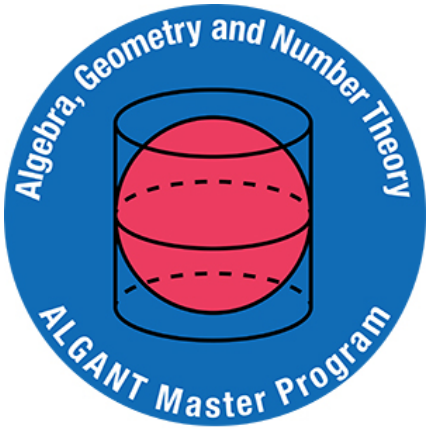


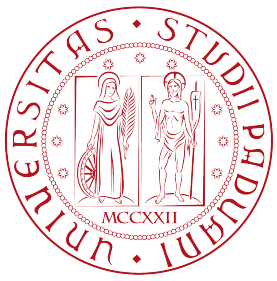
ALGANT Master Thesis in Mathematics

The RRDPS Encoding and Entropic Uncertainty Relations for Single Measurements

---

Guillermo Carlo Nuñez Ponasso

Advised by Prof. Serge Fehr



Università degli Studi di Padova

Universiteit Leiden

Academic year 2017/2018
25 June 2018

Acknowledgements

First of all, I would like to thank my advisor, Serge Fehr, for his constant guidance during the writing of this thesis. I found his suggestions very useful and the work environment at CWI very pleasant.
Also I would like to thank the members of the reading committee in Leiden, Peter Bruin, Bart de Smit and Marco Streng for their excellent review of this document.

I want to thank as well, my fellow ALGANT students at Padova and Leiden, with whom I shared really nice moments. Specially I would like to thank Matteo Pintonello and Sergej Monavari for being my second family in Leiden. Grazie coinquilini!

To the wonderful people of the Board Games Club at Leiden, and to his amazing organiser Uğur Derin, I would also like to say thank you, the days at Leiden have been really special thanks to all of you.

To my friends, Ieva Savickytė (a.k.a. Šiknutė), Priyanka Bahal and Dayu Kim, that have brightened my days at Leiden I want to say Ačiū! धन्यवाद! 고마워!

Finally I would like to thank my family. Gracias a mi tío guitarrista loco, a mis abuelos por las conversaciones tomando mate en el jardín y sobre todo a mi madre, gracias a ella soy hoy la persona que quiero ser.

Contents

Chapter 1

Introduction

Uncertainty relations establish fundamental lower bounds to the information we can extract from a physical phenomenon. These relations are useful in the context of quantum cryptography, since they provide means to prove the security of quantum cryptographic protocols.
The first formulation of an uncertainty relation was given by Heisenberg [9], and it stated that we cannot decrease our uncertainty about the position of a particle without increasing our uncertainty about its momentum, and vice versa. As a lower bound one can express this as,

$$\sigma(P)\sigma(Q) \geq \frac{\hbar}{2}, \tag{1.1}$$

where P and Q are the position and momentum of a particle, $\hbar$ is the Planck constant and σ denotes the standard deviation.
In fact, there will be such relations for any two measurements, as long as there is some degree of incompatibility between both of them.

Instead of using standard deviations, a more convenient way to express uncertainty relations is by means of measures of entropy. Relations expressed in terms of entropy measures are then known as entropic uncertainty relations. A typical choice of entropy for this is the Shannon entropy, this entropy is a particular case of the α-Rényi entropies, a uniparametric family of entropies, that also generalise other well-known measures of entropy.
All previous work considered mostly uncertainty relations for two measurements, and there are only a few relations for single measurements in the literature. For a survey on uncertainty relations, the reader can refer to [5]. The interest behind single measurement relations, is that certain measurements have an inherent uncertainty. In this thesis we initiate a systematic study of entropic uncertainty relations for single measurements.
From the point of view of the applications, these kind of measurements arise in quantum cryptography, for instance in the Round-Robin Differential Phase Shift (RRDPS) encoding [19]. This encoding expresses one string of bits as a quantum state, and an honest decoder obtains one and exactly one bit parity from said string of bits.

In this work we obtain single measurement uncertainty relations, that show the resilience of the RRDPS encoding against a dishonest decoder trying to obtain two bit parities.
Thereafter we study uncertainty relations for a single measurement in a general setting. We introduce a measure $H_\alpha(\mathbf{X})$ which expresses the best possible uncertainty relation for the α-Rényi entropy and study its behaviour under operations like $\oplus$ and $\otimes$.

We provide an efficiently computable, explicit lower bound for $H_\alpha(\mathbf{X})$, which turns out to be an equality, for measurements given by simultaneously diagonalisable operators.
The bounds obtained are, up to our knowledge, the first bounds of uncertainty for a single general measurement, in terms of any α-Rényi entropy.

Finally, we introduce the notion of the space of distributions of a measurement. We discuss how the geometry of this space can convey information about the uncertainty of the associated measurements, and we give sufficient conditions for the dominance of the uncertainty of one measurement over another, with respect to any α-Rényi entropy.

Chapter 2

Preliminaries

In this chapter, we introduce the necessary notions for the study of uncertainty relations. Most importantly we will consider the quantum formalism, which allows us to describe the properties of quantum systems in an abstract mathematical setting. The advantage of this, is that we do not need to know the particular realisation of a quantum system, we just need to know how a general quantum system behaves.

Here we will use the Dirac notation, in which for every vector x in $\mathcal{H}$, a $\mathbb{C}$-Hilbert space, we write $|x\rangle := x$, and $\langle x| := x^\dagger$, where $x^\dagger$ denotes the element of the dual Hilbert space defined by taking the inner product with x. With this notation we have that for $|x\rangle, |y\rangle \in \mathcal{H}$, $\langle x|y\rangle$ denotes the inner product of $|x\rangle$ and $|y\rangle$ in $\mathcal{H}$.

2.1 The quantum formalism

Quantum mechanics deals with quantum systems, in our formalism these systems are described by a finite dimensional Hilbert space over $\mathbb{C}$, which we will usually denote by $\mathcal{H}$. One of the most famous examples of a quantum system is the qubit, which is defined to be $\mathcal{H} = \mathbb{C}^2$, where the inner product is given by $\langle x|y\rangle = \sum_i \bar{x}_i y_i$. The logic behind this choice is that qubits are in a superposition of two states, $|0\rangle$ and $|1\rangle$[1]. Then in contrast to the case of classic information in which a bit is either 0 or 1, qubits can be in a superposition state between these two values. In general we have that the state of a qubit can be expressed as

$$|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle , \tag{2.1}$$

where $\alpha_0, \alpha_1 \in \mathbb{C}$ are such that $|\alpha_0|^2 + |\alpha_1|^2 = 1$. This condition on the scalars α_0, α_1 is equivalent to $|\psi\rangle$ having norm 1 in $\mathcal{H}$.
The reason behind the imposition of $|\psi\rangle$ to have norm 1 comes from another dichotomy between quantum and classical information, namely the fact that measuring a quantum state destroys information. The above definition of a qubit state suggest that one could theoretically store an infinite amount of information in a single state. This paradox however, is resolved by the fact that, when measured to be $|0\rangle$ or $|1\rangle$, the state is going to collapse to one of the pure states, i.e. become either $|0\rangle$ or $|1\rangle$, thus a qubit can store at most one bit of information.
With this in mind the magnitudes $|\alpha_0|^2$ and $|\alpha_1|^2$ of α_0 and α_1 can be interpreted as the probability of measuring the state $|\psi\rangle$ to be $|0\rangle$ and $|1\rangle$ respectively.
This same construction can be generalised to $\mathcal{H} = \mathbb{C}^d$, and in this case we will call the associated quantum system a d-qudit. In a general quantum system the set of states is given by the following definition.

Definition 1. Given a quantum system expressed by a finite dimensional $\mathbb{C}$-Hilbert space $\mathcal{H}$, we define its set of states to be,

$$\mathcal{S}(\mathcal{H}) = \{|\psi\rangle \in \mathcal{H} : \langle\psi|\psi\rangle = 1\}. \tag{2.2}$$

In other words, $\mathcal{S}(\mathcal{H})$ is the sphere in $\mathcal{H}$.

[1] These vectors are defined to be $|0\rangle = (1, 0)$ and $|1\rangle = (0, 1)$, the canonical basis orthonormal vectors.

Two important examples of qubit states are,

$$|+\rangle := \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle, \tag{2.3}$$

$$|-\rangle := \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle. \tag{2.4}$$

These states have the property that they are equally likely to be measured to be either $|0\rangle$ or $|1\rangle$, and conversely $|0\rangle$ and $|1\rangle$ are equally likely to be measured to be $|+\rangle$ or $|-\rangle$. The states $|+\rangle$ and $|-\rangle$ form an orthonormal basis. We will call $\{|0\rangle, |1\rangle\}$ the computational basis and $\{|+\rangle, |-\rangle\}$ the Hadamard basis. One can express the relation between these two bases through the so called Hadamard gate,

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{2.5}$$

which is a unitary matrix, i.e. $H^\dagger H = \mathbb{I}$. Unitary matrices play an important role in the quantum formalism, as the *time evolution of quantum systems* are expressed by them.

We have been discussing measurements of quantum states, to give a precise definition of this we consider the following, for a vector $|\phi\rangle \in \mathcal{H}$ we can define the operator $|\phi\rangle\langle\phi| : \mathcal{H} \to \mathcal{H}$, by

$$|\phi\rangle\langle\phi|(|\psi\rangle) = \langle\phi|\psi\rangle|\phi\rangle. \tag{2.6}$$

The operator $|\phi\rangle\langle\phi|$ is bounded linear, and is positive semi-definite since for every $|\psi\rangle \in \mathcal{H}$

$$\langle\psi|\phi\rangle\langle\phi|\psi\rangle = |\langle\psi|\phi\rangle|^2 \geq 0.$$

Notice furthermore that the image of $|\phi\rangle\langle\phi|$ is equal to the span of the vector $|\phi\rangle$. Now if we have an orthonormal basis $\{|\phi_1\rangle, \ldots, |\phi_d\rangle\}$ in $\mathcal{H}$, then the operators $|\phi_i\rangle\langle\phi_i|$ add up to the identity, this is because we can write any $|\psi\rangle \in \mathcal{H}$ as

$$|\psi\rangle = \sum_i \langle\phi_i|\psi\rangle|\phi_i\rangle = \sum_i |\phi_i\rangle\langle\phi_i|\psi\rangle, \tag{2.7}$$

which is equivalent to $\sum_i |\phi_i\rangle\langle\phi_i| = \mathbb{I}$.
If for instance we take $\{|0\rangle, \ldots, |d-1\rangle\}$, the computational basis in $\mathcal{H} = \mathbb{C}^d$, we have that if $|\psi\rangle$ is a superposition of the shape

$$|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle + \cdots + \alpha_{d-1}|d-1\rangle,$$

then $\langle\psi|i\rangle\langle i|\psi\rangle = |\alpha_i|^2$. So we can retrieve the probability distribution corresponding to measuring $|\psi\rangle$ in one of the basis vectors by using the operators $|i\rangle\langle i|$. In fact this operator corresponds to a projection onto $\mathrm{span}_{\mathbb{C}}(|i\rangle)$, since

$$|i\rangle\langle i|i\rangle\langle i| = \langle i|i\rangle|i\rangle\langle i| = |i\rangle\langle i|, \tag{2.8}$$

and the probabilities are given by the magnitudes of said projections.
More generally, since the $|i\rangle$ form an orthonormal basis, we have furthermore that,

$$|i\rangle\langle i|j\rangle\langle j| = \langle i|j\rangle|i\rangle\langle j| = \delta_{ij}|i\rangle\langle j|, \tag{2.9}$$

where δ_{ij} is the Kronecker delta function.
Therefore the projections $|i\rangle\langle i|$ are pairwise orthogonal as operators. We can generalise these kind of measurements in the following way

Definition 2. A finite collection $\mathbf{P} = (P_1, \ldots, P_n) \subseteq \mathcal{L}(\mathcal{H})^n$, where $\mathcal{L}(\mathcal{H})$ denotes the space of linear bounded operators from $\mathcal{H}$ into $\mathcal{H}$, is called a Projection-Valued Measure, or PVM, if

(i) P_i is a projection for every i, i.e. $P_i^2 = P_i$,

(ii) The P_i are pairwise orthogonal, i.e. $P_iP_j = 0$ for $i \neq j$, and

(iii) The P_i add up to the identity operator, i.e. $\sum_i P_i = \mathbb{I}$.

PVMs are also sometimes called Von Neumann Measurements.

The operators $|i\rangle\langle i|$ defined above, form then a PVM when $\{|0\rangle, \ldots, |d-1\rangle\}$ are an orthonormal basis, and one can also check that the converse is true. This kind of PVMs are called Rank-1 Projection-Valued Measure.
This definition, allows us to formally define the notions of *post-measurement state*, and obtain the probabilities with which they occur. These probabilities are given by the so called Born's Rule, namely we have the following.

Definition 3. Given a PVM $\mathbf{P}$ on $\mathcal{H}$, we define for any state $|\psi\rangle$ the *post-measurement state* by P_i as

$$|\psi^{(i)}\rangle = \frac{1}{\sqrt{p_i}} P_i |\psi\rangle , \tag{2.10}$$

where p_i is the *probability of measuring* $|\psi\rangle$ *to be* $|\psi^{(i)}\rangle$, and is given by the Born's Rule,

$$p_i = \langle\psi|P_i|\psi\rangle . \tag{2.11}$$

The projection valued measurement formalism has the fundamental drawback that it is only a good representation of measurement for a closed quantum system. In a realistic setting, quantum systems will interact with other systems, and therefore we need a more suitable definition that captures this.

Definition 4. We say that $\mathbf{M} \in \mathcal{L}(\mathcal{H})^n$ is a *general measurement*, or simply a *measurement*, if

$$\sum_i M_i^\dagger M_i = \mathbb{I}. \tag{2.12}$$

We denote by $\mathcal{M}eas_{\mathcal{I}}(\mathcal{H})$ the set of measurements on $\mathcal{H}$ indexed by a non-empty set $\mathcal{I}$.

The definitions of post-measurement state and induced distribution can be generalised to general measurements.

Definition 5. Given a measurement $\mathbf{M} \in \mathcal{M}eas_{\mathcal{I}}(\mathcal{H})$, we define for any state $|\psi\rangle \in \mathcal{S}(\mathcal{H})$ the *post-measurement state by* M_i as

$$|\psi^{(i)}\rangle = \frac{1}{\sqrt{p_i}} M_i |\psi\rangle , \tag{2.13}$$

where,

$$p_i = \langle\psi|M_i^\dagger M_i|\psi\rangle . \tag{2.14}$$

We will say then, that measuring a state $|\psi\rangle \in \mathcal{S}(\mathcal{H})$ with $\mathbf{M} \in \mathcal{M}eas_{\mathcal{I}}(\mathcal{H})$ will produce the *post-measurement state* $|\psi^{(i)}\rangle$, with probability p_i.

Because of the condition $\sum_i M_i^\dagger M_i = \mathbb{I}$, the Born's rule for general measurements in (2.14) gives again a probability measure on the set $\mathcal{I}$.
There are some situations in which we are just interested in the induced probability distribution and not in the particular post-measurement states, for instance when we measure the state only once. In these cases we can work instead with the so-called *Positive Operator Valued Measures* or *POVMs*.

Definition 6. An n-tuple of linear bounded operators $\mathbf{X} = (X_1, \ldots, X_n) \in \mathcal{L}(\mathcal{H})^n$ is called a *POVM* if

(i) X_i is positive semi-definite for every i, and

(ii) $\sum_i X_i = \mathbb{I}$.

We will denote the set of POVMs on a Hilbert space $\mathcal{H}$ indexed by $\mathcal{I}$ as $\mathcal{POVM}_{\mathcal{I}}(\mathcal{H})$.

Sometimes we will also write $\mathbf{X} = \{X_i\}_{i\in\mathcal{I}}$ for a POVM or a measurement. Although one should take into account that this notation represents a multiset, since we allow such measurements to include repeated operators.

We include here some simple examples of POVMs,

Example 1. Let $\mathbf{I} = (\mathbb{I}, \dots, \mathbb{I})$ be the n-fold identity vector, and let $p = (p_1, \dots, p_n)$ be a stochastic vector, i.e. $p_i \geq 0$ for all i and $\sum_i p_i = 1$. Then

$$p\mathbf{I} = (p_1, \dots, p_n)\mathbf{I} = \{p_i \mathbb{I}\}_{i=1}^n, \tag{2.15}$$

is a POVM. Indeed every $p_i\mathbb{I}$ is positive semi-definite because $p_i \geq 0$ for every i, and $\sum_i p_i\mathbb{I} = (\sum_i p_i)\mathbb{I} = \mathbb{I}$.

