{12},P_{12}}U) \leq U^*f(\Delta_{Q_{12},P_{12}})U$$

which implies the following inequality

$$\langle TP_1^{1/2}, f(U^*\Delta_{Q_{12},P_{12}}U)TP_1^{1/2}\rangle \leq \langle TP_1^{1/2}, U^*f(\Delta_{Q_{12},P_{12}})U(TP_1^{1/2})\rangle.$$

In addition, $U(TP_1^{1/2}) = (TP_1^{1/2}P_1^{-1/2} \otimes V_2)P_{12}^{1/2} = (T \otimes V_2)P_{12}^{1/2}$. Thus

$$\begin{aligned}\langle TP_1^{1/2}, U^*f(\Delta_{Q_{12},P_{12}})U(TP_1^{1/2})\rangle &= \langle (T\otimes V_2)P_{12}^{1/2}, f(\Delta_{Q_{12},P_{12}})(T\otimes V_2)P_{12}^{1/2}\rangle \\ &= S_f^{T\otimes V_2}(P_{12}\|Q_{12}).\end{aligned}$$

This completes the proof. □

One of the fundamental properties of the von Neumann entropy is the strong subadditivity:

Theorem 4.3.5. *Let* $\rho_{123} \in M_m(\mathbb{C}) \otimes M_n(\mathbb{C}) \otimes M_r(\mathbb{C})$ *be a density matrix and* $\rho_{12} \in M_m(\mathbb{C}) \otimes M_n(\mathbb{C})$, $\rho_2 \in M_m(\mathbb{C})$, $\rho_{23} \in M_n(\mathbb{C}) \otimes M_r(\mathbb{C})$ *be its reduced densities. Then*

$$S(\rho_{123}) + S(\rho_2) \leq S(\rho_{12}) + S(\rho_{23}) \tag{4.21}$$

or equivalently

$$\operatorname{Tr}\rho_{123}(\log\rho_{123} - \log\rho_{12} - \log\rho_{23} + \log\rho_2) \geq 0$$

where $\rho_{12} = \operatorname{Tr}_3\rho_{123}$, $\rho_{23} = \operatorname{Tr}_1\rho_{123}$ *and* $\rho_2 = \operatorname{Tr}_1\rho_{12}$.

There are many different proofs of the above theorem in the literature. The original proof was due to Lieb and Ruskai (see Lieb & Ruskai (1973)). Petz proved this theorem by using the mononicity of the relative entropy (Petz (1986)). In the following, we try to explain how to obtain the strong subadditivity inequality (4.21) by using the monoton icity of the relative entropy:

Proof. (based on a proof in Petz & Virosztek (2014)) By the monotonicity of relative entropy we have

$$S(\rho_{23}\|\rho_2 \otimes I_3) \leq S(\rho_{123}\|\rho_{12} \otimes I_3)$$

or equivalently

$$\operatorname{Tr}\rho_{23}(\log\rho_{23} - \log(\rho_2 \otimes I_3)) \leq \operatorname{Tr}\rho_{123}(\log\rho_{123} - \log(\rho_{12} \otimes I_3)). \tag{4.22}$$

Writing the left and the right hand side of (4.22) explicitly, we have

$$\operatorname{Tr}\rho_{23}\log\rho_{23} - \operatorname{Tr}\rho_{23}\log(\rho_2 \otimes I_3) \leq \operatorname{Tr}\rho_{123}\log\rho_{123} - \operatorname{Tr}\rho_{123}\log(\rho_{12} \otimes I_3)$$

which implies the following strong subadditivity inequality

$$-S(\rho_{23}) + S(\rho_2) \leq -S(\rho_{123}) + S(\rho_{12}).$$

□

Strong subadditivity is much deeper result than the subadditivity. Taking $M_n(\mathbb{C})$ to be one dimensional in the inequality (4.21) we obtain the subadditivity, that is

$$S(\rho_{12}) \leq S(\rho_1) + S(\rho_2). \tag{4.23}$$

Remember that in Chapter 2 we investigated a new type of inequality which is called partial strong subadditivity. Analogous to the classical entropy it is possible to define partial strong subaddivity for quantum entropies. In this case instead of probability distributions we have density matrices :

Theorem 4.3.6. *Let $\rho_{123} \in M_m(\mathbb{C}) \otimes M_n(\mathbb{C}) \otimes M_r(\mathbb{C})$ be a density matrix and $\rho_{12} \in M_m(\mathbb{C}) \otimes M_n(\mathbb{C})$, $\rho_2 \in \mathbb{M}_n$, $\rho_{23} \in \mathbb{M}_n \otimes \mathbb{M}_r$ be its reduced densities. Then*

$$\operatorname{Tr}_{12}\rho_{123}(\log\rho_{123} - \log\rho_{12} - \log\rho_{23} + \log\rho_2) \geq 0. \tag{4.24}$$

The inequality (4.24) can be called the partial strong subadditivity of von Neumann entropy.

Proof. This theorem is proved by Kim (2012). He uses the results of Effros (2009) based on convexity and functional analysis. For more details see Ruskai (2013) and Kim (2012). □

Remark 4.3.7. *In (4.24), the partial traces of the operator* $\rho_{123}(\log\rho_{123} - \log\rho_{12} - \log\rho_{23} + \log\rho_2)$ *with respect to the first component and the second component are not even self-adjoint at all.*

Remark 4.3.8. *In the inequality (4.24), taking* $M_n(\mathbb{C})$ *as one dimensional, one obtains the partial subadditivity, which is the matrix version of the classical partial subadditivity.*

CHAPTER FIVE
RESULTS ON QUANTUM TSALLIS ENTROPY

5.1 Definition And Some Basic Properties Of Quantum Tsallis Entropy

The one parameter extension of the von Neumann entropy is the quantum Tsallis entropy. It is quantum analogue of the classical Tsallis entropy.

