ENTROPIC UNCERTAINTIES IN QUANTUM MEASUREMENTS

by
BİLAL CANTÜRK

Submitted to the Graduate School of Engineering and Natural Sciences
in partial fulfilment of
the requirements for the degree of Doctor of Philosophy

Sabancı University
December 2021

ENTROPIC UNCERTAINTIES IN QUANTUM MEASUREMENTS

Approved by:

Date of Approval: December 16, 2021

ABSTRACT

ENTROPIC UNCERTAINTIES IN QUANTUM MEASUREMENTS

BİLAL CANTÜRK

PHYSICS Ph.D. DISSERTATION, DECEMBER 2021

Dissertation Supervisor: Prof. Mehmet Zafer Gedik

Keywords: Quantum measurement, Uncertainty principle, Entropic uncertainties, Time reversal, Information energy

We have studied entropic uncertainty relation for two types of quantum measurements in quantum information theory. One of them is the projective measurements that are constructed from the mutually unbiased bases and the other one is the symmetric informationally complete positive operator-valued measure. We present an optimal upper bound of entropic uncertainty relation for these two types of measurements. We have obtained a criterion for the extendibility of mutually unbiased bases in terms of Shannon entropy by means of the optimal upper bound of entropic uncertainty relation. We study time reversal operation for the latter type of measurement. We reveal that the notions of time reversal in quantum mechanics and in quantum operation formalism are not compatible with each other. We propose a harmonization of the notions, according to which symmetric informationally complete positive operator-valued measure is time reversal invariant. We also study on the algebraic relation between the two measurements; we provide an algebraic relation by which an analytical search of the existence of mutually unbiased bases could be studied in six-dimensional Hilbert space. Finally, a physical ground of the use of information energy in quantum information theory has been provided with recourse to Stokes parameters.

ÖZET

KUANTUM ÖLÇÜMLERİNDE ENTROPİSEL BELİRSİZLİKLER

BİLAL CANTÜRK

FİZİK DOKTORA TEZİ, ARALIK 2021

Tez Danışmanı: Prof. Dr. Mehmet Zafer Gedik

Anahtar Kelimeler: Kuantum Ölçümü, Belirsizlik ilkesi, Entropisel belirsizlikler, Zaman tersinirliği, Enformasyon enerjisi

Bu tez kapsamında, iki tür kuantum ölçümü olan karşılıklı adil tabanlardan oluşturulan izdüşümsel ölçümler ve simetrik tam enformasyonlu pozitif operatör değerli ölçümler için entropisel belirsizlikler çalışılmıştır. Bu iki tür ölçüm için entropisel belirsizliğin en küçük üst sınırı elde edilmiş; bu üst sınır kullanılarak karşılıklı adil tabanların sayısının genişletilebilirliği için Shannon entropisi cinsinden bir ölçüt ortaya koyulmuştur. Buna ek olarak, simetrik tam enformasyonlu pozitif operatör değerli ölçümler üzerinden kuantum mekaniğindeki ve kuantum enformasyon teorisindeki zaman tersinirliği kavramı çalışılmış; kavramın bu iki alandaki kullanımının uyumlu olmadığı tespit edilmiştir. Bunun sonucu olarak, zaman tersinirliğinin bu iki alandaki kullanımını örtüştüren bir öneri sunulmuştur. Bu öneriye göre, adı geçen ölçümün, zaman tersinirliği altında değişmez olduğu gösterilmiştir. Ayrıca, söz konusu iki ölçüm türü arasındaki cebirsel ilişkiler çalışılmış; çalışmalar sonucunda, altı boyutlu Hilbert uzayında karşılıklı adil tabanların varlığını analitik olarak çalışılmasını sağlayacak cebirsel bir ilişki bulunmuştur. Son olarak, enformasyon enerjisinin kuantum enformasyon teorisindeki kullanımına, Stokes parametreleri vasıtasıyla fiziksel bir içerik sağlanmıştır.

ACKNOWLEDGEMENTS

I would like to start with expressing my gratitude to my supervisor Prof. Zafer Gedik for his guidance and support. His knowledge, discipline and stimulating questions inspired me to go deeper and deeper about my questions and studies. I am heartily indebted and thankful to him for putting his faith in me.