Example 2. Let $\{|0\rangle, \dots, |d-1\rangle\}$ be an orthonormal basis of $\mathcal{H}$, and let $\mathbf{p} = (p_j^{(i)})$ where $i = 1, \dots, n$ and $j = 0, \dots, d-1$, be an $n \times d$ matrix with non-negative entries, such that $\sum_i p_j^{(i)} = 1$, for every j. Then the operators

$$X_i = \sum_j p_j^{(i)} |j\rangle \langle j|, \tag{2.16}$$

form a POVM. The positive semi-definiteness comes from the fact that $p_j^{(i)} \geq 0$ for every i and j, and we have

$$\sum_i X_i = \sum_{i,j} p_j^i |j\rangle \langle j| = \sum_j (\sum_i p_j^{(i)}) |j\rangle \langle j| = \sum_j |j\rangle \langle j| = \mathbb{I}.$$

In fact, we have that any POVM with $X_i X_j = X_j X_i$ for every i, j can be written like (2.16) in a suitable orthonormal basis.

We can also obtain new POVMs from two POVMs, by taking their direct sum, namely we have the following definition

Definition 7. Let $\mathbf{X} = \{X_i\}_{i \in \mathcal{I}}$ and $\mathbf{Y} = \{Y_i\}_{i \in \mathcal{I}}$ be two POVMS, then we can define the *direct sum POVM* as

$$\mathbf{X} \oplus \mathbf{Y} = \{X_i \oplus Y_i\}_{i \in \mathcal{I}}. \tag{2.17}$$

It is clear that $\mathbf{X} \oplus \mathbf{Y}$ is a POVM on the direct sum space, since

$$\sum_i X_i \oplus Y_i = \sum_i X_i \oplus \sum_i Y_i = \mathbb{I} \oplus \mathbb{I},$$

and $X_i \oplus Y_i \geq 0$ for every i, because the set of eigenvalues of $X_i \oplus Y_i$ equals the union of the set of eigenvalues of X_i and Y_i.

It is clear that given a measurement $\mathbf{M}$, we can obtain a POVM from it by simply defining $X_i = M_i^\dagger M_i$. Conversely for a POVM $\mathbf{X}$, if we let $M_i = X_i^{1/2}$ be the positive semi-definite square-root of X_i, then $\mathbf{M} = \{M_i\}$ is a measurement. However, there is not a unique measurement induced by a given POVM, since we can obtain measurements inducing different post-measurement states from a single POVM.

Another fundamental difference between quantum and classical information arises in the case of multiple quantum systems. When we consider the joint quantum system given by two subsystems, new properties appear, like that of entanglement.
In our formalism we will define the joint system of two quantum systems as follows

Definition 8. Given two quantum systems A and B with associated Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, then the joint quantum system AB has the associated Hilbert space

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B, \tag{2.18}$$

where $\otimes$ denotes the tensor product of both spaces.

This definition provides in particular, a new way to act on a state $|\psi\rangle \in \mathcal{S}(\mathcal{H}_A)$. Just consider an arbitrary quantum system B, and a default state $|0\rangle \in \mathcal{S}(\mathcal{H}_B)$, which will be known as an *ancilla*. We can act then on the first part of $|\psi\rangle \otimes |0\rangle \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$, via $U \otimes \mathbb{I}_{\mathcal{H}_B}$ or $M \otimes \mathbb{I}_{\mathcal{H}_B}$, where U is a unitary and M is a measurement.
Another interesting remark is that actions on different parts of the system commute, namely we have

$$(U \otimes \mathbb{I}_{\mathcal{H}_B})(\mathbb{I}_{\mathcal{H}_A} \otimes M) |\Psi\rangle = (\mathbb{I}_{\mathcal{H}_A} \otimes M)(U \otimes \mathbb{I}_{\mathcal{H}_B}) |\Psi\rangle, \tag{2.19}$$

for every $|\Psi\rangle \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$.
In general, an element $|\Psi\rangle \in \mathcal{H}_{AB}$ can be written as a finite $\mathbb{C}$-linear combination of tensor products, i.e.

$$|\Psi\rangle = \sum_i \alpha_i |\psi_i\rangle \otimes |\phi_i\rangle , \tag{2.20}$$

where $\alpha_i \in \mathbb{C}$, $|\psi_i\rangle \in \mathcal{H}_A$ and $|\phi_i\rangle \in \mathcal{H}_B$ for all i. Sometimes we omit the $\otimes$ symbol and just write $|\psi\rangle |\phi\rangle := |\psi\rangle \otimes |\phi\rangle$.
When the quantum state $|\Psi\rangle$ equals the tensor product of two states, we say $|\Psi\rangle$ is a *product state*, otherwise we say $|\Psi\rangle$ is an *entangled state*. Equation (2.19) is clear then for product states, and follows for entangled states by linearity.
As an example of entangled state we take the EPR (Einstein-Podolsky-Rosen) pair

$$|\Phi\rangle = \frac{1}{\sqrt{2}} |0\rangle |0\rangle + \frac{1}{\sqrt{2}} |1\rangle |1\rangle . \tag{2.21}$$

Suppose we measure the state of this joint quantum system on one of its subsystems in the computational basis $\{|0\rangle , |1\rangle\}$. Take for instance $|0\rangle \langle 0| \otimes \mathbb{I}$, it is easy to check that we will measure $|0\rangle$ in the first subsystem with probability $1/2$, explicitly we have

$$\langle \Phi| |0\rangle \langle 0| \otimes \mathbb{I} |\Phi\rangle = \frac{1}{2} \langle 0|0\rangle \langle 0|0\rangle + \frac{1}{2} \langle 0|1\rangle \langle 0|1\rangle = \frac{1}{2}.$$

Therefore, assuming that the outcome of measuring in the firs subsystem is $|0\rangle$, then with probability 1 the post-measurement state of $|\Phi\rangle$ will be

$$|\Phi^{(0)}\rangle = |0\rangle |0\rangle . \tag{2.22}$$

From which we conclude, that if we now measure the resulting state on the second part with $|0\rangle \langle 0|$ we will measure $|0\rangle$ with probability 1.
The above discussion has an interesting physical interpretation, which Einstein called "spooky action at a distance", that even if the two subsystems are geographically distant from each other, measuring one part of it, has an instantaneous effect on the other, and this happens when and only when the measurement is performed.
The properties of entangled states can be exploited in many ways in quantum computing and quantum cryptography, here we will use entanglement to analyse a quantum encoding scheme from a certain cryptographic perspective.

We considered measurements on subsystems of a joint quantum system by taking the tensor product with the identity operator. We can formalise also the notion of measuring independently on both subsystems as follows. In general given two POVMs on the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ associated to two quantum systems, we can define a new POVM in their joint system with associated Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$.

Definition 9. Let $\mathbf{X} = \{X_i\}_i$ and $\mathbf{Y} = \{Y_j\}_j$ be two POVMs, then we define the *tensor product POVM* as

$$\mathbf{X} \otimes \mathbf{Y} = \{X_i \otimes Y_j\}_{i,j}. \tag{2.23}$$

This is again a POVM, since

$$\sum_{i,j} X_i \otimes Y_j = \sum_i X_i \otimes \left(\sum_j Y_j \right) = \sum_i X_i \otimes \mathbb{I}_{\mathcal{H}_2} = \mathbb{I}_{\mathcal{H}_1} \otimes \mathbb{I}_{\mathcal{H}_2} = \mathbb{I}_{\mathcal{H}_1 \otimes \mathcal{H}_2},$$

and also we have $X_i \otimes Y_j \geq 0$ for all i, j because the eigenvalues of $X_i \otimes Y_j$ are precisely $\{\lambda_k \cdot \mu_\ell\}_{k,\ell}$ where $\{\lambda_k\}_k$ and $\{\mu_\ell\}_\ell$ are the eigenvalues of X_i and Y_j respectively, and therefore $\lambda_k \cdot \mu_\ell \geq 0$ for all k, and ℓ.

In particular, we might consider carrying one measurement independently in a quantum system consisting of n subsystems. This gives rise to the following definition

Definition 10. Let $\mathbf{X}$ be a POVM, we define the n-th tensor power POVM as

$$\mathbf{X}^{\otimes 1} := \mathbf{X},$$

for $n = 1$. And,

$$\mathbf{X}^{\otimes n} := \mathbf{X}^{\otimes n-1} \otimes \mathbf{X},$$

for $n > 1$.

For completeness, we finalise this discussion by introducing the following result due to Naimark,

Theorem 1. *(Naimark's Dilation Theorem) Let A be a quantum system with associated Hilbert space $\mathcal{H}_A$, and let $\mathbf{M} \in \mathcal{M}eas_{\mathcal{I}}(\mathcal{H}_A)$ be a measurement. Consider the rank 1 projection-valued measurement $|i\rangle\langle i|$, where $\{|i\rangle\}_{i\in\mathcal{I}}$ is a basis of $\mathcal{H}_B = \mathbb{C}^{|\mathcal{I}|}$. Let also B be a quantum system with $\mathcal{H}_B$ as an associated Hilbert space. Then there exists a unitary $U \in \mathcal{U}(\mathcal{H}_A \otimes \mathcal{H}_B)$, such that for every $|\psi\rangle \in \mathcal{H}_A$,*

$$(M_i |\psi\rangle) \otimes |i\rangle = (\mathbb{I}_{\mathcal{H}_A} \otimes |i\rangle\langle i|)U(|\psi\rangle \otimes |0\rangle), \tag{2.24}$$

where $|0\rangle$ is any ancilla.

In other words, the theorem states that applying a general measurement to a state, is equivalent to applying a rank 1 projection-valued measurement to an ancilla $|0\rangle$, after unitary time evolution of $|\psi\rangle \otimes |0\rangle$. Therefore, a general measurement can be obtained from projection-valued measures by just considering them acting in a multipartite state.

2.2 Entropies

In this section, we will recall the different notions of entropy and what they signify. The aim of an entropy notion, is to provide a measure of surprisal, or lack of information about the possible outcomes of an experiment, or equivalently the amount of information revealed, when realising such an experiment. To exemplify this, consider the experiment of flipping a single coin.
If the coin has a head on both sides, then we know with probability 1 that the outcome will be heads. Therefore we expect that a good entropy will assign a value of 0, to random variables with all their distribution mass concentrated in a single point.
If the coin is fair, i.e. if $P(Tails) = P(Heads) = \frac{1}{2}$, then realising a value of heads or tails, provides more information than in any other case, since our original belief didn't favour any of the two possibilities. Therefore, we should expect a notion of entropy to be maximal for uniform random variables. If there are more events in an equally likely distribution, we should also expect the entropy to be monotonically increasing in the number of events.
In his 1948 paper, Shannon [6] introduced entropy in the context of information theory. His definition of the amount of uncertainty concerning the outcome of an experiment, the possible results of which have the probability distribution $(p_1, \ldots, p_n)$ is given by,

$$H(p_1, \ldots, p_n) = -\sum_{k=1}^{n} p_k \log(p_k). \tag{2.25}$$

Different postulates have been given to characterise this function. The following postulates by Fadeev [7] characterise H up to a choice of basis for the logarithm.

(i) $H(p_1, \ldots, p_n)$ is a symmetric function on its variables for every n,

(ii) $H(p, 1-p)$ is a continuous function of p, for $0 \leq p \leq 1$.

(iii) $H(tp_1, (1-t)p_1, p_2, \ldots, p_n) = H(p_1, \ldots, p_n) + p_1 H(t, 1-t)$ for any distribution $(p_1, \ldots, p_n)$ and $0 \leq t \leq 1$.

For every choice of logarithm base we have then a different unit of measure for information. If we take $\log = \log_2$ then the unit will be called *bit*, if one chooses $\log = \log_e$ then the unit of information is called *nat*. Having set a choice of logarithm we can give the following definition,

Definition 11. If X is a finite random variable, we define the *Shannon entropy* of X as

$$H(X) = -\sum_i p_i \cdot \log p_i, \tag{2.26}$$

where $p = (p_1, \ldots, p_n)$ is the distribution of X.

If the outcome of an experiment occurs with probability p, we call $-\log(p)$ the *information content* of the outcome. Then, the above definition quantifies the average information content of the random variable X.
In other cases however, we might be interested in measures of entropy that give more weight to events with higher or lower information content. Take for instance the following notions of entropy,

- The collision entropy is defined as

$$H_{coll} = -\log p_{coll}(X), \tag{2.27}$$

  where $p_{coll}(X) = \sum_i p_i^2$ is the so-called collision probability, which corresponds to the probability of realising an event of X two times consecutively.

- The min-entropy measures the information content of the probability of guessing a value of a random variable X correctly, or the minimal surprisal of X, namely

$$H_{min} = -\log p_{guess}(X), \tag{2.28}$$

  where $p_{guess}(X) = \max_i p_i$.

- The Hartley entropy is given by

$$H_0(X) = \log |\{i : p_i > 0,\ i = 1, \ldots, n\}|, \tag{2.29}$$

  which is basically the support of the random variable X.

In his paper [16] Rényi introduced a family of entropies which generalise all of the above mentioned entropies.

Definition 12. Let $\alpha \in [0, \infty) - \{1\}$, for a finite random variable X we define the α-Rényi entropy of X as,

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \sum_i p_i^\alpha, \tag{2.30}$$

where $p = (p_1, \ldots, p_n)$ is the distribution of X, and by convention for $\alpha = 0$ we put $0^0 = 0$.

An alternative way of expressing the α-Rényi entropy comes from realising that the term $\sum_i p_i^\alpha$, equals $\|p\|_\alpha^\alpha$. We write then

$$H_\alpha(X) = \frac{\alpha}{1-\alpha} \log \|p\|_\alpha. \tag{2.31}$$

Now with this definition it becomes clear that $H_2 = H_{coll}$, and the Hartley entropy coincides with the 0-Rényi Entropy. One can also show that for every random variable X,

$$\lim_{\alpha \to 1} H_\alpha(X) = H(X),$$

so the α-Rényi entropy also generalises the Shannon entropy. Therefore we can write $H_1(X) := H(X)$. Also,

$$\lim_{\alpha \to \infty} H_\alpha(X) = H_{min}(X), \tag{2.32}$$

so we set $H_\infty(X) = H_{min}(X)$.
The α-Rényi entropies for $\alpha > 1$ give more weight then to the events with higher information content, and those with $\alpha < 1$ give more weight to events with lower information content. The α-Rényi entropy has the following properties,

- $H_\alpha(X) = 0$, if X has all its weight concentrated in one event.

- If X has a uniform distribution, then $H_\alpha(X)$ is maximal and $H_\alpha(X) = H_0(X) = \log |X|$, where $|X|$ denotes the support of X.

- $H_\alpha(X)$ is a monotone non-increasing function on α, i.e.

$$H_\alpha(X) \leq H_\beta(X) \quad \text{if } \alpha \geq \beta. \tag{2.33}$$

- We have for every $\alpha > 1$, that
$$\frac{\alpha}{\alpha - 1} H_\infty(X) \geq H_\alpha(X), \tag{2.34}$$
to see this just notice that $\sum_j p_j^\alpha \geq p_i^\alpha$ for every i, the inequality $\sum_j p_j^\alpha \geq \max_i p_i^\alpha$ holds. Now since $1/(1-\alpha)\log x$ is decreasing for $x > 1$ we have that
$$\frac{\alpha}{1-\alpha} \log \max_i p_i^\alpha \geq H_\alpha(X), \tag{2.35}$$
from which we get the result.

2.3 Majorisation

One of the most important properties of entropies is the Schur-concavity. To define this property we need to introduce the notion of majorisation.

Definition 13. Let $x \in \mathbb{R}^n$, we denote by $x^\downarrow$ the vector of coordinates of x arranged in descending order, so $x_1^\downarrow \geq x_2^\downarrow \geq \cdots \geq x_n^\downarrow$. Analogously we define $x^\uparrow$ as the vector of coordinates of x arranged in ascending order.

With the above notation we can express majorisation in the following way.

Definition 14. Let $x, y \in \mathbb{R}^n$, we say that x is majorised by y, and write $x \prec y$ if,
$$\sum_{j=1}^k x_j^\downarrow \leq \sum_{j=1}^k y_j^\downarrow, \quad \text{for all} \quad 1 \leq k \leq n, \tag{2.36}$$
and $\sum_j x_j = \sum_j y_j$.

We also have the following weaker conditions,

Definition 15. Let $x, y \in \mathbb{R}^n$, we say that x is weakly submajorised by y, and write $x \prec_\omega y$ if,
$$\sum_{j=1}^k x_j^\downarrow \leq \sum_{j=1}^k y_j^\downarrow, \quad \text{for all} \quad 1 \leq k \leq n.$$
Notice that we do not require that the sum of the coordinates of x equals the sum of coordinates of y. We say that x is weakly supermajorised by y, and write $x \prec^\omega y$, if
$$\sum_{j=1}^k x_j^\uparrow \geq \sum_{j=1}^k y_j^\uparrow \quad \text{for all} \quad 1 \leq k \leq n. \tag{2.37}$$
Therefore $x \prec y$ if and only if $x \prec_\omega y$ and $x \prec^\omega y$.