Definition 5.1.1. *Let $\rho \in \mathbb{M}_n(\mathbb{C})$ be a density matrix. Then the quantum Tsallis entropy of ρ is defined by*

$$S_q(\rho) = -\operatorname{Tr} \rho \log_q(\rho). \tag{5.1}$$

By definition of the function $\log_q x$ (see definition 2.3.1), the above formula can be written as

$$S_q(\rho) = \frac{\operatorname{Tr}(\rho^q - \rho)}{1-q} \qquad (q > 1). \tag{5.2}$$

In an analogy to the classical case, taking the limit as $q \to 1$ in (5.1) we obtain the von Neumann entropy.

Note that taking the density operator as a diagonal matrix in the equality (5.1) we obtain the classical Tsallis entropy. Hence the classical Tsallis entropy is a special case of quantum Tsallis entropy.

For the sake of simplicity, from now on we will use the term ‘Tsallis entropy’ instead of the term ‘quantum Tsallis entropy’. In the following, we collect some important and useful facts of Tsallis entropy, most of which are analoguous to the von Neumann entropy:

Proposition 5.1.2. *Tsallis entropy is basis independent. If we choose the basis consisting of the eigenvectors of the density matrix ρ, we obtain the following* *formula:*

$$S_q(\rho) = \frac{\sum_i {\lambda_i}^q - 1}{1-q}.$$

Proposition 5.1.3. *Tsallis entropy is nonnegative and its maximum value is* $-\log_q \frac{1}{n}$. *It attains its maximum value at the completely mixed state* $\rho = (\frac{1}{n})I$, *where* $\rho \in \mathbb{M}_n(\mathbb{C})$.

It is possible to prove this property in many different ways. Here we will use the concavity of the function $x \longmapsto -x\log_q x$. The following proof is due to (Carlen, 2010):

Proof. Applying the function $-x\log_q x$ to the term $\sum_{i=1}^{n} \frac{1}{n}\lambda_i$ we obtain

$$\Big(-\sum_{i=1}^{n}\frac{1}{n}\lambda_i\Big)\log_q\Big(\sum_{i=1}^{n}\frac{1}{n}\lambda_i\Big).$$

Since the function $-x\log_q x$ is concave

$$\begin{aligned}\Big(-\sum_{i=1}^{n}\frac{1}{n}\lambda_i\Big)\log_q\Big(\sum_{i=1}^{n}\frac{1}{n}\lambda_i\Big) &\geq -\sum_{i=1}^{n}\frac{1}{n}\lambda_i\log_q\lambda_i\\ -\frac{1}{n}\log_q\frac{1}{n} &\geq -\frac{1}{n}\sum_{i=1}^{n}\lambda_i\log_q\lambda_i\\ -\log_q\frac{1}{n} &\geq S_q(\rho).\end{aligned}$$

Equality occurs at the points $\lambda_1 = \lambda_2 = ... = \lambda_n = \sum_{i=1}^{n}\frac{1}{n}\lambda_i$, which means that $\rho = (\frac{1}{n})I$. □

Proposition 5.1.4. *If the composite system is in a pure state* ρ_{12} *with reduced densities* ρ_1 *and* ρ_2, *then* $S_q(\rho_1) = S_q(\rho_2)$. *This property can be deduced from Remark 3.3.4.*

Recall from Proposition 2.3.5 that the classical Tsallis entropy is pseduo-additive for two independent random variables. Keeping in mind that the product states are quantum analogous to the independent random variables, one can see the following result immediately:

Proposition 5.1.5. *For a product state* $\rho_1 \otimes \rho_2$, *Tsallis entropy is pseudo-additive (Petz*

& Virosztek (2014)), that is

$$S_q(\rho_1 \otimes \rho_2) = S_q(\rho_1) + S_q(\rho_2) + (1-q)S_q(\rho_1)S_q(\rho_2)$$

where ρ_{12} is a density matrix with reduced densities ρ_1 and ρ_2.

In Chapter 4, we saw that the von Neumann entropy is an operator concave function of the density matrix ρ. The same is true for the Tsallis entropy:

Proposition 5.1.6. *Tsallis entropy is operator concave. That is*

$$S_q(t\rho_1 + (1-t)\rho_2) \geq tS_q(\rho_1) + (1-t)S_q(\rho_2) \tag{5.3}$$

where $0 \leq t \leq 1$ and $\rho_1, \rho_2 \in \mathbb{M}_n(\mathbb{C})$ are density matrices.

Proof. The inequality (5.3) follows from the concavity of the function $x \longmapsto -x \log_q x$ and from theorem 4.1.9. □

5.2 Subadditivity Of Quantum Tsallis Entropy

One of the remarkable results related to the Tsallis entropy is the following theorem:

Theorem 5.2.1. *Let $\rho_{12} \in \mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$ be a density matrix and let $\rho_1 \in \mathbb{M}_m(\mathbb{C})$ and $\rho_2 \in \mathbb{M}_n(\mathbb{C})$ be its reduced densities. For $q > 1$*

$$S_q(\rho_{12}) \leq S_q(\rho_1) + S_q(\rho_2) \tag{5.4}$$

From Definition 5.1.1, the inequality (5.4) can be written as

$$\operatorname{Tr} \rho_{12} \log_q \rho_{12} \geq \operatorname{Tr} \rho_1 \log_q \rho_1 + \operatorname{Tr} \rho_2 \log_q \rho_2$$

or equivalently

$$\operatorname{Tr} \rho_{12}(\log_2 \rho_{12} - \log_2 \rho_1 \otimes I - I \otimes \log_2 \rho_2) \geq 0. \tag{5.5}$$

The inequality (5.4) is called the subadditivity of Tsallis entropy and it is proved by Audenaert (2007). Before giving the details of his proof it will be useful to give some definitions.

For each $A \in \mathbb{M}_n(\mathbb{C})$, A^*A is a positive definite matrix and the absolute value of A is defined by $|A| = (A^*A)^{1/2}$. The eigenvalues of $|A|$ are called the singular values of the matrix A.