Prof. Emrah Kalemci and Assoc. Prof. Emre Erdem encouraged me in my work by their questions and help during my thesis progress that clarified and enriched my studies. Prof. Emrah Kalemci helped me to correct typing errors in the thesis and to make the thesis more understandable. I would like to express my appreciation to them.

Another person who had an impact on me is Prof. Barış Bağcı. He is not only an advisor or academician for me, but also a friend with whom I can get in conversation friendly. He always encouraged and guided me in my work and life whenever I fell into perplexity. His knowledge on entropies and fundamental concepts of physics broadened my vision and enriched my understanding. For that, I am deeply thankful to him.

I am always inclined to write the reasons and the results of an idea in one sentence that is inherent in my thinking. This naturally gives rise to long sentences, so thus, ambiguity in my writing for readers. Assoc. Prof. Levent Subaşı has made many suggestions to me about this issue and about some concepts that I have used throughout the thesis. They were helpful to clarify the thesis. I kindly thank to him.

I would like to keenly express my gratitude to my friends, at the same time my colleagues, Ali Mürteza Altıngün and Solomon Brhanu Samuel. We together had lots of useful and suggestive discussions on diverse questions about my work and the foundations of physics and other related fields.

I acknowledge the support and full scholarship from Sabanci University.

Bu çalışmayı, varlığı daima benim ruh dünyam için bir güç ve ilham kaynağı olan güzel insan, anneannem Saadet Sönmez'e ithaf ediyorum.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATONS

1. INTRODUCTION

Heisenberg uncertainty principle is considered as one of the cornerstones of quantum mechanics. At the beginning, It was proposed by Heisenberg as the physical content of the commutation relation between position and momentum variables. Position X and momentum P are considered two observables whose commutation relation is $[X, P] = \mathrm{i}\hbar I$, that is, position and momentum do not commute, similar to two matrices not commuting. In matrix formalism, two non-commutative matrices do not have common eigenvectors. Heisenberg interpreted this fact physically as the impossibility of the simultaneous determination of the observables. For Heisenberg, determining a physical observable means to have, at least in principle, an experimental setup by which we can perform the measurement of the observable. Therefore, for Heisenberg, it is impossible to set up an experiment by which measurements of non-commutative observable is performed with a desired accuracy. Heisenberg himself expressed uncertainty principle for position and momentum in terms of their deviation as $\Delta(X)\Delta(P) \sim h$, where h is the Planck constant (Heisenberg, 1927). Later, Robertson formulated it mathematically for any two observables A and B as $\Delta(A)\Delta(B) \geq \frac{1}{2}|\langle\psi|[A, B]|\psi\rangle|$ for the quantum state $|\psi\rangle$ of the system of inquiry (Robertson, 1929). In the course of time, the conceptual and mathematical critics of these formulation gave rise to many formulations of uncertainty (Deutsch, 1983; Ozawa, 2003; Bush et al., 2013). One of these formulations is entropic uncertainty relation. To express it formally, we try to formulate the incompatibility of two observables A and B by means of an entropy function H. In other words, we seek for an inequality $H(A) + H(B) \geq C > 0$ (Deutsch, 1983). The lower bound C becomes zero if the observables are compatible, that is, if they do commute.

Entropic uncertainty relation has gained a central position in quantum information theory due to their useful role in many issues such as detecting entanglement (Spengler et al., 2012; Wang & Zheng, 2021) and quantum cryptography (Mafu et al., 2013). On the other hand, determining an unknown quantum state so-called quantum state tomography is one of the fundamental tasks in quantum information theory. The quantification, storage, and communication of information require using resources such as devices, energy and limited time. Therefore, when determining the quantum state, we always need to use the resources minimally and wish to determine the quantity under quest as best as possible in the sense that the statistical error in the determination must be minimum. In line with this task, there are two important measurement types in quantum information theory. First one is a set of $d+1$ projective measurements that are constructed from a set of $d+1$ *Mutually Unbiased Bases* (MUBs) (Wootters & Fields, 1989) in d-dimensional Hilbert space. The operators corresponding to MUBs are the generalization of complementary observables such as position and momentum in every possible dimension. For example, the eigenbases of Pauli operators are such bases in 2-dimensional Hilbert space. The other type of measurement is the *Symmetric Informationally Complete Positive Operator-Valued Measure* (SIC-POVM) in d-dimensional Hilbert space, which consists of d^2 measurement elements.