If we interpret the magnitudes of the coordinates of the vectors as weights, majorisation then means that the weights of x are more evenly distributed in its coordinates than the weights of y.
Indeed for an arbitrary stochastic vector $p = (p_1, \ldots, p_n)$, we have
$$p^\star = (\frac{1}{n}, \ldots, \frac{1}{n}) \prec p \prec (1, 0, \ldots, 0). \tag{2.38}$$
There are several equivalent ways of characterising majorisation [3]. Geometrically $x \prec y$ means that x is contained in the convex hull defined by the orbit of y under the group action on $\mathbb{R}^n$ defined as follows:
$$\begin{aligned} S_n \times \mathbb{R}^n &\to \mathbb{R}^n \\ (\sigma, x) &\mapsto \sigma(x), \end{aligned}$$
where $\sigma(x)_i := x_{\sigma^{-1}(i)}$ for all $i = 1, \ldots, n$. If we write then $\langle y \rangle := \{\sigma(y) : \sigma \in S_n\}$, we have that,
$$x \prec y \quad \text{ if and only if } \quad x \in \mathrm{Conv}\,\langle y \rangle\,, \tag{2.39}$$
where $\mathrm{Conv}(\mathcal{X})$ denotes the convex hull of any $\mathcal{X} \subseteq \mathbb{R}^n$.
For completeness, we also include another important characterisation of majorisation.

Definition 16. Let $A = (a_{ij})$ be an $n \times n$ matrix, then we say that A is a *doubly stochastic matrix* if $\sum_i a_{ij} = 1$ for all j, $\sum_j a_{ij} = 1$ for all i, and $a_{ij} \geq 0$ for all i, and j. Equivalently A is said to be doubly stochastic, if

(i) A is positivity-preserving, i.e. for every $x \in \mathbb{R}^n_{\geq 0}$, we have $Ax \in \mathbb{R}^n_{\geq 0}$.

(ii) A is trace-preserving, i.e. $\mathrm{tr}(Ax) = \mathrm{tr}(x)$, where $\mathrm{tr}(x) := \sum_i x_i$.

(iii) A is unital, i.e. $Ae = e$, where $e = (1, \ldots, 1)$.

Then, we have that $x \prec y$ if and only if $x = Ay$ for some doubly stochastic matrix A. For a proof of the equivalence of the characterisations of majorisation see [3].

Definition 17. Let $A \subseteq \mathbb{R}^n$, a function $f : A \to \mathbb{R}$ is said to be Schur-convex when,

$$f(x) \leq f(y) \quad \text{if} \quad x \prec y. \tag{2.40}$$

Similarly f is said to be Schur-concave when,

$$f(x) \geq f(y) \quad \text{if} \quad x \prec y. \tag{2.41}$$

If we see then the α-Rényi entropy as a function from $\Delta_{n-1} \subseteq \mathbb{R}^n \to \mathbb{R}$, where $\Delta_{n-1} = \{(p_1, \ldots, p_n) \in \mathbb{R}^n : \sum_i p_i = 1\}$, we have the following theorem,

Theorem 2. The α-Rényi entropy $H_\alpha(x)$ is Schur-concave on Δ_{n-1}.

Proof. For a proof the reader can refer to [14]. □

2.4 Matrix Functions

We will need some results in matrix analysis that we will list here. These will prove useful further on when we will try to derive properties for the entropies of POVMs, and find explicit expressions for uncertainty relations.

One basic result of linear algebra is that for any hermitian operator A, there is a suitable choice of an orthonormal basis $\{|0\rangle, \ldots, |n-1\rangle\}$ such that we can write A in the form,

$$A = \sum_i \lambda_i |i\rangle \langle i| , \tag{2.42}$$

where $\lambda_i \in \mathbb{R}$ for every i, if A is positive semi-definite, then $\lambda_i \in \mathbb{R}_{\geq 0}$. This is called the *spectral decomposition* of A, and corresponds to the usual notion of diagonalisation. Note that this choice of $|i\rangle$ is not in general the computational basis but instead they form an orthonormal basis of eigenvectors of A.
With this in mind, we can extend functions $f : \mathbb{R} \to \mathbb{R}$ to hermitian matrices by simply taking,

$$f(A) := \sum_i f(\lambda_i) |i\rangle \langle i| . \tag{2.43}$$

In particular this defines A^α for every $\alpha \geq 0$. We will be interested in the logarithm of positive semi-definite matrices, which we will define as follows,

Definition 18. Let A be a positive semi-definite hermitian matrix, we define the *logarithm of* A as

$$\log(A) := \sum_{\substack{i \\ \lambda_i \neq 0}} \log(\lambda_i) |i\rangle \langle i| . \tag{2.44}$$

Note that the above definition doesn't correspond to the logarithm of A in the sense that $e^{\log A} = A$, since A may be a singular matrix. However if A is nonsingular then both notions coincide.

Lemma 1. For any positive semi-definite matrix A, we have that

$$\frac{\partial}{\partial\alpha}A^\alpha = A^\alpha \cdot \log A. \tag{2.45}$$

Furthermore, A^α and $\log A$ are diagonalisable in the same basis, i.e. we can take their orthonormal bases in their spectral decomposition to be the same.

Proof. Consider the spectral decomposition of A,

$$A = \sum_i \lambda_i \left|i\right\rangle \left\langle i\right|,$$

then by definition,

$$A^\alpha = \sum_i \lambda_i^\alpha \left|i\right\rangle \left\langle i\right|,$$

so we have that

$$\begin{aligned}\frac{\partial}{\partial\alpha}A^\alpha &= \sum_i \frac{\partial}{\partial\alpha}\lambda_i^\alpha \left|i\right\rangle \left\langle i\right| \\ &= \sum_i \lambda_i^\alpha \log(\lambda_i) \left|i\right\rangle \left\langle i\right| \\ &= A^\alpha \log A.\end{aligned}$$

It is clear by the above proof, that the choice of the orthonormal basis is the same in the spectral decomposition of A, A^α and $\log A$.

□

Definition 19. For an hermitian operator A, with spectral decomposition $A = \sum_i \lambda_i \left|i\right\rangle \left\langle i\right|$. We write,

$$\lambda(A) = (\lambda_1, \ldots, \lambda_n), \tag{2.46}$$

for the vector consisting of its eigenvalues with multiplicities.

We can choose any order for the above definition, since we will always consider λ rearranged in ascending or descending order.

In many cases, we will need to maximise or minimise the values of the form

$$f_A\left|\psi\right\rangle = \left\langle\psi|A|\psi\right\rangle, \tag{2.47}$$

over all state vectors $\left|\psi\right\rangle$, for a given hermitian operator A.
The minimax principle provides us with an explicit solution to this optimisation problem.

Theorem 3. *(The minimax Principle)*
Let A be an Hermitian operator on an n-dimensional Hilbert space $\mathcal{H}$, then for every $1 \leq k \leq n$,

$$\lambda_k^\downarrow(A) = \max_{\substack{\mathcal{M}\subseteq\mathcal{H} \\ \dim\mathcal{M}=k}} \min_{\substack{\left|\psi\right\rangle\in\mathcal{M} \\ \|\left|\psi\right\rangle\|=1}} \left\langle\psi|A|\psi\right\rangle \tag{2.48}$$

$$= \min_{\substack{\mathcal{M}\subseteq\mathcal{H} \\ \dim\mathcal{M}=n-k+1}} \max_{\substack{\left|\psi\right\rangle\in\mathcal{M} \\ \|\left|\psi\right\rangle\|=1}} \left\langle\psi|A|\psi\right\rangle. \tag{2.49}$$

Recall that $\mathcal{S}(\mathcal{H}) = \{\left|\psi\right\rangle \in \mathcal{H} : \|\left|\psi\right\rangle\| = 1\}$. Then, we will usually use this theorem in the following form,

Corollary 1. For every Hermitian operator A we have,

$$\lambda_1^\downarrow(A) = \max_{\left|\psi\right\rangle\in\mathcal{S}(\mathcal{H})} \left\langle\psi|A|\psi\right\rangle, \tag{2.50}$$

and,

$$\lambda_n^\downarrow(A) = \min_{\left|\psi\right\rangle\in\mathcal{S}(\mathcal{H})} \left\langle\psi|A|\psi\right\rangle. \tag{2.51}$$

The biggest eigenvalue, $\lambda_1^{\downarrow}(A)$ of an Hermitian operator A is related to the so-called operator norm of A.
In general for a linear bounded operator $A \in \mathcal{L}(\mathcal{H})$, we define the operator norm to be the infimum over all $k > 0$ such that

$$\|Ax\|_2 \leq k\|x\|_2, \quad \text{for all } x \in \mathcal{H} \tag{2.52}$$

where $\|\cdot\|_2$ denotes the 2-norm in $\mathcal{H}$.
We denote then

$$\|A\| = \inf k = \sup_{\|x\|_2=1} \|Ax\|_2. \tag{2.53}$$

It becomes clear now that for a matrix A, the operator norm equals the biggest eigenvalue of A. We have that $\|\cdot\|$ is indeed a norm [18].
The operator norm is also submultiplicative, namely we have

Lemma 2. *We have that if A and B are two linear bounded operators, then*

$$\|AB\| \leq \|A\|\|B\| \tag{2.54}$$

Proof. Let $|\psi\rangle \in \mathcal{H}$, then we have that

$$\begin{aligned} \|AB\,|\psi\rangle\,\|_2 &\leq \|A\|\|B\,|\psi\rangle\,\|_2 \\ &\leq \|A\|\|B\|\|\,|\psi\rangle\,\|_2. \end{aligned}$$

Therefore we conclude $\|AB\| \leq \|A\|\|B\|$. □

We will extensively consider positive semi-definite matrices, as they appear associated to the operators of a POVM.
The set of positive semi-definite matrices forms a convex cone contained in $\mathcal{L}(\mathcal{H})$. Indeed any positive linear combination of positive semi-definite matrices is in turn going to be a positive semi-definite matrix. Therefore, there is a partial ordering induced by the cone which is the following,

Definition 20. *(Loewner Order)*
Let A, B be Hermitian positive semi-definite matrices, then we write,

$$A \prec B \quad \text{if and only if} \quad B - A \geq 0. \tag{2.55}$$

This notion of ordering will be very useful because with it we can derive inequalities for the expressions $\langle\psi|A|\psi\rangle$ where $|\psi\rangle \in \mathcal{H}$. Suppose for instance that $A \prec B$, then we have that $B - A$ is positive semi-definite, therefore for every $|\psi\rangle \in \mathcal{H}$,

$$\langle\psi|A|\psi\rangle = \langle\psi|B|\psi\rangle - \langle\psi|B - A|\psi\rangle \leq \langle\psi|B|\psi\rangle\,,$$

by the positive semi-definiteness of $B - A$.

For the Loewner order we have an analogue of the triangle inequality,

$$|A + B| \prec U|A|U^\dagger + V|B|V^\dagger, \tag{2.56}$$

for some $U, V \in \mathcal{U}(\mathcal{H})$, where $|A| = \sqrt{A^\dagger A}$ is the matrix modulus of A.
Also for the Loewner order there is an analogue of Young's inequality for matrices

$$U|A \cdot B|U^\dagger \prec \frac{1}{p}|A|^p + \frac{1}{q}|B|^q, \tag{2.57}$$

where $U \in \mathcal{U}(\mathcal{H})$, and p, q are positive real and such that, $\frac{1}{p} + \frac{1}{q} = 1$.
Therefore if we have a norm for matrices that is invariant under products by unitary matrices, we can obtain analogues of Hölder's inequality [3].

Chapter 3

The RRDPS Encoding

3.1 Introduction to the RRDPS encoding

Most of Quantum Cryptography relies on the BB84 (*Bennet and Brassard 1984*) encoding [2] or variations thereof. In this encoding one works with two bases orthogonal to each other in a two dimensional Hilbert space, for instance the computational and Hadamard bases. The Round Robin Differential Phase Shift or RRDPS is a different encoding [19] introduced in the context of a new *Quantum Key Distribution scheme* (QKD). This encoding uses instead d-dimensional qudits, and it is easy to implement despite being a d-qudit protocol. Its main difference with respect to BB84 is that the encoding is fixed and does not depend on the choice of bases.
The RRDPS encoding, has the potential to be useful beyond QKD, since when carried by an honest encoder (or sender) and decoder (or receiver), it allows the decoder to retrieve exactly one out of several possible bits of the encoded message. This is exactly what an *Oblivious Transfer* (OT) protocol is supposed to achieve. OT is an important cryptographic primitive that allows a user to extract one and exactly one element from a database, without letting the server know which element did the user extract. In order to possibly obtain security against dishonest parties in the RRDPS, we need to better understand how well the encoding hides information from a dishonest receiver.

In this chapter, we will study how the RRDPS encoding hides information, and we will see how this question relates to entropic uncertainty relations, namely bounds on the entropy of the induced distributions given by Born's rule.
The approach we will use, exploits the commutativity of operations in joint quantum systems, and it is a standard technique for proving the security of quantum cryptographic protocols. Namely we will obtain an equivalent formulation of the protocol, where the information of interest is obtained as the result of a measurement.

The encoding procedure on RRDPS can be formalised in the following way. Take $\{|0\rangle, \ldots, |d-1\rangle\}$, to be an orthonormal basis of $\mathcal{H} = \mathbb{C}^d$. Now for every d-bit string $z \in \{0,1\}^d$, we set the encoding of z to be,

$$|\psi_z\rangle := \frac{1}{\sqrt{d}} \sum_i (-1)^{z_i} |i\rangle , \tag{3.1}$$

where z_i is the i-th coordinate of z. It is clear that $|\psi_z\rangle \in \mathcal{S}(\mathcal{H})$ since $\langle\psi_z|\psi_z\rangle = \frac{1}{d}\sum_i (-1)^{2z_i} \langle i|i\rangle = 1$. And therefore one can perform measurements on this kind of quantum states. Now in order to define the decoding procedure to find information about parities of the coordinates of z, define for every choice of $R \in \{1, \ldots, d-1\}$ the following states,

$$|t^R_{s,k}\rangle := \frac{1}{\sqrt{2}} \left[|k\rangle + (-1)^s |k \oplus R\rangle \right], \tag{3.2}$$

where $0 \leq k < d$, $s \in \{0,1\}$ and $\oplus$ denotes addition modulo d; Then we can consider for every such choice of R, a POVM $\mathbf{X}^R = \{X^R_{s,k}\}_{s,k} = \{\frac{1}{2} |t^R_{s,k}\rangle \langle t^R_{s,k}|\}_{s,k}$, that will give us the decoding step in the protocol.

Indeed we have that $\mathbf{X}^R$ forms a POVM,

$$\begin{aligned}\sum_{s,k}\frac{1}{2}\left|t^R_{s,k}\right\rangle\left\langle t^R_{s,k}\right| &= \frac{1}{4}\sum_{s,k}\left[|k\rangle\langle k| + (-1)^s(|k\rangle\langle k\oplus R| + |k\oplus R\rangle\langle k|) + |k\oplus R\rangle\langle k\oplus R|\right]\\ &= \frac{1}{4}\left[4\mathbb{I} + \sum_{s,k}(-1)^s(|k\rangle\langle k\oplus R| + |k\oplus R\rangle\langle k|)\right]\\ &= \mathbb{I},\end{aligned}$$

and $X^R_{s,k} \geq 0$ for every s, k by construction. To see that this measurements can reveal information about parities, take any state $|\psi_z\rangle$, then

$$\begin{aligned}\sqrt{X_{s,k}}\,|\psi_z\rangle &= \frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{d}}\left(\sum_i\left[(-1)^{z_i}\langle k|i\rangle + (-1)^s\cdot(-1)^{z_i}\langle k\oplus R|i\rangle\right]\right)|t^R_{s,k}\rangle\\ &= \frac{1}{2}\cdot\frac{1}{\sqrt{d}}\left[(-1)^{z_k} + (-1)^{s+z_{k\oplus R}}\right]\cdot|t^R_{s,k}\rangle\,,\end{aligned}$$

which vanishes if $z_k + s + z_{k\oplus R} = 1 \pmod 2$. And therefore we observe $s = 1 + z_k + z_{k\oplus R}$ with probability 0. Thus the measurement outcome s, k is uniformly random subject to $s = z_k + z_{k\oplus R} \pmod 2$, i.e. s being the parity between the k-th and $k\oplus R$-th bits of z.
So for every R, the R-decoding gives us one bit of information about z in the form of a parity, where we learn which parity we are obtaining, but we have no control on which one in particular, only on the distance R between the bits.