Definition 5.2.2. *(Bhatia (2013), Hiai & Petz (2014)) For any matrix A in* $\mathbb{M}_n(\mathbb{C})$ *Schatten q-norm of A is defined by*

$$\|A\|_q = \begin{cases} \left(\sum_{j=1}^n s_j(A)^q\right)^{1/q} = (\operatorname{Tr}|A|^q)^{1/q}, & (1 \leq q < \infty) \\ s_1(A) = \|A\|, & q = \infty \end{cases}$$

where $s_j(A)$ *are the singular values of A (or the eigenvalues of* $|A|$*). In particular,* $\|A\|_1 = \operatorname{Tr}|A|$ *is called the trace norm (or Schatten 1-norm),* $\|A\|_2 = (\operatorname{Tr} A^*A)^{1/2}$ *is called the Hilbert Schmidt norm and* $\|A\|_\infty = \|A\| = s_1(A)$ *is called the operator norm.*

For $x \in \mathbb{R}$ consider the following functions

$$f_+(x) = \max\{x, 0\},$$
$$f_-(x) = \max\{-x, 0\}.$$

One can easily see that

$$f_+(x), f_-(x) \geq 0, \qquad f_+(x) - f_-(x) = x \quad \text{and} \quad f_+(x) + f_-(x) = |x|. \tag{5.6}$$

For any self adjoint matrix B it is possible to define the following functions

$$B_+ = f_+(B), \qquad B_- = f_-(B),$$

B_+ is called the positive part and B_- is called the negative part of the matrix B. By (5.6) we have

$$B_+, B_- \geq 0, \quad B_+ - B_- = B \quad \text{and} \quad B_+ + B_- = |B|$$

and $B_+ - B_- = B$ is called the Jordan decomposition of the matrix B. Before proving the subadditivity, Audenaert first proved the following lemma:

Lemma 5.2.3. *Let $X \in \mathbb{M}_m(\mathbb{C})$ and $Y \in \mathbb{M}_n(\mathbb{C})$ be positive definite matrices such that $\|X\|_q, \|Y\|_q \leq 1$. Then the following inequality*

$$\|(X \otimes I_n + I_m \otimes Y - I_m \otimes I_n)_+\|_q \leq 1 \tag{5.7}$$

holds for $q \geq 1$.

Proof. *(Audenaert (2007), Hiai & Petz (2014))*
We will prove the case $\|X\|_q, \|Y\|_q = 1$. The case $\|X\|_q, \|Y\|_q < 1$ follows immediately. Let x_i and y_j be the elements of the spectrum of X and Y respectively, where $i = 1, ..., m$ and $j = 1, ..., n$. Since $\|X\|_q, \|Y\|_q = 1$ we have

$$\sum_{i=1}^{m} x_i^q = 1 \quad \text{and} \quad \sum_{j=1}^{n} y_j^q = 1.$$

We will prove the following inequality

$$\|(X \otimes I_n + I_m \otimes Y - I_m \otimes I_n)_+\|_q^q = \sum_{i,j} ((x_i - y_j - 1)_+)^q \leq 1.$$

The function $a \longmapsto (a + b - 1)_+$ is a convex function of a for any $b \in \mathbb{R}$. Let us define a vector valued function f as

$$f(a) = (a + y - 1)_+$$

where $y = (y_1, y_2, ..., y_n)$. The function f is also convex in a. That is

$$f(ta_1 + (1-t)a_2) \leq tf(a_1) + (1-t)f(a_2)$$

or equivalently

$$(ta_1 + (1-t)a_2 + y - 1)_+ \leq t(a_1 + y - 1)_+ + (1-t)(a_2 + y - 1)_+.$$

By the monotonicity and convexity of the l_q norm we have

$$\begin{aligned} \|(ta_1 + (1-t)a_2 + y - 1)_+\|_q &\leq \|t(a_1 + y - 1)_+ + (1-t)(a_2 + y - 1)_+\|_q \\ &\leq t\|(y + a - 1)_+\|_q + (1-t)\|(y + b - 1)_+\|_q. \end{aligned}$$

Hence the function

$$g(a) = \|(a + y - 1)_+\|_q = \left(\sum_j (a + y_j - 1)_+^q \right)^{1/q}$$

is convex in a. Moreover $g(0) = 0$ and $g(1) = 1$ which means that $g(a) \leq a$ for $0 \leq a \leq 1$. Since $0 \leq x_i \leq 1$ we have

$$g(x_i) = \left(\sum_j (x_i + y_j - 1)_+^q \right)^{1/q} \leq x_i$$

and summing over the index i we obtain

$$\sum_i g(x_i)^q = \sum_i \sum_j (x_i + y_j - 1)_+^q \leq \sum_{i=1}^m x_i^q = 1$$

which proves the desired inequality. □

It is possible to prove Theorem 5.2.1 by using the above lemma. The following proof is due to Hiai & Petz (2014):

Proof. *(of Theorem 5.2.1)* By (5.2) the inequality (5.4) can be written as

$$\operatorname{Tr}\rho_1^q + \operatorname{Tr}\rho_2^q \leq \operatorname{Tr}\rho_{12}^q + 1$$

or equivalently

$$\|\rho_1\|_q^q + \|\rho_2\|_q^q \leq \|\rho_{12}\|_q^q + 1. \tag{5.8}$$

To prove (5.8) first we will prove the following inequality:

$$\|\rho_1\|_q + \|\rho_2\|_q \leq \|\rho_{12}\|_q + 1. \tag{5.9}$$

Let $q' \in \mathbb{R}$ be such that $1/q + 1/q' = 1$. Then for any positive matrix A we have

$$\|A\|q = \max\left\{\operatorname{Tr} AB : \ B \geq 0, \ \|B\|_{q'} \leq 1\right\}$$