The existence of $d+1$ MUBs is known if the dimension d is a power of a prime number, while for a composite dimension such as six, their existence is still an open problem. We also know by construction that there are at least 3 MUBs in every dimension. Therefore, as the first step to the proof of the existence of $(d+1)$ MUBs, the existence of fourth MUB is a challenge. On the other hand, the existence of SIC-POVMs in every dimension was conjectured by Zauner (Zauner, 2011); however, despite many solutions in many dimensions as high as 844, their existence in every dimension has not been proven yet (Appleby & Bengtsson, 2019).

These two types of quantum measurements are important not only because of their use for quantum state tomography but also for their many applications such as in quantum cryptography (Spengler et al., 2012; Renes, 2005), quantum key distribution (Cerf et al., 2002; Bouchard et al., 2018), quantum channels and foundations of quantum theory (Durt et al., 2010; Fuchs et al., 2017).

In the context of this thesis, we have studied entropic uncertainty relations for MUBs and SIC-POVMs based on Shannon, Rényi and Tsallis entropies.

In Chapter 2, we have first summarized the notion of quantum operation followed by a presentation of MUBs and SIC-POVMs with some examples. Subsequently, we have furnished an outlook of aforementioned entropies that are common in literature together with the requirements that are propounded for the accessibility of any proposed entropy.

In Chapter 3, we have first presented deviation-based uncertainty relations and a critics of them based on some concrete examples. Afterwards, in Section 3.2, we have explored entropic uncertainty relations for continuous observables such as position and momentum, and for the observables in finite dimensions such as Pauli observables.

The rest of the thesis is an exploration of our work on several subjects that we studied. In the first place, we studied optimal upper bound of entropic uncertainty relations for MUBs and SIC-POVMs in terms of Shannon entropy. Accordingly, Section 3.3 is based on our work (Canturk & Gedik, 2021), in which we have found an optimal upper bound of entropic uncertainty relation for MUBs and SIC-POVMs. It enables us to provide a criterion for the existence of MUBs. As the second subject we studied, Section 3.4 consists of our work on time reversal operation in quantum operation formalism in the context of SIC-POVM. We revealed that the notions of time reversal in quantum mechanics and in quantum operation formalism are not compatible with each other. We have proposed a harmonization of them. Thirdly, we studied algebraic relation between MUBs and SIC-POVMs; we found an algebraic relation between them by means of which it is possible to search the existence of MUBs analytically. We explored our results in Section 3.5. Using information energy as the measure of the information content of a quantum system was firstly proposed by Brukner and Zeilinger (Brukner & Zeilinger, 1999). However, the physical motivation for the use of information energy given by Brukner and Zeilinger is not satisfactory. Regarding to this problem, in Section 3.6, we studied information energy, and provided a physical motivation for the use of it in quantum information theory with recourse to Stokes parameters. Finally, in Conclusion, we have summarized our results and their possible consequences.

We have used the word "relations" instead of "principle" when expressing entropic uncertainties for two reasons. First of all, the word "principle" for us deserves a rather indisputable character for its content and meaning. However, there has not been any agreement on the content and meaning of "uncertainty principle" since it was proposed firstly by Heisenberg. Due to this situation, many alternative expressions of "uncertainty principle" were proposed in the course of time (Schrödinger, 1930; Ozawa, 2003; Bush et al., 2013). Indeed, if it was a principle that was immune to any debate about its content and expression, all of these attempts would not be proposed as the best candidate for the expression of the uncertainty principle. Secondly, not only one but many entropies are used for expressing uncertainty; that is why we have used the word "relation" in the plural form throughout the thesis.

2. QUANTUM MEASUREMENTS AND ENTROPY

In this chapter, we shall give a concise outline of the quantum measurement and entropy that are two fundamental concepts underpinning the content of the thesis. In the following section, we first present the quantum measurement concept in the context of quantum information theory and then, MUBs and SIC-POVMs. Afterwards, we present the concept of entropy from its origin to its current status and usage. Our goal is not to dive into a discussion of these two concepts, but rather to draw a clear frame of them, which is going to be helpful to follow the thesis.