With this we can describe the following simple communication scheme, in which we assume both the sender and the receiver to be honest,

$$\{0,1\}^d \xrightarrow{r} z \longrightarrow |\psi_z\rangle \longrightarrow \boxed{\mathbf{X}^R} \longrightarrow s, k$$

Protocol 1.

- *The sender $\mathcal{S}$, takes $z \in \{0,1\}^d$ uniformly at random, then prepares the quantum state $|\psi_z\rangle$ and sends it through a channel to the receiver $\mathcal{R}$.*
- *The receiver $\mathcal{R}$ measures the state $|\psi_z\rangle$ in $\mathbf{X}^R$, and retrieves a uniformly random tuple (s, k), consisting of a parity $s = z_k \oplus z_{k\oplus R}$ together with its position k.*

To see to some extent how much information a dishonest receiver can get by performing, instead of $\mathbf{X}^R$ a different measurement χ, we consider the diagram,

$$\{0,1\}^d \xrightarrow{r} z \longrightarrow |\psi_z\rangle \longrightarrow \boxed{\chi} \longrightarrow i,$$

and the question then becomes: How much information can we obtain about z given the information i obtained by using a dishonest measurement? We can express this amount of information about z in terms of i as

$$H(z|i),$$

where H is an entropy function.

3.2 Purifying the protocol

In order to study this entropy, one can reformulate the communication protocol in such a way that it will be possible to analyse it. We can do so, by defining a measurement which will give us the encodings $|\psi_z\rangle$ when measuring the *fully entangled state*,

$$|\Psi\rangle = \sum_i \frac{1}{\sqrt{d}}|i\rangle|i\rangle\,. \tag{3.3}$$

We will define this POVM as follows, for every $z \in \{0,1\}^d$ we write,

$$Q_z := \frac{d}{2^d} |\psi_z\rangle \langle\psi_z| , \tag{3.4}$$

and we let $\mathbf{Q} = \{Q_z\}_{z\in\{0,1\}^d}$. By construction again, the Q_z are positive semi-definite, and we have that

$$\begin{aligned}
\sum_z Q_z &= \frac{d}{2^d} \sum_z |\psi_z\rangle \langle\psi_z| \\
&= \frac{1}{2^d} \sum_{z,i,j} (-1)^{z_i+z_j} |i\rangle \langle j| \\
&= \frac{1}{2^d} \sum_{z,i} |i\rangle \langle i| + \frac{1}{2^d} \sum_{z,i\neq j} (-1)^{z_i+z_j} |i\rangle \langle j| \\
&= \frac{1}{2^d} \sum_{z,i} |i\rangle \langle i| \\
&= \frac{1}{2^d} \sum_z \mathbb{I} = \mathbb{I},
\end{aligned}$$

where in the fourth equality, we used the fact that there are exactly 2^{d-1} summands with $z_i = 0$ and 2^{d-1} summands with $z_i = 1$ in $\sum_{z,i\neq j}(-1)^{z_i+z_j} |i\rangle \langle j|$, so they cancel out.
In order to obtain the post-measurement states under $\mathbf{Q}$, we need to consider the unique positive semi-definite square-root of Q_z for every z. This can be easily checked to be

$$\sqrt{Q_z} = \frac{\sqrt{d}}{\sqrt{2^d}} |\psi_z\rangle \langle\psi_z| . \tag{3.5}$$

Now we can see that when the fully entangled state is measured on one of its subsystems by $\sqrt{Q_z}$ we obtain the states $|\psi_z\rangle$ in both subsystems, namely we have,

$$\begin{aligned}
(\sqrt{Q_z} \otimes \mathbb{I}) |\Psi\rangle &= \sum_i \frac{1}{\sqrt{d}} Q_z |i\rangle \otimes |i\rangle \\
&= \sum_i \frac{\sqrt{d}}{\sqrt{2^d}} \frac{1}{\sqrt{d}} \langle\psi_z|i\rangle |\psi_z\rangle \otimes |i\rangle \\
&= \frac{1}{\sqrt{2^d}} \sum_i \frac{(-1)^{z_i}}{\sqrt{d}} |\psi_z\rangle \otimes |i\rangle \\
&= \frac{1}{\sqrt{2^d}} |\psi_z\rangle \otimes \left(\sum_i \frac{(-1)^{z_i}}{\sqrt{d}} |i\rangle \right) \\
&= \frac{1}{\sqrt{2^d}} |\psi_z\rangle \otimes |\psi_z\rangle ,
\end{aligned}$$

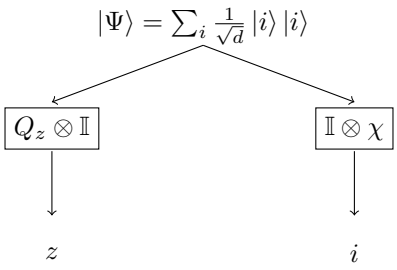
and since $\langle\Psi|Q_z \otimes \mathbb{I}|\Psi\rangle = \frac{1}{2^d}$, we finally conclude that the post-measurement state of $|\Psi\rangle$ under $\sqrt{Q_z} \otimes \mathbb{I}$ is $|\psi_z\rangle \otimes |\psi_z\rangle$. Also since $\langle\Psi|Q_z \otimes \mathbb{I}|\Psi\rangle = \frac{1}{2^d}$, we see that by performing the measurement $\mathbf{Q}$ we will obtain a z drawn uniformly at random from $\{0,1\}^d$.

By the commutativity of the operators in the separate subsystems we conclude that the order of application of $\sqrt{Q_z} \otimes \mathbb{I}$ and $\mathbb{I} \otimes \chi$, does not alter the final distribution. Namely we have that

$$(\sqrt{Q_z} \otimes \mathbb{I}) |\psi^{(i)}\rangle = (\sqrt{Q_z} \otimes \mathbb{I})(\mathbb{I} \otimes \chi_i) |\Psi\rangle = (\mathbb{I} \otimes \chi)(\sqrt{Q_z} \otimes \mathbb{I}) |\Psi\rangle = (\mathbb{I} \otimes \chi_i) |\psi_z\rangle \otimes |\psi_z\rangle , \tag{3.6}$$

where $|\psi^{(i)}\rangle$ is the post-measurement state of $|\Psi\rangle$ after $\mathbb{I} \otimes \chi$, for some i. Notice now that the information obtained by measuring with χ in the protocol is captured by the distributions $\langle\psi_z|\chi_i|\psi_z\rangle$ which by the above equal the distributions $\langle\psi^{(i)}|Q_z|\psi^{(i)}\rangle$. In this way we don't need to study an arbitrary measurement χ anymore, and instead we can analyse Q_z and obtain lower bounds of entropy for it, i.e. uncertainty relations, so that we know how much uncertainty a dishonest receiver has in the worst case scenario.

This gives rise to a new protocol, which is called a purified version, that preserves the probability distribution of the measurement outcomes,



Protocol 2.

- *The sender $\mathcal{S}$, prepares a fully entangled state $|\Psi\rangle$, keeps one part of the system, and sends the other to $\mathcal{R}$.*
- *$\mathcal{R}$ measures her part of the subsystem with χ (if she is honest she performs the measurement $\mathbf{X}^R$).*
- *$\mathcal{S}$ measures her part of the subsystem with $\mathbf{Q}$.*

In this purified protocol, the last two steps can be performed in any order by the two parties. We can summarise the above in the following result,

Theorem 4. *Let $p = (p_{z,i})_{z,i}$ be the joint distribution induced by the measurement χ and the random choice of z when carrying Protocol 1, and let $q = (q_{z,i})_{z,i}$ be the joint distribution induced by the measurements $\mathbf{Q}$ and χ, obtained by carrying Protocol 2. Then $p = q$.*

Proof. In Protocol 1 we draw z uniformly at random from $\{0,1\}^d$, and we have seen that $q_z = \langle\Psi|Q_z \otimes \mathbb{I}|\Psi\rangle = \frac{1}{2^d}$ so we have that the marginal probabilities of z coincide, i.e. $p_z = q_z$ for all $z \in \{0,1\}^d$. Now notice that for every i, $p_{i|z} = \langle\psi_z|\chi_i|\psi_z\rangle$ and $q_{i|z} = \langle\Psi|Q_z \otimes \chi_i|\Psi\rangle$, are the conditional probabilities of i given z. But then by applying equation (3.6),

$$\begin{aligned} q_{i|z} &= \langle\Psi|(\sqrt{Q_z}^\dagger \sqrt{Q_z} \otimes \mathbb{I})(\mathbb{I} \otimes \chi_i)|\Psi\rangle \\ &= \langle\Psi|(\sqrt{Q_z}^\dagger \otimes \mathbb{I})(\mathbb{I} \otimes \chi_i)(\sqrt{Q_z} \otimes \mathbb{I})|\Psi\rangle \\ &= \langle\psi_z, \psi_z|\mathbb{I} \otimes \chi_i|\psi_z, \psi_z\rangle \\ &= \langle\psi_z|\psi_z\rangle \langle\psi_z|\chi_i|\psi_z\rangle \\ &= \langle\psi_z|\chi_i|\psi_z\rangle = p_{i|z}. \end{aligned}$$

Therefore, since $p_z = q_z$ and $p_{i|z} = q_{i|z}$ we conclude that the joint distributions are equal, i.e. $p_{z,i} = q_{z,i}$. □

In particular we have that the conditional probabilities $p_{z|i}$ and $q_{z|i}$ coincide, so we have

$$H(z|i) = H(p_{z|i}) = H(q_{z|i}). \tag{3.7}$$

We can check the soundness of this modified protocol by studying what information we obtain when $\mathcal{R}$ measures first and is honest, i.e. when $\chi = \mathbf{X}^R$, and we apply first χ and then $\mathbf{Q}$ to the respective subsystems. Applying $\mathbb{I} \otimes \mathbf{X}^R$ to the fully entangled state gives the following,

$$\begin{aligned} \frac{1}{\sqrt{2}} |t^R_{s,k}\rangle \langle t^R_{s,k}| |\Psi\rangle &= \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{d}} \sum_i |i\rangle \otimes (\langle t^R_{s,k}|i\rangle |t^R_{s,k}\rangle) \\ &= \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{d}} \cdot \frac{1}{\sqrt{2}} (|k\rangle |t^R_{s,k}\rangle + (-1)^s |k \oplus R\rangle |t^R_{s,k}\rangle) \\ &= \frac{1}{\sqrt{2d}} |t^R_{s,k}\rangle |t^R_{s,k}\rangle . \end{aligned}$$

And since $\langle\Psi|\mathbb{I} \otimes X^R_{s,k}|\Psi\rangle = \frac{1}{2d}$, we conclude that the post-measurement state of $|\Psi\rangle$ under $\mathbb{I} \otimes \sqrt{X^R_{s,k}} = \mathbb{I} \otimes \frac{1}{\sqrt{2}} |t^R_{s,k}\rangle \langle t^R_{s,k}|$ is $|t^R_{s,k}\rangle |t^R_{s,k}\rangle$. It is enough then to study $\sqrt{Q_z} |t^R_{s,k}\rangle$, to describe the information $\mathcal{R}$ obtains when $\mathcal{S}$ measures in second place. We have

$$\begin{aligned}\sqrt{Q_z}\,|t^R_{s,k}\rangle &= \frac{\sqrt{d}}{\sqrt{2^d}}\,\langle\psi_z|t^R_{s,k}\rangle\,|\psi_z\rangle \\ &= \frac{1}{\sqrt{2^{d+1}}}\left[\sum_i (-1)^{z_i}(\langle i|k\rangle + (-1)^s\,\langle i|k\oplus R\rangle)\right]|\psi_z\rangle \\ &= \frac{1}{\sqrt{2^{d+1}}}\left[(-1)^{z_k} + (-1)^s(-1)^{z_{k\oplus R}}\right]|\psi_z\rangle\,.\end{aligned}$$

And again since we will measure $s = 1 + z_k + z_{k\oplus R}$ with probability 0, we conclude that $s = z_k + z_{k\oplus R} \pmod 2$, as expected.

3.3 Bounds on the entropy of bit pairs

It is clear that there is uncertainty in the random bit string z given the information i. This is due to Holevo's bound [10], which informally states that a d dimensional state carries at most $\log_2 d$ bits of information. But, whether carrying the protocol, allows us to extract information about two parities of z is not so trivial. In order to study this possible vulnerability, we want to know how much uncertainty there is in a given tuple of parities of z, that is to what extent a dishonest receiver can make a correct guess on such parities.
Suppose we don't have uncertainty on parities of the type $\mathcal{X} = (z_0 \oplus z_1, z_2 \oplus z_3)$, i.e. two parities with no overlaps. Then we can obtain $d/2$ such parities but then this would be a contradiction to Holevo's bound, since we cannot obtain more than $\log d$ bits of information from z.
Consider then, without loss of generality, the parities $\mathcal{X} = (z_0 \oplus z_1, z_0 \oplus z_2)$. To find how much entropy these parities have, we can either extract the distribution of $\mathcal{X}$ from the distribution given by Q_z or by linearity we can consider the following operators,

$$S_{ab} := \sum_{\substack{z_0+z_1=a\\ z_0+z_2=b}} Q_z, \tag{3.8}$$

where z again ranges over $\{0,1\}^d$, and $a, b \in \{0,1\}$. These operators form in turn a POVM $\chi = \mathbf{S} =: \{S_{ab}\}_{a,b\in\{0,1\}}$. The positive semi-definiteness of the S_{ab} follows from that of the Q_z, and we have $\sum_{a,b} S_{ab} = \mathbb{I}$, since there are exactly 2^{n-1} bit strings z with $z_0 + z_1 = 0$ and exactly 2^{n-1} with $z_0 + z_1 = 1$, thus $\sum_{a,b} S_{ab} = \sum_z Q_z = \mathbb{I}$. We can find an explicit expression for S_{ab} in the computational basis $\{|0\rangle, \ldots, |d-1\rangle\}$,

$$\begin{aligned}S_{00}\,|0\rangle &= \sum_{z_0=z_1=z_2} Q_z\,|0\rangle \\ &= \frac{d}{2^d}\sum_{z_0=z_1=z_2} \langle\psi_z|0\rangle\,|\psi_z\rangle \\ &= \frac{1}{2^d}\sum_{z_0=z_1=z_2}\sum_i (-1)^{z_0+z_i}\,|i\rangle \\ &= \frac{2^{d-2}}{2^d}(|0\rangle + |1\rangle + |2\rangle) = \frac{1}{4}(|0\rangle + |1\rangle + |2\rangle),\end{aligned}$$

and analogously $S_{00}\,|1\rangle = S_{00}\,|2\rangle = \frac{1}{4}(|0\rangle + |1\rangle + |2\rangle)$. For $i \geq 3$ we have

$$S_{00}\,|i\rangle = \frac{1}{4}\,|i\rangle\,.$$

Therefore S_{00} is of the shape

$$S_{00} = \frac{1}{4}\left(\begin{array}{c|c} M_{00} & 0 \\ \hline 0 & \mathbb{I}_{d-3}\end{array}\right),$$

where $\mathbb{I}_{d-3}$ is the $d-3 \times d-3$ identity matrix, and

$$M_{00} = \begin{pmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{pmatrix} = (|0\rangle + |1\rangle + |2\rangle)(\langle 0| + \langle 1| + \langle 2|).$$

We can proceed analogously with S_{01}, S_{10} and S_{11}, just noticing that the difference in parity implies a change of sign. We have for $a, b \in \{0,1\}$,

$$S_{ab} = \frac{1}{4}\left(\begin{array}{c|c} M_{ab} & 0 \\ \hline 0 & \mathbb{I}_{d-3} \end{array}\right),$$

where

$$M_{00} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = (|0\rangle + |1\rangle + |2\rangle)(\langle 0| + \langle 1| + \langle 2|),$$

$$M_{01} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix} = (|0\rangle + |1\rangle - |2\rangle)(\langle 0| + \langle 1| - \langle 2|),$$

$$M_{10} = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} = (|0\rangle - |1\rangle + |2\rangle)(\langle 0| - \langle 1| + \langle 2|),$$

$$M_{11} = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} = (|0\rangle - |1\rangle - |2\rangle)(\langle 0| - \langle 1| - \langle 2|).$$

Notice we can write $\mathbf{S} = \mathbf{M} \oplus \frac{1}{4}\mathbf{I}$, where $\mathbf{M} = \{\frac{1}{4}M_{ab}\}_{a,b\in\{0,1\}}$. For reasons we will discuss in Chapter 4, it will be enough to find bounds for the entropy of $\mathbf{M}$. We are interested for instance in the guessing entropy H_∞ and the collision entropy H_2 of this measurement, so we wish to obtain uncertainty relations for them. We have the following result,

Proposition 1. *For every state $|\psi\rangle \in \mathcal{S}(\mathcal{H})$, the min-entropy $H_\infty(p)$ where p is obtained by applying Born's rule when measuring the state $|\psi\rangle$ with $\mathbf{S}$ is,*

$$H_\infty(p) \geq -\log(3/4) \tag{3.9}$$

and this bound is achieved at some state. For the collision entropy $H_2(p)$ we have the lower bound

$$H_2(p) \geq -\log(7/12), \tag{3.10}$$

which is also achieved at some state. In particular this implies, that there is uncertainty in a tuple of parities when carrying Protocol 1.