(This property is called the duality of Schatten $q-$norm.) Hence

$$\|\rho_1\|_q = \operatorname{Tr} X\rho_1 \quad \text{and} \quad \|\rho_2\|_q = \operatorname{Tr} Y\rho_2$$

for some positive matrices X, Y such that $\|X\|_{q'}, \|Y\|_{q'} \leq 1$. Then by Lemma 5.2.3 we have

$$\|(X \otimes I_n + I_m \otimes Y - I_m \otimes I_n)_+\|_{q'} \leq 1.$$

From the Weyl's monotonicity principle we can find a matrix $Z \geq 0$ with $\|Z\|_{q'} = 1$ such that

$$Z \geq X \otimes I_n + I_m \otimes Y - I_m \otimes I_n.$$

Then

$$\rho_{12}^{1/2}(Z + I_m \otimes I_n)\rho_{12}^{1/2} \geq \rho_{12}^{1/2}(X \otimes I_n + I_m \otimes Y)\rho_{12}^{1/2}$$

and

$$\begin{aligned}\mathrm{Tr}[Z\rho_{12}] + 1 &\geq \mathrm{Tr}[(X \otimes I_n + I_m \otimes Y)\rho_{12}] \\ &= \mathrm{Tr}[X\rho_1] + \mathrm{Tr}[Y\rho_2] \\ &= ||\rho_1||_q + ||\rho_2||_q.\end{aligned}$$

Since $||\rho_{12}||_q \geq \mathrm{Tr}[Z\rho_{12}]$ we have

$$||\rho_{12}||_q + 1 \geq ||\rho_1||_q + ||\rho_2||_q$$

which proves (5.9).

Now we prove the inequality (5.8). To do this we use the function $f(x, y) = x^q + y^q$ for $q > 1$. Let us look at the maximum value of f in the domain

$$D = \Big\{(x, y) : \quad 0 \leq x \leq 1, \quad 0 \leq y \leq 1, \quad x + y \leq 1 + ||\rho_{12}||_q\Big\}.$$

Since f is a convex function of x and y we examine only the extreme points $(0,0)$, $(1,0)$, $(0,1)$, $(1, ||\rho_{12}||_q)$, $(||\rho_{12}||_q, 1)$. The value of f at these points are

$$\begin{gathered} f(0,0) = 0, \quad f(1,0) = 1, \quad f(0,1) = 1 \\ f(1, ||\rho_{12}||_q) = 1 + ||\rho_{12}||_q^q, \quad f(||\rho_{12}||_q, 1) = 1 + ||\rho_{12}||_q^q. \end{gathered}$$

Hence $f(x, y) \leq 1 + ||\rho_{12}||_q^q$. But the point $(||\rho_1||_q, ||\rho_2||_q) \in D$, so we have $||\rho_1||_q^q + ||\rho_2||_q^q \leq ||\rho_{12}||_q^q + 1$. □

We saw in Section 4.3 that von Neumann entropy is strongly subadditive. The following example shows that this is not true for Tsallis entropy:

Example 5.2.4. *Let*

$$\rho_{123} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.4 & 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.4 & 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

It is easy to see that $\rho_{123} \in M_2(\mathbb{C}) \otimes M_2(\mathbb{C}) \otimes M_2(\mathbb{C})$ *is a positive matrix with a trace equals to* 1, *that is,* ρ_{123} *is a density matrix.*

Then the corresponding reduced densities are

$$\rho_{12} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \rho_{23} = \begin{bmatrix} 0.1 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.4 \end{bmatrix}, \quad \rho_2 = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}.$$

For $q = 2$ *we have* $S_2(\rho_{123}) = 0.32,\ S_2(\rho_{12}) = 0,\ S_2(\rho_2) = 0.5,\ S_2(\rho_{23}) = 0.66$ *which shows that*

$$S_2(\rho_{123}) + S_2(\rho_2) > S_2(\rho_{12}) + S_2(\rho_{23}).$$

Therefore Tsallis entropy does not satisfy the SSA property!

5.3 Partial Subadditivity Of Quantum Tsallis Entropy

In Chapter 2 we made a diagram to show the relations between the strong subadditivity (SSA), partial strong subadditivity (PSSA), subadditivity (SA) and partial subadditivity (PSA) of classical entropies. The same diagram applies to quantum analogues of these entropies:

$$
\begin{array}{ccc}
PSSA & \Longrightarrow & SSA \\
\Downarrow & & \Downarrow \\
PSA & \Longrightarrow & SA
\end{array}
$$

Remark 5.3.1. *Recall from Example 5.2.4 that quantum Tsallis entropy is not strongly subadditive and the above diagram shows that it can not be partially strong subadditive. The table 1.1 of the introduction part (see page 3) summarizes all the cases we have seen so far.*

At the end of Chapter 4 we gave a theorem of partial strong subadditivity of von Neumann entropy. We observe by the above remark that the partial strong subadditivity is not in the case for the Tsallis entropy.

In Besenyei & Petz (2013), the following inequality (which is related to the SSA of quantum Tsallis entropy) was conjectured:

If ρ_{12} is a density operator in $\mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{K})$ with reduced densities ρ_1 and ρ_2, one has

$$\mathrm{Tr}(T \otimes I_2)\rho_{12}(\log_q \rho_{12} - \log_q \rho_1 \otimes I_2 - I_1 \otimes \log_q \rho_2) \geq 0 \tag{5.10}$$

whenever $T \geq 0$ and $q > 1$.

The inequality (5.10) is a kind of partial subadditivity of Tsallis entropy and is quite similar to the inequality (5.5).

The above conjecture was proved in the same article for the following cases:

1. $\rho_{12} = \rho_1 \otimes \rho_2$ (The case of product states), and
2. The case $m = n = 2$ when $q = 2$. ($\dim(\mathcal{H}) = m$ and $\dim(\mathcal{K}) = n$)

The authors give the following example for the second case:

Example 5.3.2. *Let ρ_{12} be a density matrix in $\mathbb{M}_2(\mathbb{C}) \otimes \mathbb{M}_2(\mathbb{C})$ and $T \in \mathbb{M}_2(\mathbb{C})$ be a positive semidefinite matrix such that*

$$\rho_{12} = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix}, \qquad T = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

with the reduced densities

$$\rho_1 = \begin{bmatrix} TrA & TrB \\ TrB^* & TrC \end{bmatrix}, \qquad \rho_2 = A + C.$$

The key point of the proof is the following inequality:

$$\operatorname{Tr} BB^* - \operatorname{Tr} B \operatorname{Tr} B^* - \operatorname{Tr} AC + \operatorname{Tr} A \operatorname{Tr} C \geq 0.$$

This inequality is proved in Besenyei (2013):

Theorem 5.3.3. ***(Trace Inequality For Positive Block Matrices):*** *Let A, B, C be $n \times n$ matrices with complex entries and the block matrix $\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \in \mathbb{M}_{2n}(\mathbb{C})$ be positive semidefinite. Then*

$$\operatorname{Tr} AC - \operatorname{Tr} BB^* \leq \operatorname{Tr} A \operatorname{Tr} C - \operatorname{Tr} B \operatorname{Tr} B^*. \tag{5.11}$$

We restate Besenyei's proof :

Proof. Since trace function is unitarily invariant we can assume that the matrix A is diagonal. By simple calculations, (5.11) reduces to

$$\sum_{i=1}^{n} a_{ii}c_{ii} - \sum_{1\leq i,j\leq n} |b_{ij}|^2 \leq \sum_{i=1}^{n} a_{ii} \sum_{i=1}^{n} c_{ii} - \left|\sum_{i=1}^{n} b_{ii}\right|^2$$

or equivalently

$$2\sum_{i>j} \mathfrak{Re}(b_{ii}\overline{b}_{jj}) - \sum_{i\neq j} |b_{ij}|^2 \leq \sum_{i>j} (a_{ii}c_{jj} + a_{jj}c_{ii}).$$

If we show that

$$2\sum_{i>j} \mathfrak{Re}(b_{ii}\overline{b}_{jj}) \leq a_{ii}c_{jj} + c_{jj}c_{ii}$$

then we are done.

Since $\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \geq 0$, then all principal minors of ρ_{12} are nonnegative. Hence the determinant of the matrix $\begin{bmatrix} a_{ii} & b_{ii} \\ \overline{b_{ii}} & c_{ii} \end{bmatrix}$ is nonnegative for all $i = 1, 2, .., n$. So

$$a_{ii}c_{ii} - |b_{ii}|^2 \geq 0.$$

Similarly,

$$a_{22}c_{22} - |b_{22}|^2 \geq 0.$$

By the arithmetic and geometric mean inequality we have

$$a_{ii}c_{jj} + a_{jj}c_{ii} \geq 2\sqrt{a_{ii}c_{jj}a_{jj}c_{ii}} \geq 2\sqrt{|b_{ii}|^2|b_{jj}|^2} \geq 2\mathfrak{Re}(b_{ii}b_{jj})$$

and this proves the inequality (5.11). □

5.4 Some Obtained Results Related To The Quantum Tsallis Entropy

In the following theorem we present a proof of the inequality (5.10) for arbitrary m and n when $q = 2$. This completes Table 1.1 when $q = 2$.

Theorem 5.4.1. *Let $\rho_{12} \in \mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$ be a density matrix, $\rho_1 \in \mathbb{M}_m(\mathbb{C})$, $\rho_2 \in \mathbb{M}_n(\mathbb{C})$ be its reduced densities and $0 \leq T \in \mathbb{M}_m(\mathbb{C})$. Then*

$$\mathrm{Tr}\,(T \otimes I_2)\rho_{12}(\log_2 \rho_{12} - \log_2 \rho_1 \otimes I - I \otimes \log_2 \rho_2) \geq 0$$

or equivalently

$$\mathrm{Tr}\,T\rho_1 + \mathrm{Tr}(T \otimes I_2)\rho_{12}^2 - \mathrm{Tr}\,T\rho_1^2 - \mathrm{Tr}(T \otimes \rho_2)\rho_{12} \geq 0. \tag{5.12}$$

Before starting the proof it will be useful to give the following remark:

Remark 5.4.2. *The inequality*

$$\mathrm{Tr}\,T\rho_1 + \mathrm{Tr}(T \otimes I_2)\rho_{12}^2 - \mathrm{Tr}\,T\rho_1^2 - \mathrm{Tr}(T \otimes \rho_2)\rho_{12} \geq 0$$

is unitarily invariant. That is, if $UTU^* = D$, where D is diagonal and U is unitary and if $\rho'_{12} = (U \otimes U)\rho_{12}(U^* \otimes U^*)$ with reduced densities $\rho'_1 = U\rho_1 U^*$, $\rho'_2 = U\rho_2 U^*$, then we have

$$\begin{aligned}&\mathrm{Tr}(UTU^*U\rho_1U^*) + \mathrm{Tr}(U \otimes U)(T \otimes I)(U^* \otimes U^*)[(U \otimes U)\rho_{12}(U^* \otimes U^*)]^q \\ &= \mathrm{Tr}\,T\rho_1 + \mathrm{Tr}(T \otimes I_2)\rho_{12}^q\end{aligned}$$

and

$$\begin{aligned}&\mathrm{Tr}\,UTU^*(U\rho_1U^*)^q + \mathrm{Tr}[(U \otimes U)(T \otimes \rho_2^{q-1})(U^* \otimes U^*)(U \otimes U)\rho_{12}(U^* \otimes U^*)] \\ &= \mathrm{Tr}\,T\rho_1^q + \mathrm{Tr}(T \otimes \rho_2^{q-1})\rho_{12}.\end{aligned}$$

Note that $\rho'_{12} = (U \otimes U)\rho_{12}(U^* \otimes U^*)$ is a density matrix and $U\rho_1 U^*, U\rho_2 U^*$ are reduced densities of ρ'_{12}. Therefore in (5.12) we may assume that T is a diagonal matrix with nonnegative diagonal elements.