Proof. To find the min-entropy of a distribution it is enough to find the maximum of its event probabilities. Then we are interested in the greatest value we can achieve when applying Born's rule to $\mathbf{M}$, namely we need to study the expressions

$$\langle\psi|(|0\rangle \pm |1\rangle \pm |2\rangle)(\langle 0| \pm \langle 1| \pm \langle 2|)|\psi\rangle\,. \tag{3.11}$$

This however corresponds to computing the magnitude of the projection of $|\psi\rangle$ into $\mathrm{span}(|0\rangle \pm |1\rangle \pm |2\rangle)$. And therefore this maximum is going to be achieved when we take a unitary vector in this subspace of $\mathcal{H}$. For instance if we choose

$$|\psi_{ab}\rangle = \frac{1}{\sqrt{3}}(|0\rangle + (-1)^a |1\rangle + (-1)^b |2\rangle),$$

we get for every $\frac{1}{4}M_{ab}$, that $\langle\psi_{ab}|\frac{1}{4}M_{ab}|\psi_{ab}\rangle = 3/4$. Therefore the min-entropy of any distribution induced by the Born's rule is going to be bounded by below by $-\log(3/4)$.
Alternatively we can obtain this result by checking that the largest eigenvalue of $\frac{1}{4}M_{ab}$ for every a and b is $3/4$.

For the second claim one can check by the method of Lagrange Multipliers that the vector $|\psi_{ab}\rangle$ again maximises the expression

$$\sum_{a,b} \langle\psi|\frac{1}{4}M_{ab}|\psi\rangle^2,$$

and for every such vector, the induced distribution is $(3/4, 1/12, 1/12, 1/12)$ up to permutation of coordinates. This vector has a collision entropy of $-\log\frac{7}{12}$, from which we obtain the result. □

Notice, that for parities of the type $\mathcal{X} = (z_0 \oplus z_1, z_1 \oplus z_2)$, if we let $S'_{ab} = \sum_{\substack{z_0+z_1=a \\ z_1+z_2=b}} Q_z$, then

$$S'_{0b} = S_{0b}, \tag{3.12}$$
$$S'_{1b} = S_{1(1-b)}. \tag{3.13}$$

Therefore the POVM defined by the S'_{ab} is a permutation of the one defined by the S_{ab}, and the above proposition is still valid for this type of parities.

For completeness, we finish with a lower bound for parities of the type $\mathcal{X} = \{z_0 \oplus z_1, z_2 \oplus z_3\}$. We can do the analogous computations as in the previous case to check that the operators,

$$S''_{ab} = \sum_{\substack{z_0+z_1=a \\ z_2+z_3=b}} Q_z, \tag{3.14}$$

are of the shape,

$$S''_{ab} = \frac{1}{4} M_a \oplus M_b \oplus \mathbb{I}. \tag{3.15}$$

Where $M_a = (|0\rangle + (-1)^a |1\rangle)(\langle 0| + (-1)^a \langle 1|)$. Therefore the best lower bound for the min-entropy of the distributions obtained by these operators, is the one achieved at $|\psi_a\rangle = \frac{1}{\sqrt{2}} |0\rangle + (-1)^a |1\rangle$, which gives $\langle \psi_a | \frac{1}{4} M_a | \psi_a \rangle = \frac{1}{2}$. Therefore we have for every p obtained by Born's rule applied to S'_{ab},

$$H_\infty(p) \geq -\log \frac{1}{2}, \tag{3.16}$$

which for the logarithm in base 2, corresponds to 1 bit of uncertainty.

We can also consider the case of n bit parities of the type $\mathcal{X} = (z_0 \oplus z_1, \ldots, z_0 \oplus z_n)$ with $n < d$. In this case, we can define the measurement $\mathbf{S}_n$ by the operators,

$$S_{a_1,\ldots,a_n} = \sum_{\substack{z_0 \oplus z_1 = a_1 \\ \cdots \\ z_0 \oplus z_n = a_n}} Q_z. \tag{3.17}$$

And by proceeding analogously as in the case $n = 2$ we conclude that, $S = \frac{1}{2^n}(M_{a_1,\ldots,a_n} \oplus \mathbb{I})$, where

$$M_{a_1,\ldots,a_n} = (|0\rangle + \sum_{i=1}^{n} (-1)^{a_i} |i\rangle)(\langle 0| + \sum_{i=1}^{n} (-1)^{a_i} \langle i|). \tag{3.18}$$

Now the vectors that maximise the expressions $\langle \psi | M_{a_1,\ldots,a_n} | \psi \rangle$ are the

$$|\tilde{\psi}_a\rangle = \frac{1}{\sqrt{n+1}} (|0\rangle + \sum_{i=1}^{n} (-1)^{a_i} |i\rangle), \tag{3.19}$$

for every $a \in \{0,1\}^n$. And these give us $\langle \tilde{\psi}_a | M_{a_1,\ldots,a_n} | \tilde{\psi}_a \rangle = n + 1$, by which we conclude that,

$$H_\infty(p) \geq -\log \frac{n+1}{2^n}, \tag{3.20}$$

for any distribution p obtained from Born's rule applied to the measurement $\mathbf{S}_n$.
In particular (3.20), implies that the best bound for $H_\infty(p)$ is an $O(n)$ function, where n is the number of parities considered.

In conclusion, a dishonest user can reduce the uncertainty about two bit parities by more than one bit in the worst case scenario. This is still not enough to retrieve two parities with probability 1. This suggests then, that RRDPS can be applied to implement an OT protocol.

Chapter 4

Uncertainty Relations for Single POVMs

4.1 Introduction

All the previous considerations, give motivation for the study of uncertainty relations for single measurements. Our aim is then, to find a framework in which we can talk about how uncertainty relations behave under operations like the direct sum or the direct product, and to see up to which extent we can find information about the uncertainty of a general POVM.

In the following, we will consider $\mathcal{H}$ to be a d-dimensional $\mathbb{C}$-Hilbert space, e.g. $\mathcal{H} = \mathbb{C}^d$. Unless stated otherwise, $|\psi\rangle$ will refer to a state in the Hilbert space, i.e. $|\psi\rangle \in \mathcal{S}(\mathcal{H})$.
We have seen that by Born's rule, a measurement induces a probability distribution for every state $|\psi\rangle$. Therefore one can talk about the entropy of such induced distributions. This motivates the following definition,

Definition 21. If $\mathbf{X}$ is a POVM, and $|\psi\rangle$ is any state, we write

$$H_\alpha(\mathbf{X}|\psi) = H_\alpha(p_{|\psi\rangle}) = \frac{1}{1-\alpha} \log \sum_i \langle\psi|X_i|\psi\rangle^\alpha \,, \tag{4.1}$$

where $p_{|\psi\rangle} = (\langle\psi|X_1|\psi\rangle \,, \ldots, \langle\psi|X_n|\psi\rangle)$, is the probability distribution induced by $\mathbf{X}$ when measuring $|\psi\rangle$.

For a PVM, we can always find a state vector which will be fixed under one of the projections. For these vectors their induced probability distribution will be, (up to permutation of the coordinates), equal to $(1, 0, \ldots, 0)$. Then we conclude that for a projective measurement we can always find an induced distribution with zero entropy. Given random variables $X_1, \ldots, X_n$, will call an inequality of the type $H(X_1) + \cdots + H(X_n) \geq c$, an *uncertainty relation* for the random variables.

It becomes clear then that the best possible uncertainty relations for a single PVM are trivial, namely $H(\mathbf{P}|\psi) \geq 0$, for every $|\psi\rangle$. However we have seen in the previous chapter that for general measurements this is not true, and that we can have uncertainty in every induced distribution by measuring a state. Therefore it is interesting to consider uncertainty relations of the type $H(\mathbf{X}|\psi) \geq c$, where c does not depend on $|\psi\rangle$.

The theory of uncertainty relations can be traced back to Heisenberg's uncertainty principle, which informally states that we cannot prepare a quantum state in which we can tightly control both its position and momentum. These considerations made by Heisenberg were purely heuristic, and were formalised for the first time by Kennard [11], giving

$$\sigma(P)\sigma(Q) \geq \frac{\hbar}{2}, \tag{4.2}$$

where $\sigma(P)$, $\sigma(Q)$ denote the standard deviation of the position and momentum respectively, and $\hbar$ is the Planck constant.
In general, for any pair of projective measurements X and Z, we have the following state-dependent relation due to Robertson [17],

$$\sigma(X)\sigma(Z) \geq \frac{1}{2} |\, \langle\psi|[X, Z]|\psi\rangle \,|, \tag{4.3}$$

where $[A, B] = AB - BA$ denotes the commutator. These relations are interesting because in the former $\hbar$ appears as a limit to our knowledge, and in the later the commutator is identified as an object of interest when considering two-measurements uncertainty relations.
The above uncertainty relations have their entropic analogues. A formulation employing entropy is preferable, since the standard deviation is a measure of dispersion rather than uncertainty. For the PVMs associated to the position and momentum observables, we have the result due to Białnicki-Birula and Mycielski [4],

$$h(Q) + h(P) \geq \log(e\pi\hbar). \tag{4.4}$$

It turns out that this entropic relation implies (4.2). The corresponding generalisation of (4.3) is due to Massen and Uffink [13], and it is one of the most well known entropic uncertainty relations,

$$H(X) + H(Z) \geq \log \frac{1}{c} \tag{4.5}$$

where H is the Shannon entropy and $c = \max_{j,k} |\langle a_j | b_k \rangle|$, is the maximum overlap between any two eigenvectors of the PVMs X and Z.
Subsequently, Krishna and Parthasarathy [12] extended this result to general measurements by using Naimark's theorem. Namely they showed that for every pair of general measurements $\mathbf{X}$ and $\mathbf{Y}$ the following state-independent relation holds,

$$H(\mathbf{X}|\psi) + H(\mathbf{Y}|\psi) \geq -2 \log \max_{i,j} \|X_i^{1/2} Y_j^{1/2}\|, \tag{4.6}$$

where H is the Shannon entropy. With this they obtained the first uncertainty relation for a single POVM, namely

$$H(\mathbf{X}|\psi) \geq -\log \max_{i,j} \|X_i^{1/2} X_j^{1/2}\|. \tag{4.7}$$

4.2 Elementary properties

We will be interested in sharp and state-indpendent uncertainty relations for a given POVM, i.e. lower bounds of entropy valid for every state, and attainable at a particular state. Therefore we are drawn to consider for any $\alpha \geq 0$, the following optimisation problem,

$$\inf_{|\psi\rangle} H_\alpha(\mathbf{X}|\psi). \tag{4.8}$$

But now, since $\mathcal{S}(\mathcal{H})$ is compact, we have

$$\inf_{|\psi\rangle} H_\alpha(\mathbf{X}|\psi) = \min_{|\psi\rangle} H_\alpha(\mathbf{X}|\psi). \tag{4.9}$$

Therefore by construction, $\min_{|\psi\rangle} H_\alpha(\mathbf{X}|\psi)$ is the sharp uncertainty relation we were looking for, and we may give the following definition.

Definition 22. Given a POVM $\mathbf{X} = \{X_i\}_{i \in \mathcal{I}}$ we define for every $\alpha > 1$, the minimal α-Rényi Entropy of $\mathbf{X}$ is defined as

$$H_\alpha(\mathbf{X}) = \min_{|\psi\rangle} H_\alpha(\mathbf{X}|\psi). \tag{4.10}$$

With this definition, we will study how the function $H_\alpha(\mathbf{X})$ behaves under operations like $\oplus$, $\otimes$.

First, one can see that the function $H_{(\cdot)}(\cdot) : [0, \infty] \times \mathcal{POVM}_\mathcal{I} \to \mathbb{R}$, inherits many properties of $H_\alpha(\mathbf{X}|\psi)$ and thus from the Rényi entropy. We can prove for instance that it is a continuous function,

Proposition 2. The function $H_\alpha(\mathbf{X})$ is continuous on $[0, \infty] \times \mathcal{POVM}_\mathcal{I}(\mathcal{H})$.

Proof. Let $n = |\mathcal{I}|$, then the function $\mathcal{POVM}_\mathcal{I} \times \mathcal{S}(\mathcal{H}) \to \Delta_{n-1}$ given by

$$|\psi\rangle \to (\langle\psi|X_1|\psi\rangle, \ldots, \langle\psi|X_n|\psi\rangle), \tag{4.11}$$

is continuous, since it is a polynomial function in each one of its coordinates.
Therefore since $H_\alpha(p)$ is continuous on $(\alpha, p) \in [0, \infty] \times \Delta_{n-1}$, we have that its composition with (4.11) is a continuous function from $[0, \infty] \times \mathcal{POVM}_\mathcal{I} \times \mathcal{S}(\mathcal{H})$ to $\mathbb{R}$. But this composition is precisely the

α-Rényi entropy $H_\alpha(\mathbf{X}|\psi)$. Therefore it is a continuous function on $(\alpha, \mathbf{X}, |\psi\rangle)$ with range in $\mathbb{R}$, and by compactness of $\mathcal{S}(\mathcal{H})$, the function

$$H_\alpha(\mathbf{X}) = \min_{|\psi\rangle} H_\alpha(\mathbf{X}|\psi),$$

is continuous. □

Remark 1. The minimal α-Rényi entropy of $\mathbf{X}$ inherits also the following properties from the α-Rényi entropy

- For $\mathbf{X}$ fixed, $H_\alpha(\mathbf{X})$ is a monotone non-increasing function on α. If we let $\alpha \leq \beta$, and $|\psi_\alpha\rangle$ be a state vector such that $H_\alpha(\mathbf{X}) = H_\alpha(\mathbf{X}|\psi_\alpha)$, then we have,

$$H_\alpha(\mathbf{X}) = H_\alpha(\mathbf{X}|\psi_\alpha) \geq H_\beta(\mathbf{X}|\psi_\alpha) \geq H_\beta(\mathbf{X}). \tag{4.12}$$

- For $\mathbf{X}$ fixed, we have for every $\alpha > 1$,

$$\frac{\alpha}{\alpha - 1} H_\infty(\mathbf{X}) \geq H_\alpha(\mathbf{X}) \geq H_\infty(\mathbf{X}). \tag{4.13}$$

The function $\frac{1}{1-\alpha}\log(x)$ is decreasing in x for $\alpha > 1$, and increasing in x for $\alpha < 1$. Therefore we can rewrite the minimal entropy as

$$H_\alpha(\mathbf{X}) = \frac{1}{1-\alpha} \log\left(\max_{|\psi\rangle} \sum_i \langle\psi|X_i|\psi\rangle^\alpha\right), \tag{4.14}$$

in the case $\alpha > 1$, and as

$$H_\alpha(\mathbf{X}) = \frac{1}{1-\alpha} \log\left(\min_{|\psi\rangle} \sum_i \langle\psi|X_i|\psi\rangle^\alpha\right), \tag{4.15}$$

in the case $\alpha < 1$. We will use this simple remark extensively as it allows us to work instead with the expression $\sum_i \langle\psi|X_i|\psi\rangle^\alpha$.