Proof. (of Theorem 5.4.1) We will prove the inequality (5.12). One can write the density matrix ρ_{12} as a block matrix:

$$\rho_{12} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \dots & A_{1m} \\ A_{12}^* & A_{22} & A_{23} & \dots & A_{2m} \\ A_{13}^* & A_{23}^* & A_{33} & \dots & A_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{1m}^* & A_{2m}^* & A_{3m}^* & \dots & A_{mm} \end{bmatrix}.$$

Then

$$\rho_1 = \begin{bmatrix} \operatorname{Tr} A_{11} & \operatorname{Tr} A_{12} & \operatorname{Tr} A_{13} & \dots & \operatorname{Tr} A_{1m} \\ \operatorname{Tr} A_{12}^* & \operatorname{Tr} A_{22} & \operatorname{Tr} A_{23} & \dots & \operatorname{Tr} A_{2m} \\ \operatorname{Tr} A_{13}^* & \operatorname{Tr} A_{23}^* & \operatorname{Tr} A_{33} & \dots & \operatorname{Tr} A_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \operatorname{Tr} A_{1m}^* & \operatorname{Tr} A_{2m}^* & \operatorname{Tr} A_{3m}^* & \dots & \operatorname{Tr} A_{mm} \end{bmatrix}, \qquad \rho_2 = A_{11} + A_{22} + \dots + A_{mm}.$$

By Remark 5.4.2 we may assume that T is a diagonal matrix with nonnegative diagonal entries $t_{11}, t_{22}, \dots, t_{mm}$. After some calculations we obtain the following formulas:

$$\operatorname{Tr} T\rho_1 = \sum_{i=1}^{m} t_{ii} \operatorname{Tr} A_{ii} \tag{5.13}$$

$$\operatorname{Tr}(T \otimes I)\rho_{12}{}^2 = \sum_{i=1}^{m} t_{ii} \operatorname{Tr}(A_{ii}^2) + \sum_{j>i} (t_{ii} + t_{jj}) \operatorname{Tr} A_{ij}^* A_{ij} \tag{5.14}$$

$$\operatorname{Tr} T\rho_1^2 = \sum_{i=1}^{m} t_{ii}(\operatorname{Tr} A_{ii})^2 + \sum_{j>i}(t_{ii}+t_{jj})|\operatorname{Tr} A_{ij}|^2 \tag{5.15}$$

$$\operatorname{Tr}(T\otimes\rho_2)\rho_{12} = \sum_{i=1}^{m} t_{ii}\operatorname{Tr}(A_{ii}^2) + \sum_{j>i}(t_{ii}+t_{jj})\operatorname{Tr} A_{ii}A_{jj} \tag{5.16}$$

where $i, j = 1, ..., m$. By using formulas (5.13), (5.14), (5.15), (5.16); the left hand side of the inequality (5.12) becomes

$$\sum_{j>i}(t_{ii}+t_{jj})\Big[\operatorname{Tr} A_{ij}^{*}A_{ij} - |\operatorname{Tr} A_{ij}|^2 - \operatorname{Tr} A_{ii}A_{jj}\Big] + \sum_{i=1}^{m} t_{ii}\Big[\operatorname{Tr} A_{ii} - (\operatorname{Tr} A_{ii})^2\Big].$$

Using the formula $\operatorname{Tr} A_{11} + \operatorname{Tr} A_{22} + ... + \operatorname{Tr} A_{mm} = 1$ we obtain $\operatorname{Tr} A_{ii} - (\operatorname{Tr} A_{ii})^2 = \sum_{\substack{j=1\\ i\neq j}}^{m} \operatorname{Tr} A_{ii}\operatorname{Tr} A_{jj}$, for all i. Then we substitute this formula into the sum

$$\sum_{i=1}^{m} t_{ii}\Big[\operatorname{Tr} A_{ii} - (\operatorname{Tr} A_{ii})^2\Big]$$

and we obtain

$$\sum_{i=1}^{m} t_{ii}\Big[\sum_{\substack{j=1\\ i\neq j}}^{m} \operatorname{Tr} A_{ii}\operatorname{Tr} A_{jj}\Big] = \sum_{j>i}(t_{ii}+t_{jj})\operatorname{Tr} A_{ii}\operatorname{Tr} A_{jj}.$$

Then,

$$\operatorname{Tr} T\rho_1 + \operatorname{Tr}(T\otimes I_2)\rho_{12}^2 - \operatorname{Tr} T\rho_1^2 - \operatorname{Tr}(T\otimes\rho_2)\rho_{12}$$

is equal to

$$\sum_{j>i}(t_{ii}+t_{jj})\Big[\operatorname{Tr} A_{ij}^{*}A_{ij} - |\operatorname{Tr} A_{ij}|^2 - \operatorname{Tr} A_{ii}A_{jj} + \operatorname{Tr} A_{ii}\operatorname{Tr} A_{jj}\Big].$$

Now we will show that

$$(t_{ii}+t_{jj})\Big[\operatorname{Tr} A_{ij}^*A_{ij} - |\operatorname{Tr} A_{ij}|^2 - \operatorname{Tr} A_{ii}A_{jj} + \operatorname{Tr} A_{ii}\operatorname{Tr} A_{jj}\Big] \geq 0, \quad j>i.$$

Since $t_{11}, t_{22}, ..., t_{mm} \geq 0$ then $t_{ii}+t_{jj} \geq 0$ for any i, j. Also since $\rho_{12} \geq 0$, the principal submatrix $\begin{bmatrix} A_{ii} & A_{ij} \\ A_{ij}^* & A_{jj} \end{bmatrix} \geq 0$. The proof of the trace inequality

$$\operatorname{Tr} A_{ij}^*A_{ij} - |\operatorname{Tr} A_{ij}|^2 - \operatorname{Tr} A_{ii}A_{jj} + \operatorname{Tr} A_{ii}\operatorname{Tr} A_{jj} \geq 0, \quad j>i$$

follows from Theorem 5.3.3. Hence,

$$\sum_{j>i}(t_{ii}+t_{jj})\Big[\operatorname{Tr} A_{ij}^*A_{ij} - |\operatorname{Tr} A_{ij}|^2 - \operatorname{Tr} A_{ii}A_{jj} + \operatorname{Tr} A_{ii}\operatorname{Tr} A_{jj}\Big] \geq 0.$$

This completes the proof. □

Corollary 5.4.3. *In the above theorem putting $T = |\psi\rangle\langle\psi|$, where $|\psi\rangle \in \mathbb{C}^m$ is any vector, we obtain the following inequality*

$$\langle\psi|\operatorname{Tr}_2 \rho_{12}(\log_2\rho_{12} - \log_2\rho_1 \otimes I - I \otimes \log_2\rho_2)|\psi\rangle \geq 0$$

which means that the operator

$$\operatorname{Tr}_2 \rho_{12}(\log_2\rho_{12} - \log_2\rho_1 \otimes I - I \otimes \log_2\rho_2)$$

is positive semidefinite on $\mathbb{C}^m$.

This corollary shows that the partial trace of the operator $\rho_{12}(\log_2\rho_{12} - \log_2\rho_1 \otimes I - I \otimes \log_2\rho_2)$ is positive semidefinite on $\mathbb{C}^m$. In fact this operator has some other properties not only for $q=2$ but also for $q \in (1,\infty)$. Hence in the rest of this section we concentrate on the operator

$$\rho_{12}(\log_q\rho_{12} - \log_q\rho_1 \otimes I - I \otimes \log_q\rho_2) \qquad (q>1) \tag{5.17}$$

which is equivalent to

$$\frac{1}{q-1}[\rho_{12}^{q}-\rho_{12}(I_1\otimes\rho_2^{q-1})-\rho_{12}(\rho_1^{q-1}\otimes I_2)+\rho_{12}] \tag{5.18}$$

by (5.2).

Lemma 5.4.4. *Partial traces*

$$\operatorname{Tr}_1\rho_{12}(\log_q\rho_{12}-\log_q\rho_1\otimes I-I\otimes\log_q\rho_2)\in\mathbb{M}_m(\mathbb{C}) \tag{5.19}$$

$$\operatorname{Tr}_2\rho_{12}(\log_q\rho_{12}-\log_q\rho_1\otimes I-I\otimes\log_q\rho_2)\in\mathbb{M}_n(\mathbb{C}) \tag{5.20}$$

of the operator (5.17) are Hermitian.

We need the following proposition to prove the lemma:

Proposition 5.4.5. *Let* $\rho_{12}\in\mathbb{M}_m(\mathbb{C})\otimes\mathbb{M}_n(\mathbb{C})$ *be a density matrix,* $\rho_1\in\mathbb{M}_m(\mathbb{C})$*,* $\rho_2\in\mathbb{M}_n(\mathbb{C})$ *be its reduced densities and* $q\in(1,\infty)$*. Then the operators* $\operatorname{Tr}_2\rho_{12}(I_1\otimes\rho_2^{q-1})$*,* $\operatorname{Tr}_1\rho_{12}(\rho_1^{q-1}\otimes I_2)$ *are positive semidefinite.*

Proof. We will show that

$$\langle x,\operatorname{Tr}_2\rho_{12}(I_1\otimes\rho_2^{q-1})x\rangle\geq 0,\qquad \forall x\in\mathbb{C}^m.$$

Let $\{f_j\}_{j=1}^n$ be an orthonormal basis of $\mathbb{C}^n$. By Definition 3.3.8 we have

$$\langle x,\operatorname{Tr}_2\rho_{12}(I_1\otimes\rho_2^{q-1})x\rangle=\sum_{j=1}^{n}\langle(x\otimes f_j),\rho_{12}(I_1\otimes\rho_2^{q-1})(x\otimes f_j)\rangle.$$

This definition is independent of the choice of the orthonormal basis. Thus we may assume that the basis $\{f_j\}_{j=1}^n$ consists of the eigenvectors of the density operator ρ_2.

Writing $\rho_2 = \sum_j \lambda_j |f_j\rangle\langle f_j|$ we have

$$\begin{aligned}\langle x, \mathrm{Tr}_2\, \rho_{12}(I_1 \otimes \rho_2^{q-1})x\rangle &= \sum_{j=1}^{n} \langle (x\otimes f_j), \rho_{12}(I_1 \otimes \rho_2^{q-1})(x\otimes f_j)\rangle \\ &= \sum_{j=1}^{n} \langle (x\otimes f_j), \rho_{12}(x \otimes \rho_2^{q-1} f_j)\rangle \\ &= \sum_{j=1}^{n} \lambda_j^{q-1} \langle (x\otimes f_j), \rho_{12}(x\otimes f_j)\rangle\end{aligned}$$

where $0 \leq \lambda_j \in \mathrm{sp}(\rho_2)$ for all j. And $\langle (x\otimes f_j), \rho_{12}(x\otimes f_j)\rangle \geq 0$ for all j since ρ_{12} is positive. Hence

$$\langle x, \mathrm{Tr}_2\, \rho_{12}(I_1 \otimes \rho_2^{q-1})x\rangle \geq 0, \qquad \forall x \in \mathbb{C}^m.$$

The positivity of the operator $\mathrm{Tr}_1\, \rho_{12}(\rho_1^{q-1} \otimes I_2)$ can be proved in an analogous way. □

Proof. (of Lemma (5.4.4)) In order to show the assertion, we will prove that the partial traces of the operator (5.18) are Hermitian. Hence we will show that the operators

$$\mathrm{Tr}_1(\rho_{12}^q - \rho_{12}(I_1 \otimes \rho_2^{q-1}) - \rho_{12}(\rho_1^{q-1} \otimes I_2) + \rho_{12}) \tag{5.21}$$

$$\mathrm{Tr}_2(\rho_{12}^q - \rho_{12}(I_1 \otimes \rho_2^{q-1}) - \rho_{12}(\rho_1^{q-1} \otimes I_2) + \rho_{12}) \tag{5.22}$$

are Hermitian.