As we have seen in Chapter 2 it is interesting to consider the tensor product of two POVMs since this is related to taking measurements in the separate subsystems of a joint quantum system. We can obtain an uncertainty relation for the entropy of the tensor product of two POVMs from the uncertainty of the individual POVMs. We will need the following lemma,

Lemma 3. Let $\mathbf{X}$ be a POVM on $\mathcal{H}_1$ and let $\mathbb{I}$ be the identity operator on $\mathcal{H}_2$. Then we have that for every $|\psi\rangle \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$, and every $\alpha > 1$

$$\sum_i \langle\psi|X_i \otimes \mathbb{I}|\psi\rangle^\alpha \leq \max_{|\psi\rangle\in\mathcal{S}(\mathcal{H}_1)} \sum_i \langle\psi|X_i|\psi\rangle^\alpha . \tag{4.16}$$

For $0 \leq \alpha < 1$ we have

$$\sum_i \langle\psi|X_i \otimes \mathbb{I}|\psi\rangle^\alpha \geq \min_{|\psi\rangle\in\mathcal{S}(\mathcal{H}_1)} \sum_i \langle\psi|X_i|\psi\rangle^\alpha . \tag{4.17}$$

Proof. If $|\psi\rangle \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$, then we can write $|\psi\rangle$ as

$$|\psi\rangle = \sum_i \alpha_i |\psi_i\rangle |i\rangle , \tag{4.18}$$

where the $\alpha_i \in \mathbb{C}$ are such that $\sum_i |\alpha_i|^2 = 1$, and $|\psi_i\rangle \in \mathcal{S}(\mathcal{H}_1)$, for every i. So we have for every k,

$$\begin{aligned}\langle\psi|X_k \otimes \mathbb{I}|\psi\rangle &= \sum_{i,j} \overline{\alpha_i}\alpha_j \langle\psi_i|X_k|\psi_j\rangle \langle i|j\rangle \\ &= \sum_i |\alpha_i|^2 \langle\psi_i|X_k|\psi_i\rangle \leq \max_i \langle\psi_i|X_k|\psi_i\rangle .\end{aligned}$$

We have then

$$\sum_k \langle\psi|X_k \otimes \mathbb{I}|\psi\rangle^\alpha \leq \max_i \sum_k \langle\psi_i|X_k|\psi_i\rangle^\alpha \leq \max_{|\psi\rangle} \sum_k \langle\psi|X_k|\psi\rangle .$$

The proof in the case $0 \leq \alpha < 1$ is analogous. □

Proposition 3. *Let $\mathbf{X} = \{X_i\}_i$ and $\mathbf{Y} = \{Y_j\}_j$ be two POVMs over $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, then for every α,*

$$H_\alpha(\mathbf{X} \otimes \mathbf{Y}) = H_\alpha(\mathbf{X}) + H_\alpha(\mathbf{Y}).$$

Proof. Let $\alpha > 1$, then clearly

$$\begin{aligned}\left[\max_{|\psi_1\rangle} \sum_i \langle\psi_1|X_i|\psi_1\rangle^\alpha\right] \cdot \left[\max_{|\psi_2\rangle} \sum_j \langle\psi_2|Y_j|\psi_2\rangle^\alpha\right] &= \max_{|\psi_1\rangle\otimes|\psi_2\rangle} \sum_{i,j} \langle\psi_1\psi_2|X_i \otimes Y_j|\psi_1\psi_2\rangle^\alpha \\ &\leq \max_{|\psi\rangle\in\mathcal{S}(\mathcal{H}_1\otimes\mathcal{H}_2)} \sum_{i,j} \langle\psi|X_i \otimes Y_j|\psi\rangle^\alpha .\end{aligned}$$

For the reverse inequality, let X_i' be such that

$$X_i'^\dagger X_i' = X_i,$$

and for every $|\psi\rangle \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$, let

$$|\psi'^{(i)}\rangle = \frac{1}{\sqrt{\langle\psi|X_i|\psi\rangle}} X_i' \otimes \mathbb{I}\,|\psi\rangle ,$$

be the post-measurement state obtained by measuring $|\psi\rangle$ with $X_i' \otimes \mathbb{I}$.
Then for every i, j

$$\langle\psi|X_i \otimes Y_j|\psi\rangle = \langle\psi|X_i \otimes \mathbb{I}|\psi\rangle \cdot \langle\psi'^{(i)}|\mathbb{I} \otimes Y_j|\psi'^{(i)}\rangle .$$

Therefore for all $|\psi\rangle \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$, we have by the previous lemma that,

$$\begin{aligned}\sum_{i,j} \langle\psi|X_i \otimes Y_j|\psi\rangle^\alpha &= \sum_{i,j} \langle\psi|X_i \otimes \mathbb{I}|\psi\rangle^\alpha \cdot \langle\psi'^{(i)}|\mathbb{I} \otimes Y_j|\psi'^{(i)}\rangle^\alpha \\ &= \sum_i \langle\psi|X_i \otimes \mathbb{I}|\psi\rangle^\alpha \cdot \sum_j \langle\psi'^{(i)}|\mathbb{I} \otimes Y_j|\psi'^{(i)}\rangle^\alpha \\ &\leq \sum_i \langle\psi|X_i \otimes \mathbb{I}|\psi\rangle^\alpha \cdot \max_{|\phi\rangle} \sum_j \langle\phi|Y_j|\phi\rangle^\alpha \\ &\leq \max_{|\varphi\rangle} \sum_i \langle\varphi|X_i|\varphi\rangle^\alpha \cdot \max_{|\phi\rangle} \sum_j \langle\phi|Y_j|\phi\rangle^\alpha ,\end{aligned}$$

and thus

$$\left[\max_{|\varphi\rangle} \sum_i \langle\varphi|X_i|\varphi\rangle^\alpha\right] \cdot \left[\max_{|\phi\rangle} \sum_j \langle\phi|Y_j|\phi\rangle^\alpha\right] = \max_{|\psi\rangle\in\mathcal{S}(\mathcal{H}_1\otimes\mathcal{H}_2)} \sum_{i,j} \langle\psi|X_i \otimes Y_j|\psi\rangle^\alpha ,$$

which proves the claim.
We can prove the result for $0 \leq \alpha < 1$ analogously, by writing min instead of max and changing the direction of the inequalities. Finally by continuity of $H_\alpha(\mathbf{X})$, we can extend the result to $\alpha = 1$ and $\alpha = \infty$.

□

Corollary 2. *Let $\mathbf{X} = \{X_i\}_i$ be a POVM, and $n \geq 1$ then for every $\alpha \geq 0$,*

$$H_\alpha(\mathbf{X}^{\otimes n}) = nH_\alpha(\mathbf{X}).$$

4.3 An explicit uncertainty relation for general measurements

We can use Jensen's inequality to obtain a closed formula for a state-independent uncertainty relation for the α-Rényi entropy of general POVMs, for any $\alpha \in [0, \infty]$.
This relation turns out to be sharp in the case of simultaneously diagonalisable measurements.

Theorem 5. *Let* $\mathbf{X}$ *be a POVM in a* d*-dimensional* $\mathbb{C}$*-Hilbert space* $\mathcal{H}$*, then for every* $\alpha > 1$*,*

$$H_\alpha(\mathbf{X}) \geq \frac{1}{1-\alpha} \log \lambda_1^{\downarrow} \left(\sum_i X_i^\alpha \right) = \frac{1}{1-\alpha} \log \left\| \sum_i X_i^\alpha \right\|, \tag{4.19}$$

and for every $0 < \alpha < 1$*,*

$$H_\alpha(\mathbf{X}) \geq \frac{1}{1-\alpha} \log \lambda_d^{\downarrow} \left(\sum_i X_i^\alpha \right). \tag{4.20}$$

Here $\|\cdot\|$ *denotes the operator norm, and* $\lambda^{\downarrow}(A)$ *is the vector of eigenvalues of* A*, counting multiplicities, in descending order.*

Proof. Consider for every i the spectral decomposition,

$$X_i = \sum_j \lambda_j^{(i)} \, |e_j^{(i)}\rangle \langle e_j^{(i)}| \,,$$

where $\{|e_1^{(i)}\rangle, \ldots, |e_d^{(i)}\rangle\}$ is an orthonormal basis for $\mathcal{H}$. Then we can write

$$\langle\psi|X_i|\psi\rangle^\alpha = \left(\sum_j \lambda_j^{(i)} | \langle\psi|e_j^{(i)}\rangle|^2 \right)^\alpha .$$

Notice that $\sum_j |\langle\psi|e_j^{(i)}\rangle|^2 = 1$ for every i. Therefore the $|\langle\psi|e_j^{(i)}\rangle|^2$ define a probability distribution on the set $\{1, \ldots, j, \ldots, d\}$, and so by convexity of $x \mapsto x^\alpha$ when $\alpha > 1$, we have by Jensen's inequality that

$$\left(\sum_j \lambda_j^{(i)} | \langle\psi|e_j^{(i)}\rangle|^2 \right)^\alpha \leq \sum_j (\lambda_j^{(i)})^\alpha | \langle\psi|e_j^{(i)}|\psi\rangle|^2. \tag{4.21}$$

And then by the definition of the X_i^α, we can rewrite the above inequality as,

$$\langle\psi|X_i|\psi\rangle^\alpha \leq \langle\psi|X_i^\alpha|\psi\rangle \,. \tag{4.22}$$

Then we by applying this bound for every i we get,

$$\sum_i \langle\psi|X_i|\psi\rangle^\alpha \leq \sum_i \langle\psi|X_i^\alpha|\psi\rangle = \langle\psi| \sum_i X_i^\alpha |\psi\rangle \leq \left\| \sum_i X_i^\alpha \right\|.$$

Therefore, by taking the α-Rényi entropy,

$$H_\alpha(\mathbf{X}|\psi) = \frac{1}{1-\alpha} \log \left(\sum_i \langle\psi|X_i|\psi\rangle^\alpha \right) \geq \frac{1}{1-\alpha} \log \left\| \sum_i X_i^\alpha \right\|,$$

which proves (4.19).
Now for $0 < \alpha < 1$ we have that $x \mapsto x^\alpha$ is concave and, arguing analogously, Jensen's inequality provides us with the inequality in the other direction,

$$\langle\psi|X_i|\psi\rangle^\alpha \geq \langle\psi|X_i^\alpha|\psi\rangle \,.$$

So we have that by the minimax principle

$$\sum_i \langle\psi|X_i|\psi\rangle^\alpha \geq \sum_i \langle\psi|X_i^\alpha|\psi\rangle = \langle\psi| \sum_i X_i^\alpha |\psi\rangle \geq \lambda_d^{\downarrow} \left(\sum_i X_i^\alpha \right).$$

Thus by taking the α-Rényi entropy,

$$H_\alpha(\mathbf{X}|\psi) = \frac{1}{1-\alpha} \log \left(\sum_i \langle\psi|X_i|\psi\rangle^\alpha \right) \geq \frac{1}{1-\alpha} \log \lambda_d^{\downarrow} \left(\sum_i X_i^\alpha \right).$$

□

As stated before, we can show that the above bounds for $H_\alpha(\mathbf{X})$ turns out to be an equality for simultaneously diagonalisable measurements. Namely we will find for such measurements, a state vector that achieves the bound.

Corollary 3. If $\mathbf{X} = \{X_i\}_{i\in\mathcal{I}}$ is a POVM such that $X_iX_j = X_jX_i$ for all $i, j \in \mathcal{I}$, then for $\alpha > 1$

$$H_\alpha(\mathbf{X}) = \frac{1}{1-\alpha}\log\left\|\sum_i X_i^\alpha\right\|,$$

and for $0 < \alpha < 1$,

$$H_\alpha(\mathbf{X}) = \frac{1}{1-\alpha}\log\lambda_d^\downarrow\left(\sum_i X_i^\alpha\right).$$

Proof. Since $X_iX_j = X_jX_i$ for all $i, j \in \mathcal{I}$, the X_i are simultaneously diagonalisable. Therefore for $\alpha > 1$,

$$\max_{|\psi\rangle}\sum_i \langle\psi|X_i|\psi\rangle^\alpha = \max_{|\psi\rangle}\sum_{i,j}(\lambda_j^{(i)})^\alpha(\langle\psi|\,|j\rangle\,\langle j|\,|\psi\rangle)^\alpha, \tag{4.23}$$

where $X_i = \sum_j \lambda_j^{(i)}\,|j\rangle\,\langle j|$. Write $|\psi\rangle = \sum_i \psi_i\,|i\rangle$, with $\psi_i \in \mathbb{C}$, then

$$\sum_i \langle\psi|X_i|\psi\rangle^\alpha = \sum_{i,j}(|\psi_j|^2\lambda_j^{(i)})^\alpha = \sum_j |\psi_j|^{2\alpha}\sum_i(\lambda_j^{(i)})^\alpha,$$

therefore if we let $\ell = \arg\max_j \sum_i(\lambda_j^{(i)})^\alpha$, then taking $\tilde{\psi}_\ell = 1$ and $\tilde{\psi}_j = 0$ for all $j \neq \ell$, then

$$\sum_i \langle\tilde{\psi}|X_i|\tilde{\psi}\rangle^\alpha = \left\|\sum_i X_i^\alpha\right\|,$$

which proves the sharpness of relation (4.19).
We can prove the sharpness of relation (4.20), for $0 < \alpha < 1$ analogously by taking $\ell = \arg\min_j \sum_i(\lambda_j^{(i)})^\alpha$. □

Theorem 5 introduces the operator $\sum_i X_i^\alpha$ as an object of interest to find uncertainty relations. We know that for a PVM $\mathbf{P}$, we have $H_\alpha(\mathbf{P}) = 0$. Still, it is interesting to see the relation between this fact and Theorem 5, so notice that

$$\sum_i P_i^\alpha = \sum_i P_i = \mathbb{I}. \tag{4.24}$$

So we have that, since all the eigenvalues of $\mathbb{I}$ are 1, taking the logarithm will give us zero in (4.19) and (4.20). Equation (4.24), follows from the fact that the set of eigenvalues $\mathrm{eig}(P_i)$, of P_i is contained in $\{0, 1\}$, therefore for every $\alpha > 0$,

$$P_i^\alpha = \sum_j(\lambda_j^{(i)})^\alpha\,|j\rangle\,\langle j| = \sum_j \lambda_j^{(i)}\,|j\rangle\,\langle j| = P_i.$$

Corollary 3 indicates that simultaneously diagonalisable measurements achieve the bounds (4.19) and (4.20). We can also relate the fact that $H_\alpha(\mathbf{P}) = 0$ for a PVM to this result because PVMs are simultaneously diagonalisable.
We can see then that the amount of commutativity within the operators of a POVM, seems to be a key factor to quantify the amount of uncertainty that these measurements have.

Remark 2. For a POVM $\mathbf{X} = \{X_i\}$, where each X_i is expressible as dilations of projections, (i.e. $X_i^2 = \mu X_i$) we have for every $\alpha \in (0, \infty) - \{1\}$,

$$\frac{1}{1-\alpha}\log\left\|\sum_i X_i^\alpha\right\| = \frac{1}{1-\alpha}\log\lambda_d^\downarrow\left(\sum_i X_i^\alpha\right) = H_\infty(\mathbf{X}).$$

In other words the relations (4.19) and (4.20) do not give any new information for this kind of measurements, since we always have $H_\alpha(\mathbf{X}) \geq H_\infty(\mathbf{X})$.

Proof. We have that if $X_i^2 = \mu X_i$ with $\mu \neq 0$, then

$$\sum_j \lambda_j^{(i)^2} |e_j^{(i)}\rangle \langle e_j^{(i)}| = \sum_j \mu \lambda_j^{(i)} |e_j^{(i)}\rangle \langle e_j^{(i)}| . \tag{4.25}$$

Therefore for every i and j, $\lambda_j^{(i)^2} = \mu \lambda_j^{(i)}$, by which we conclude that $\lambda_j^{(i)} \in \{0, \mu\}$ for every i, j. Then we can write

$$X_i = \sum_{\lambda_j^{(i)} \neq 0} \mu |e_j^{(i)}\rangle \langle e_j^{(i)}| , \tag{4.26}$$

and then,

$$X_i^\alpha = \sum_{\lambda_j^{(i)} \neq 0} \mu^\alpha |e_j^{(i)}\rangle \langle e_j^{(i)}| = \mu^{\alpha-1} \sum_{\lambda_j^{(i)} \neq 0} \mu |e_j^{(i)}\rangle \langle e_j^{(i)}| = \mu^{\alpha-1} X_i. \tag{4.27}$$

So we have that for every $\alpha > 0$,

$$\sum_i X_i^\alpha = \mu^{\alpha-1} \sum_i X_i = \mu^{\alpha-1} \cdot \mathbb{I}.$$

And so

$$\frac{1}{1-\alpha} \log \left\| \sum_i X_i^\alpha \right\| = \frac{1}{1-\alpha} \log \lambda_d^\downarrow \left(\sum_i X_i^\alpha \right) = \frac{1}{1-\alpha} \log(\mu^{\alpha-1}) = -\log \mu = H_\infty(\mathbf{X}). \tag{4.28}$$

□

As a consequence of this remark we have that if $\mathbf{X} = \mathbf{A} \oplus \mathbf{B}$ where $\mathbf{A}$ and $\mathbf{B}$ are POVMs where $A_i^2 = \lambda A_i$ and $B_i^2 = \mu B_i$ for all i, then

$$\sum_i X_i^\alpha = \sum_i (\lambda^{\alpha-1} A_i) \oplus (\mu^{\alpha-1} B_i) = (\lambda^{\alpha-1} \cdot \mathbb{I}) \oplus (\mu^{\alpha-1} \cdot \mathbb{I}).$$

And if $1 \geq \lambda > \mu$, then for $\alpha > 1$, $\lambda^{\alpha-1} > \mu^{\alpha-1}$, and for $0 < \alpha < 1$, $\lambda^{\alpha-1} < \mu^{\alpha-1}$. So in this case the relations (4.19) and (4.20) are also trivial. Therefore for the measurement $\mathbf{S}$, defined in (3.8) the bounds do not give any new information.