We will only prove that the operator (5.22) is Hermitian, the Hermitianness of (5.21) can be proved in an anologous way. By the linearity of partial trace, the operator (5.22) is equal to

$$\mathrm{Tr}_2\, \rho_{12}^q - \mathrm{Tr}_2\, \rho_{12}(I_1 \otimes \rho_2^{q-1}) - \rho_1^q + \rho_1. \tag{5.23}$$

In (5.23) the operators $\mathrm{Tr}_2\, \rho_{12}^q$, ρ_1^q and ρ_1 are all Hermitian. So, we only have to show that $\mathrm{Tr}_2\, \rho_{12}(I_1 \otimes \rho_2^{q-1})$ is Hermitian. But in the above proposition we proved that

$\mathrm{Tr}_2\,\rho_{12}(I_1 \otimes \rho_2^{q-1}) \geq 0$. Hence the operator (5.22) is Hermitian. $\square$

CHAPTER SIX
CONCLUSION AND OPEN PROBLEMS

In this thesis we investigate the inequality $\mathrm{Tr}(T \otimes I_2)\rho_{12}(\log_q \rho_{12} - \log_q \rho_1 \otimes I_2 - I_1 \otimes \log_q \rho_2) \geq 0$, where ρ_{12} is a density matrix and $0 \leq T \in \mathbb{M}_m(\mathbb{C})$. This inequality was conjectured by Besenyei and Petz in 2013, where it was proved to hold for the density matrices in $\mathbb{M}_2(\mathbb{C}) \otimes \mathbb{M}_2(\mathbb{C})$ and for $q = 2$. Here we prove this inequality for the density matrices in $\mathbb{M}_m(\mathbb{C}) \otimes \mathbb{M}_n(\mathbb{C})$. We also obtain some new inequalities related to the operators (matrices) in this inequality.

The most important problem that remains is to understand the case of $q > 1$ with $q \neq 2$. It seems that some new ideas are needed for a general solution. Having this in mind we performed some numerical computations in Wolfram Mathematica 12. These examples suggest that the operators $\mathrm{Tr}_2\, \rho_{12}(\log_q \rho_{12} - \log_q \rho_1 \otimes I - I \otimes \log_q \rho_2)$ and $\mathrm{Tr}_1\, \rho_{12}(\log_q \rho_{12} - \log_q \rho_1 \otimes I - I \otimes \log_q \rho_2)$ are not only Hermitian but also positive. If it is true, this would imply the partial subadditivity of the Tsallis entropy. Hence our future work will be to investigate this claim.

REFERENCES

Araki, H. (1976). Relative entropy of states of von neumann algebras. *Publications of the Research Institute for Mathematical Sciences*, *11*(3), 809–833.

Audenaert, K. M. (2007). Subadditivity of q-entropies for q> 1. *Journal of mathematical physics*, *48*(8), 083507.

Besenyei, Á. (2013). A note on a trace inequality for positive block matrices. *Athens, Greece*.

Besenyei, Á., & Petz, D. (2013). Partial subadditivity of entropies. *Linear Algebra and its Applications*, *439*(10), 3297–3305.

Bhatia, R. (2009). *Positive definite matrices*. Princeton University Press.

Bhatia, R. (2013). *Matrix analysis*, volume 169. Springer Science & Business Media.

Carlen, E. (2010). Trace inequalities and quantum entropy: an introductory course. *Entropy And The Quantum*, *529*, 73–140.

Effros, E. G. (2009). A matrix convexity approach to some celebrated quantum inequalities. *Proceedings of the National Academy of Sciences*, *106*(4), 1006–1008.

Furuci, S. (2006). Information theoretical properties of tsallis entropy. *J. Math. Phys*, *47*, 023302–1.

Hiai, F. (2017). Matrix analysis and quantum information. *Sugaku Expositions*, *30*(1), 35–63.

Hiai, F., & Petz, D. (2014). *Introduction to matrix analysis and applications*. Springer Science & Business Media.

Jenčová, A., & Ruskai, M. B. (2010). A unified treatment of convexity of relative entropy and related trace functions, with conditions for equality. *Reviews in Mathematical Physics*, *22*(09), 1099–1121.

Kim, I. H. (2012). Operator extension of strong subadditivity of entropy. *Journal of mathematical physics*, *53*(12), 122204.

Lieb, E. H., & Ruskai, M. B. (1973). Proof of the strong subadditivity of quantum-mechanical entropy. *Les rencontres physiciens-mathématiciens de Strasbourg-RCP25*, *19*, 36–55.

McMahon, D. (2007). *Quantum computing explained*. John Wiley & Sons.

Nielsen, M. A., & Chuang, I. L. (2010). *Quantum Computation and Quantum Information*. Cambridge University Press.

Nielsen, M. A., & Petz, D. (2004). A simple proof of the strong subadditivity inequality. *arXiv preprint quant-ph/0408130*.

Petz, D. (1986). Quasi-entropies for finite quantum systems. *Reports on mathematical physics*, *23*(1), 57–65.

Petz, D. (2007). *Quantum information theory and quantum statistics*. Springer Science & Business Media.

Petz, D. (2008). Information and its measures. In *Quantum Information Theory and Quantum Statistics* (25–51). Springer.

Petz, D., & Virosztek, D. (2014). Some inequalities for quantum tsallis entropy related to the strong subadditivity. *arXiv preprint arXiv:1403.7062*.

Ruskai, M. B. (2013). Remarks on kim's strong subadditivity matrix inequality: Extensions and equality conditions. *Journal of Mathematical Physics*, *54*(10), 102202.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379–423.

Von Neumann, J. (2013). *Mathematische grundlagen der quantenmechanik*, volume 38. Springer-Verlag.