The measurements in the above remark are in general, not simultaneously diagonalisable when $\mu > 0$, otherwise we would have by Corollary 3 that $H_\alpha(\mathbf{X}) = H_\infty(\mathbf{X})$ for every $\alpha > 0$.
Also, the uncertainty relation given by Krishna and Parthasarathy (4.7) does not give any new information for this type of measurements. Since when we have $X_i^2 = \mu X_i$, then $X_i^{1/2} = \frac{1}{\sqrt{\mu}} X_i$. And so we have by Lemma 2 that for every i and j,

$$\|X_i^{1/2} X_j^{1/2}\| = \frac{1}{\mu} \|X_i X_j\| \leq \frac{1}{\mu} \|X_i\| \|X_j\| = \frac{1}{\mu} \mu^2 = \mu, \tag{4.29}$$

and this bound is achieved when $i = j$, so we conclude

$$H(\mathbf{X}) \geq -\log \max_{i,j} \|X_i^{1/2} X_j^{1/2}\| = -\log \mu = H_\infty(\mathbf{X}).$$

Considering the above remarks, one can ask the question of what is the correct generalisation of Theorem 5 so that we obtain more information in cases like the above. To do so it might be of interest to consider measures of the degree of noncommutativity between pairs of operators in a POVM. As we have seen in (4.3), the commutator is an interesting object to consider, it might also prove useful then, to obtain uncertainty relations for the Rényi entropy.

4.4 Norms on POVMs

The minimisation of the Rényi entropy in $\mathcal{S}(\mathcal{H})$ leads us to maximise $\sum_i \langle \psi | X_i | \psi \rangle^\alpha$. The latter expression is closely related to the norm of the induced $|\psi\rangle$ distribution, so this suggest that we may define a norm on POVMs, or in general on tuples of linear bounded operators, by maximising the state induced norms.

Definition 23. Let $\mathbf{A} \in \mathcal{L}(\mathcal{H})^n$ and $\alpha > 1$. We define its $|\psi\rangle$-*dependent* α-*Rényi pseudo-norm* as

$$\|\mathbf{A}\|_\alpha^{|\psi\rangle} = \left(\sum_{i=1}^{n} |\langle\psi|A_i|\psi\rangle|^\alpha \right)^{1/\alpha}. \tag{4.30}$$

We define its α-*Rényi norm* as

$$\|\mathbf{A}\|_\alpha = \max_{|\psi\rangle} \|\mathbf{A}\|_\alpha^{|\psi\rangle}. \tag{4.31}$$

For every $|\psi\rangle \in \mathcal{S}(\mathcal{H})$ and every $\alpha > 0$, $\|\cdot\|_\alpha^{|\psi\rangle}$ is indeed a *pseudo-norm*, i.e. absolutely homogeneous and satisfying the triangle inequality. Notice that

$$\|A\|_\alpha^{|\psi\rangle} = \|(\gamma_1, \ldots, \gamma_n)\|_\alpha,$$

where $\gamma_i := \langle\psi|A_i|\psi\rangle$. So for $|\psi\rangle$ fixed, $\|\cdot\|_\alpha^{|\psi\rangle}$ satisfies absolute homogeneity and the triangle inequality. We cannot say in general that $\|\mathbf{A}\|_\alpha^{|\psi\rangle} = 0$ if and only if $\mathbf{A} = \mathbf{0}$. It is enough to consider an n-tuple of matrices having kernel containing $|\psi\rangle$.
On the other hand, $\|\cdot\|_\alpha$ defines a norm on $\mathcal{L}(\mathcal{H})^n$; the triangle inequality and absolute homogeneity are inherited from $\|\cdot\|_\alpha^{|\psi\rangle}$, and if we have $\|\mathbf{A}\|_\alpha = 0$ then

$$\|\mathbf{A}\|_\alpha = \max_{|\psi\rangle} \left(\sum_i |\langle\psi|A_i|\psi\rangle|^\alpha \right)^{1/\alpha} = 0.$$

And thus for every i and every $|\psi\rangle$ we have $\langle\psi|A_i|\psi\rangle = 0$, which implies $A_i = 0$.
Also, since the α-norm on real vectors is monotone increasing in α it follows that $\|\mathbf{A}\|_\alpha$ is also monotone increasing in α.

We can also extend the definition to $0 < \alpha < 1$ by putting min instead of max, although note that this is not a norm, since $\|\cdot\|_\alpha$ does not satisfy the triangle inequality in $\mathbb{C}^n$. This allows us to write for every POVM $\mathbf{X}$ and $\alpha > 0$,

$$H_\alpha(\mathbf{X}|\psi) = \frac{\alpha}{1-\alpha} \log \|\mathbf{X}\|_\alpha^{|\psi\rangle}, \tag{4.32}$$

and,

$$H_\alpha(\mathbf{X}) = \frac{\alpha}{1-\alpha} \log \|\mathbf{X}\|_\alpha. \tag{4.33}$$

These equations are analogous to the expression for random variables

$$H_\alpha(X) = \frac{\alpha}{1-\alpha} \log \|p\|_\alpha, \tag{4.34}$$

where p is the distribution of X.

The example we considered in (3.8) motivates the study of uncertainty relations for the direct sum of two measurements. With this notion of norm we can prove the following result for any pair of POVMs,

Theorem 6. Given two POVMs $\mathbf{X} = \{X_i\}_{i\in\mathcal{I}}$ and $\mathbf{Y} = \{Y_i\}_{i\in\mathcal{I}}$, we have for every $\alpha > 0$

$$H_\alpha(\mathbf{X} \oplus \mathbf{Y}) = \min\left(H_\alpha(\mathbf{X}), H_\alpha(\mathbf{Y})\right). \tag{4.35}$$

Proof. We just need to prove that the maximum/minimum over all states $|\psi\rangle \in \mathcal{S}(\mathcal{H}_1 \oplus \mathcal{H}_2)$ of $\sum_i \langle\psi|X_i \oplus Y_i|\psi\rangle^\alpha$, is achieved at $\mathcal{S}(\mathcal{H}_1) \oplus 0$ or $0 \oplus \mathcal{S}(\mathcal{H}_2)$.
We can easily prove the inequality for $\alpha > 0$,

$$H_\alpha(\mathbf{X} \oplus \mathbf{Y}) \leq \min(H_\alpha(\mathbf{X}), H_\alpha(\mathbf{Y})), \tag{4.36}$$

this is because

$$\max_{|\psi\rangle} \sum_i \langle\psi|X_i \oplus Y_i|\psi\rangle^\alpha \geq \max\left(\max_{|\psi_1\rangle} \sum_i \langle\psi_1|X_i|\psi_1\rangle^\alpha, \max_{|\psi_2\rangle} \sum_i \langle\psi_2|Y_i|\psi_2\rangle^\alpha \right),$$

and then since $\alpha > 1$, we have (4.36), the case $0 < \alpha < 1$ is analogous by changing the direction of the inequalities and taking min instad of max. For the converse, notice that we can write any state $|\psi\rangle \in \mathcal{S}(\mathcal{H}_1 \oplus \mathcal{H}_2)$ as convex linear combination

$$|\psi\rangle = t \cdot |\psi_1\rangle \oplus \mathbf{0} + (1-t) \cdot \mathbf{0} \oplus |\psi_2\rangle\,, \tag{4.37}$$

where $t \in [0,1]$ and $|\psi_1\rangle$, $|\psi_2\rangle$ are states in $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. Now by linearity we have that for every i,

$$\begin{aligned}\langle\psi|X_i \oplus Y_i|\psi\rangle &= t\,\langle\psi_1 \oplus \mathbf{0}|X_i \oplus Y_i|\psi_1 \oplus \mathbf{0}\rangle + (1-t)\,\langle\mathbf{0} \oplus \psi_2|X_i \oplus Y_i|\mathbf{0} \oplus \psi_2\rangle\,,\\ &= t\,\langle\psi_1|X_i|\psi_1\rangle + (1-t)\,\langle\psi_2|Y_i|\psi_2\rangle\,.\end{aligned}$$

So we can write,

$$\|\mathbf{X} \oplus \mathbf{Y}\|_\alpha^{|\psi\rangle} = \|t\gamma_1 + (1-t)\gamma_2\|_\alpha, \tag{4.38}$$

where $\gamma_1 = (\langle\psi_1|X_1|\psi_1\rangle\,, \ldots, \langle\psi_1|X_n|\psi_1\rangle)$ and $\gamma_2 = (\langle\psi_2|Y_1|\psi_2\rangle\,, \ldots, \langle\psi_2|Y_n|\psi_2\rangle)$. Now by the triangle inequality, we have

$$\begin{aligned}\|\mathbf{X} \oplus \mathbf{Y}\|_\alpha^{|\psi\rangle} &\leq t\|\gamma_1\|_\alpha + (1-t)\|\gamma_2\|_\alpha\\ &= t\|\mathbf{X}\|_\alpha^{|\psi_1\rangle} + (1-t)\|\mathbf{Y}\|_\alpha^{|\psi_2\rangle}\\ &\leq t\|\mathbf{X}\|_\alpha + (1-t)\|\mathbf{Y}\|_\alpha\\ &\leq \max(\|\mathbf{X}\|_\alpha, \|\mathbf{Y}\|_\alpha),\end{aligned}$$

and since this inequality holds for any $|\psi\rangle \in \mathcal{S}(\mathcal{H}_1 \oplus \mathcal{H}_2)$, we conclude

$$\|\mathbf{X} \oplus \mathbf{Y}\|_\alpha \leq \max(\|\mathbf{X}\|_\alpha, \|\mathbf{Y}\|_\alpha), \tag{4.39}$$

which proves the claim.
For $0 < \alpha < 1$, we have that for the "α-norm" the triangle inequality holds in the reverse direction for non-negative vectors. Therefore, if the coordinates of x and y are all non-negative, we have that

$$\|tx + (1-t)y\|_\alpha \geq t\|x\|_\alpha + (1-t)\|y\|_\alpha. \tag{4.40}$$

Therefore, by proceeding analogously, we conclude that for $0 < \alpha < 1$,

$$\|\mathbf{X} \oplus \mathbf{Y}\|_\alpha^{|\psi\rangle} \geq \min(\|\mathbf{X}\|_\alpha, \|\mathbf{Y}\|_\alpha), \tag{4.41}$$

for every $|\psi\rangle \in \mathcal{S}(\mathcal{H}_1 \oplus \mathcal{H}_2)$. This implies then that

$$\|\mathbf{X} \oplus \mathbf{Y}\|_\alpha \geq \min(\|\mathbf{X}\|_\alpha, \|\mathbf{Y}\|_\alpha), \tag{4.42}$$

which proves the claim for $0 < \alpha < 1$.
By continuity of $H_\alpha(\mathbf{X})$ we can extend this result to $\alpha = 1$. The case $\alpha = \infty$ is trivial. □

Theorem 5 motivates the following definition

Definition 24. For $\alpha > 1$, and $\mathbf{A} \in \mathcal{L}(\mathcal{H})$, we define,

$$N_\alpha^{|\psi\rangle}(\mathbf{A}) = \left(\sum_i \langle\psi|\ |A_i|^\alpha\ |\psi\rangle\right)^{1/\alpha}, \tag{4.43}$$

where $|A_i| = \sqrt{A_i^\dagger A_i}$ is the matrix modulus of A_i. And by maximising over all states we can define,

$$N_\alpha(\mathbf{A}) = \max_{|\psi\rangle} N_\alpha^{|\psi\rangle}(\mathbf{A}) = \left\|\sum_i |A_i|^\alpha\right\|^{1/\alpha}, \tag{4.44}$$

where $\|\cdot\|$ denotes the operator norm.

The functions $N_\alpha^{|\psi\rangle}$ and N_α, are clearly absolutely homogeneous, and as in the case of the α-Rényi norm, $N_\alpha(\mathbf{A}) = 0$ if and only if $\mathbf{A} = \mathbf{0}$.
It would be an interesting question to determine if N_α is indeed a norm on bounded linear operators. For diagonal operators, this is true because of the triangle inequality on $\mathbb{C}^d$, but the question of whether this is true in general remains open.

Again, we might generalise the function $N_\alpha(\mathbf{A})$ to $0 < \alpha < 1$, by writing taking min instead of max in (4.44). Therefore we can restate Theorem 5 as,

$$\|\mathbf{X}\|_\alpha \leq N_\alpha(\mathbf{X}), \tag{4.45}$$

for every $\alpha > 0$, and every POVM $\mathbf{X}$. And Corollary 3 can also be rephrased as

$$\|\mathbf{X}\|_\alpha = N_\alpha(\mathbf{X}), \tag{4.46}$$

for every POVM $\mathbf{X}$ satisfying $X_i X_j = X_j X_i$ for every i, j.

One can check that for every POVM $\mathbf{X}$, the functions $N_\alpha^{|\psi\rangle}(\mathbf{X})$, and $N_\alpha(\mathbf{X})$ are continuous on α for every $\alpha \in (0,\infty) - \{1\}$. Indeed in the case of $N_\alpha^{|\psi\rangle}(\mathbf{X})$, we have that the function is a composition of continuous functions, namely for fixed $\mathbf{X}$ and $|\psi\rangle$

$$\begin{aligned} N_\alpha^{|\psi\rangle}(\mathbf{X}) \ :(0,\infty) - \{1\} &\longrightarrow \mathcal{L}(\mathcal{H})^n \longrightarrow \mathbb{R} \longrightarrow \mathbb{R} \\ \alpha \mapsto (X_1^\alpha, \ldots, X_n^\alpha) &\mapsto \sum_i \langle\psi|X_i^\alpha|\psi\rangle \mapsto \left(\sum_i \langle\psi|X_i^\alpha|\psi\rangle\right)^{1/\alpha}, \end{aligned}$$

and in the above decomposition every function is continuous.
Then we know by compactness of $\mathcal{S}(\mathcal{H})$, that $N_\alpha(\mathbf{X})$ is continuous in α on the intervals $(0,1)$ and $(1,\infty)$.

Proposition 4. For any POVM $\mathbf{X}$, the functions $N_\alpha^{|\psi\rangle}(\mathbf{X})$ and $N_\alpha(\mathbf{X})$ are decreasing in the intervals $(0,1)$ and $(1,\infty)$.

Proof. We have for every $|\psi\rangle$,

$$\begin{aligned} \frac{\partial}{\partial\alpha} N_\alpha^{|\psi\rangle}(\mathbf{X}) &= \frac{\partial}{\partial\alpha}\left(\sum_i \langle\psi|X_i^\alpha|\psi\rangle\right)^{1/\alpha} \\ &= \frac{1}{\alpha}\left(\sum_i \langle\psi|X_i^\alpha|\psi\rangle\right)^{1/\alpha-1} \cdot \sum_i \langle\psi|X_i^\alpha \cdot \log X_i|\psi\rangle\,. \end{aligned}$$

Now since $\mathbf{X}$ is a POVM, we have that for all i, the eigenvalues of X_i are all between 0 and 1. Thus the eigenvalues of $\log X_i$ are all ≤ 0, and so $\log X_i$ is negative semi-definite for every i. Therefore, since X_i is positive definite and diagonalisable in the same basis as $\log X_i$, we have that for all $\alpha > 0$ and all i, $X_i^\alpha \log X_i$ is negative semi-definite. This implies that for every $|\psi\rangle$,

$$\sum_i \langle\psi|X_i^\alpha \cdot \log X_i|\psi\rangle \leq 0,$$

and so

$$\frac{\partial}{\partial\alpha} N_\alpha^{|\psi\rangle}(\mathbf{X}) \leq 0.$$

Now let $\alpha_0 \geq \alpha_1$, and $|\psi\rangle \in \mathcal{S}(\mathcal{H})$ an arbitrary state, then

$$N_{\alpha_0}^{|\psi\rangle}(\mathbf{X}) \geq N_{\alpha_1}^{|\psi\rangle}(\mathbf{X}).$$

Now since $|\psi\rangle$ is arbitrary we have in particular that,

$$\min_{|\psi\rangle} N_{\alpha_0}^{|\psi\rangle}(\mathbf{X}) \geq \min_{|\psi\rangle} N_{\alpha_1}^{|\psi\rangle},$$

so $N_\alpha(\mathbf{X})$ is decreasing in $(0,1)$. Finally take $|\tilde{\psi}\rangle$ be such that $N_{\alpha_1}(\mathbf{X}) = N_{\alpha_1}^{|\tilde{\psi}\rangle}(\mathbf{X})$, then

$$N_{\alpha_0}(\mathbf{X}) \geq N_{\alpha_0}^{|\tilde{\psi}\rangle}(\mathbf{X}) \geq N_{\alpha_1}^{|\tilde{\psi}\rangle}(\mathbf{X}) = N_{\alpha_1}(\mathbf{X}),$$

which proves that N_α is decreasing in $(1,\infty)$. □

We know by (4.46) that in the case of POVMs with simultaneously diagonalisable operators this function is continuous at $\alpha = 1$. It would be interesting then, to see if this continuity also holds for a general POVM.
Also, we know that the α-Rényi norm is monotone non-decreasing for $\alpha > 1$, and N_α is monotone non-increasing for $\alpha > 1$, we might ask ourselves what is the gap in the limit inequality

$$\|\mathbf{X}\|_\infty \leq \lim_{\alpha\to\infty} N_\alpha(\mathbf{X}). \tag{4.47}$$

Again for POVMs with simultaneously diagonalisable operators, or for POVMs expressible as dilations of projections, the above is an equality. Does this hold for general POVMs?
Finally, another interesting question would be to determine when is $N_\alpha(\mathbf{X})$ constant in α.

4.5 The space of distributions of a measurement

The study of the space of distributions induced by a POVM opens a door for a geometric approach to analyse the uncertainty of these measurements. For example the interplay between the convex hulls that these spaces define, can give us conditions for the entropy of one measurement dominating that of the other.
We define the space of distributions as follows,

Definition 25. For a POVM $\mathbf{X} = \{X_i\}_{i=1}^m$ on a d-dimensional Hilbert space, the space of distributions of $\mathbf{X}$ is defined to be

$$\Delta_{\mathbf{X}} = \{(\langle\psi|X_{\sigma(i)}|\psi\rangle)_i \in \mathbb{R}^m : |\psi\rangle \in \mathcal{S}(\mathcal{H}),\ \sigma \in S_m\}. \tag{4.48}$$

We can derive some simple geometric properties of $\Delta_{\mathbf{X}}$. Notice that $\Delta_{\mathbf{X}} \subseteq \Delta_{m-1}$ the $m-1$ dimensional simplex in $\mathbb{R}^m$. Since the maps

$$\begin{aligned} \mathcal{S}(\mathcal{H}) &\to \mathbb{R} \\ |\psi\rangle &\mapsto \langle\psi|X_i|\psi\rangle\,, \end{aligned}$$

are continuous for every i, we have that $\Delta_{\mathbf{X}}$ is compact and connected in $\mathbb{R}^m$. We can also see that $\Delta_{\mathbf{X}}$ is invariant under the action of S_m in $\mathbb{R}^m$ given by the permutations of coordinates. We can analyse the rank of the matrix whose rows are the vectors in $\langle x\rangle = \{\sigma(x) : \sigma \in S_m\}$ for any given $x \in \Delta_{\mathbf{X}}$. We know that since $\sum_i x_i = 1$, either this rank is 1, which corresponds to the case $x_i = \frac{1}{m}$ for every i, or the rank is m. Therefore one concludes that either $\Delta_{\mathbf{X}}$ consists of a single point or the points of $\Delta_{\mathbf{X}}$ are in general position inside Δ_{m-1}.

Example 3. For the POVM $p\mathbf{I} = (p_1, \ldots, p_m)\mathbf{I}$ we have that

$$\Delta_{p\mathbf{I}} = \{\sigma(p) : \sigma \in S_m\}, \tag{4.49}$$

where $\sigma(p)_i = p_{\sigma(i)}$ for all $i = 1, \ldots, m$. This is because for every $|\psi\rangle \in \mathcal{S}(\mathcal{H})$ and every $i = 1, \ldots, m$,

$$\langle\psi|p_i\mathbb{I}|\psi\rangle = p_i \langle\psi|\psi\rangle = p_i.$$

In particular, for the POVM $\frac{1}{m}\mathbf{I} = (\frac{1}{m}, \ldots, \frac{1}{m})\mathbf{I}$ we have

$$\Delta_{\frac{1}{m}\mathbf{I}} = \{(\frac{1}{m}, \ldots, \frac{1}{m})\}. \tag{4.50}$$

Therefore, the distribution space of $\frac{1}{m}\mathbf{I}$ corresponds to one single point.

We can obtain in general more complex distribution spaces. The following image corresponds to a projection in $\mathbb{R}^3$ of the space of distributions in our example from Chapter 3.

We can compare the spaces $\Delta_{\mathbf{X}}$ and $\Delta_{\mathbf{Y}}$ of two POVMs by assuming that both have the same number of operators, we can always do this by enlarging one of the POVMs by adding $\mathbf{0}$ operators. It is clear that the resulting space of distributions will just include permutations of the original space of distributions. In this way we can also consider $\Delta_{\mathbf{X}}$ in a bigger simplex $\Delta_{m'}$ with $m' > m$ as

$$\Delta_{\mathbf{X}} = \{\sigma(\langle\psi|X_1|\psi\rangle, \ldots, \langle\psi|X_m|\psi\rangle\,; 0, \ldots, 0) : |\psi\rangle \in \mathcal{S}(\mathcal{H}),\ \sigma \in S_{m'}\}. \tag{4.51}$$

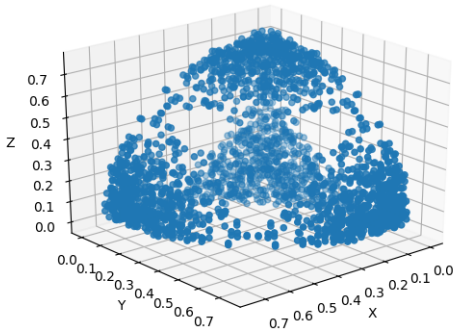


Figure 4.1: Randomly generated points in $\Delta_{\mathbf{M}}$, where $\mathbf{M} = \{M_{ab}\}_{a,b\in\{0,1\}}$.

By the property of majorisation being preserved under addition of zero coordinates, we can consider all spaces of distributions of POVMs in a big enough simplex $\Delta_{m-1} \subseteq \mathbb{R}^m$.
A way of relating entropy to the points of $\Delta_{\mathbf{X}}$, is by considering majorisation of the distributions in this space. Geometrically, majorisation $p \prec q$ of two distributions p and q means that the convex hull of the coordinate permutations of q contains that of p, [3].
It will be interesting then to consider the convex hull $\mathrm{Conv}(\Delta_{\mathbf{X}})$ of $\Delta_{\mathbf{X}}$.

Remark 3. For every $p \in \Delta_{\mathbf{X}}$, we define

$$\mathrm{Conv}_{S_m}(p) := \mathrm{Conv}\{\sigma(p) : \sigma \in S_m\}, \tag{4.52}$$

to be the convex hull of the vectors of permutations of coordinates of p.
Then with this terminology we have that $\mathrm{Conv}(\Delta_{\mathbf{X}}) \supseteq \mathrm{Conv}_{S_m}(p)$ for all $p \in \Delta_{\mathbf{X}}$. This comes from the fact that $\{\sigma(p) : \sigma \in S_m\} \subseteq \Delta_{\mathbf{X}}$, therefore $\mathrm{Conv}(\Delta_{\mathbf{X}}) \supseteq \bigcup_{p\in\Delta_{\mathbf{X}}} \mathrm{Conv}_{S_m}(p)$. Therefore

$$\mathrm{Conv}(\Delta_{\mathbf{X}}) = \mathrm{Conv}\left(\bigcup_{p\in\Delta_{\mathbf{X}}} \mathrm{Conv}_{S_m}(p)\right). \tag{4.53}$$

To have an analogous of Schur-concavity for $H_\alpha(\mathbf{X})$, the object we will be interested in analysing is $\bigcup_{p\in\Delta_{\mathbf{X}}} \mathrm{Conv}_{S_m}(p)$.
With this in mind we can formulate the following theorem, which gives us a sufficient condition for a POVM to dominate another in terms of the α-Rényi entropy for every $\alpha \geq 0$,

Theorem 7. *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be POVMs on* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *respectively, and suppose that in some big enough simplex* Δ_m*, the inclusion* $\bigcup_{q\in\Delta_{\mathbf{X}}} Conv_{S_m}(q) \supseteq \bigcup_{p\in\Delta_{\mathbf{Y}}} Conv_{S_m}(p)$ *holds. Then* $H_\alpha(\mathbf{Y}) \geq H_\alpha(\mathbf{X})$ *for all* $\alpha \geq 0$.

Proof. Let $|\tilde{\psi}\rangle \in \mathcal{S}(\mathcal{H}_2)$ such that

$$H_\alpha(\mathbf{Y}) = H_\alpha(\mathbf{Y}|\tilde{\psi}).$$

Let $\tilde{p} = (\langle\tilde{\psi}|Y_1|\tilde{\psi}\rangle, \ldots, \langle\tilde{\psi}|Y_m|\tilde{\psi}\rangle)$, i.e. $\tilde{p}$ is the distribution associated to $|\tilde{\psi}\rangle$ in $\Delta_{\mathbf{Y}} \subseteq \Delta_m$. Then we have that $\tilde{p} \in \Delta_{\mathbf{Y}} \subseteq \bigcup_{p\in\Delta_{\mathbf{Y}}} \mathrm{Conv}_{S_m}(p) = \bigcup_{q\in\Delta_{\mathbf{X}}} \mathrm{Conv}_{S_m}(q)$. Therefore there exists a $q \in \Delta_{\mathbf{X}}$ such that $\tilde{p} \in \mathrm{Conv}_{S_m}(q)$, which is equivalent to $\tilde{p} \prec q$. Now by the Schur-concavity of H_α we have that

$$H_\alpha(\mathbf{Y}) = H_\alpha(\tilde{p}) \geq H_\alpha(q) \geq H_\alpha(\mathbf{X}), \tag{4.54}$$

where the last inequality comes from the fact that $q \in \Delta_{\mathbf{X}}$. □

In fact the condition $\bigcup_{q\in\Delta_{\mathbf{X}}} \mathrm{Conv}_{S_m}(q) \supseteq \bigcup_{p\in\Delta_{\mathbf{Y}}} \mathrm{Conv}_{S_m}(p)$ defines a preorder relation on the set of POVMs.

Definition 26. Let $\mathbf{X}$ and $\mathbf{Y}$ be two POVMs. Then we write $\mathbf{Y} \prec \mathbf{X}$ if and only if

$$\bigcup_{q\in\Delta_{\mathbf{X}}} \mathrm{Conv}_{S_m}(q) \supseteq \bigcup_{p\in\Delta_{\mathbf{Y}}} \mathrm{Conv}_{S_m}(p). \tag{4.55}$$

Then we can restate Theorem 7 as a "concavity property" of the entropy of a POVM, namely if $\mathbf{Y} \prec \mathbf{X}$ then for every $\alpha \geq 0$,

$$H_\alpha(\mathbf{Y}) \geq H_\alpha(\mathbf{X}). \tag{4.56}$$

One can also ask the question if the converse of Theorem 7 holds, that is if $H_\alpha(\mathbf{Y}) \geq H_\alpha(\mathbf{X})$ holds for every $\alpha \geq 0$, then do we have $\mathbf{Y} \prec \mathbf{X}$?

We can use these notions to give a simpler proof of the claim that, for every POVM $\mathbf{X}$, one has $H_\alpha(\mathbf{X} \oplus \frac{1}{m}\mathbf{I}) = H_\alpha(\mathbf{X})$ for every $\alpha > 0$. This follows from the fact that the distribution $(1/m, \ldots, 1/m)$ is majorised by any other distribution and thus $\frac{1}{m}\mathbf{I} \prec \mathbf{X}$ for any POVM $\mathbf{X}$.

Therefore, in our example (3.8) we can always consider that the entropy is concentrated in the M_{ab} part.

Another interesting observation is that if there exists a $p \in \Delta_{\mathbf{X}}$ such that $\mathrm{Conv}_{S_m}(p) \supseteq \mathrm{Conv}(\Delta_{\mathbf{X}})$, then we can express all sharp α-Rényi entropy uncertainty relations as the entropy of p, namely

$$H_\alpha(\mathbf{X}) = H_\alpha(p). \tag{4.57}$$

This is because $q \in \mathrm{Conv}_{S_m}(p)$ for all $q \in \Delta_{\mathbf{X}}$, and so

$$H_\alpha(\mathbf{X}|\psi) \geq H_\alpha(p).$$

Now since $p \in \Delta_{\mathbf{X}}$ we get the equality (4.57). It would be interesting then to investigate under which conditions on the POVMS a situation like this can happen. For instance in our example (3.8) the distributions that minimise the min-entropy and the collision entropy coincide, this suggests that this distribution might dominate the other induced distributions by the POVM.

Finally another question one might ask is to find which kind of $\mathrm{Conv}(\Delta_{\mathbf{X}})$ can arise from a POVM, a characterisation might provide interesting insight to obtain uncertainty relations.

Conclusions and Future Work

- There seems to be a potential application of the RRDPS encoding to a 1-2 OT protocol. Whether we can devise a secure OT protocol from it, possibly in the Bounded-Quantum-Storage model, remains an open question.
- The quantity $H_\alpha(\mathbf{X})$ has a good behaviour under common operations that arise when studying measurements.
- From the study of the behaviour of the function $N_\alpha(\mathbf{X})$ on simultaneously diagonalisable POVMs one can see that the degree of commutativity within operators of a POVM can provide information about the uncertainty of said POVM.
- It remains an open question to see if $N_\alpha(\mathbf{X})$ is a norm, if it is continuous at $\alpha = 1$ and $\alpha = \infty$, and to find better bounds for $H_\alpha(\mathbf{X})$ by taking into account a degree of noncommutativity. It would be interesting then, to determine which expression carrying information about the degree of noncommutativity within the measurements is more suitable for the analysis of POVMs.
  Also determining when N_α is continuous can give us information on when Theorem 5 fails to give new information.
- The study of the space of distributions enables us to extract information about the uncertainty of a measurement by geometric means. A deeper analysis in this direction would be interesting for future study. Our results on the dominance of the entropy of one measurement over another measurement, remain true when considering any entropy $\mathscr{E}$ that is Schur-concave and has a global maximum at the uniform distribution. Consider for a general entropy the quantities,

  $$\mathscr{E}(\mathbf{X}|\psi) = \mathscr{E}(p_{|\psi\rangle}), \tag{4.58}$$

  for every state $|\psi\rangle$, and

  $$\mathscr{E}(\mathbf{X}) = \min_{|\psi\rangle} \mathscr{E}(\mathbf{X}|\psi). \tag{4.59}$$

  It remains an open problem then, to see if $H_\alpha(\mathbf{Y}) \geq H_\alpha(\mathbf{X})$ for every $\alpha > 0$ implies that $\mathbf{Y} \prec \mathbf{X}$. More generally we can ask ourselves, if the inequality $\mathscr{E}(\mathbf{Y}) \geq \mathscr{E}(\mathbf{X})$, holding for every entropy $\mathscr{E}$, implies that $\mathbf{Y} \prec \mathbf{X}$.
- Along the line of this last open problem, we can also consider the question of how can we generalise the results we obtained to general entropies, and if there exist generalisations to the norms we defined.

Bibliography

[1] T. Ando. *Matrix Young Inequalities*, pages 33–38. Birkhäuser Basel, Basel, 1995.

[2] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. In *International Conference on Computers, Systems & Signal Processing*.

[3] Rajendra Bhatia. *Matrix Analysis*. Springer New York, New York, NY, 1997.

[4] Iwo Białynicki-Birula and Jerzy Mycielski. Uncertainty relations for information entropy in wave mechanics. *Communications in Mathematical Physics*, 44(2):129–132, Jun 1975.

[5] Patrick J Coles, Mario Berta, Marco Tomamichel, and Stephanie Wehner. Entropic uncertainty relations and their applications. *Reviews of Modern Physics*, 89(1):015002, 2017.

[6] Shannon C. E. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

[7] D Fadeev. Zum begriff der Entropie einer endlichen Wahrscheinlichkeitschemes. arbeit zur Informationstheorie, 1957.

[8] Godfrey Harold Hardy, John Edensor Littlewood, and George Pólya. *Inequalities*. Cambridge university press, 1952.

[9] W. Heisenberg. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, 43(3):172–198, Mar 1927.

[10] Alexander Semenovich Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, 9(3):3–11, 1973.

[11] Earle H. Kennard. Zur Quantenmechanik einfacher Bewegungstypen. *Zeitschrift für Physik*, 44(4):326–352, Apr 1927.

[12] M. Krishna and K. R. Parthasarathy. An entropic uncertainty principle for quantum measurements. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 64(3):842–851, 2002.

[13] Hans Maassen and J. B. M. Uffink. Generalized entropic uncertainty relations. *Phys. Rev. Lett.*, 60:1103–1106, Mar 1988.

[14] Albert W Marshall, Ingram Olkin, and Barry C Arnold. Inequalities: theory of majorization and its applications. *New York*, 1979.

[15] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[16] Alfréd Rényi. On measures of entropy and information. Technical report, Hungarian Academy of Sciences Budapest Hungary, 1961.

[17] Howard P. Robertson. The uncertainty principle. *Phys. Rev.*, 34:163–164, Jul 1929.

[18] Bryan P. Rynne and Martin A. Youngson. *Linear functional analysis*. Springer Science & Business Media, 2000.

[19] Toshihiko Sasaki, Yoshihisa Yamamoto, and Masato Koashi. Practical quantum key distribution protocol without monitoring signal disturbance. *Nature*, 509(7501):475, 2014.