2.1 Quantum measurements

In quantum mechanics, we observe physical quantities that are represented mathematically by hermitian matrices. Hermitian matrices are also called hermitian operators in quantum mechanics and quantum information theory. Every physical quantity that can be observed is called observable in general. Therefore, we shall use the terms "operators" and "observables" interchangeably throughout the thesis. Observation means measuring physical quantities on an ensemble of the system under consideration. If the state of the system is ρ and the hermitian operators corresponding to the physical quantity to be measured is A, the expectation value of the physical quantity, after performing many measurements on the elements of the ensemble, is expressed as the trace of the multiplication of A and ρ: $\langle A \rangle = \mathrm{tr}(A\rho)$. A hermitian operator A in d-dimensional Hilbert space has spectral decomposition $A = \sum_k a_i \left| a_i \right\rangle \left\langle a_i \right|$ such that $\{a_i\}_{i=1}^d$ is the set of eigenvalues of A with the set of corresponding projection operators $\{\Pi_i = \left| a_i \right\rangle \left\langle a_i \right|\}_{i=1}^d$. A single measurement on an element of the ensemble results in an outcome, which is one of the eigenvalues of A, say a_i, with probability $p_i = p(a_i) = \mathrm{tr}(\Pi_i \rho)$, which is known as the *Born rule* for probabilities. Therefore, putting aside the measured physical quantities, an immediate implication of this examination is that the measurement phenomenon can be represented by a set of projections $\{\Pi_i\}_{i=1}^n$. From the relation $\sum_i p_i = 1 = \sum_i \mathrm{tr}(\Pi_i\rho)$ for all ρ, it has to be that $\sum_i \Pi_i = I_d$. This is a necessary condition for every measurement that is known as the *completeness condition of measurement*. A measurement of projections is called *projective measurement*, which is also known as *von Neumann measurement*. The following properties uniquely define a projective measurement $\{\Pi_i\}_{i=1}^n$:

i. They are hermitian, that is, they are self-adjoint: $\Pi_i^\dagger = \Pi_i$.

ii. They are positive operators: $\Pi_i \geq 0$ for all i.

iii. They are orthonormal: $\Pi_i\Pi_j = \Pi_i\delta_{ij}$

The first two properties have physical meaning. Projectors are hermitian with real eigenvalues since they represent physical, or say observable, quantities. They are positive because their expectation values are probabilities, which have to be non-negative. However, the third property does not have a convincing physical interpretation if the realistic conditions of performing a measurement are taken into account.

Projective measurements are ideal in the sense that any measurement device detects the quantum state of the system under observation perfectly and projects it ideally to the eigenspace of the eigenvalues correspondingly. To express formally, if the prior quantum state is ρ and the projective measurement $\mathcal{M} = \{\Pi_i\}_{i=1}^n$ is to be performed, the posterior quantum state in turn is set to be $\rho \mapsto \rho' = \frac{\Pi_i\rho\Pi_i}{\mathrm{tr}(\Pi_i\rho\Pi_i)}$ perfectly and with certainty, where the numerator ensures the normalization of the new quantum state. However, to be realistic and consider the measurement phenomenon in general, every measurement is not a projective measurement. For example, consider a double-slit experiment with photons. If the experimenter aims to detect which slit a photon passes through, they simply try to detect the photon such that the observed photon is absorbed by the detector, and after that, it is meaningless to talk about the posterior quantum state of the photon. In the case of such measurements, the posterior quantum state is irrelevant but the outcomes and the probabilities of their occurrence are the issue at stake. To formalize this fact, one assumes in general a set of positive operators $\{F_i\}_{i=1}^m$ with probability expression of the corresponding outcomes as $p_i = \mathrm{tr}(F_i\rho)$ provided that the quantum state of the considered system is ρ. F_i's do not have to satisfy the third property of projective measurements. This is the general form of measurement, which is called *Positive Operator-Valued Measure* (POVM). Secondly, one could not claim that, after measurement, the posterior quantum state changes to $\rho \mapsto \rho' = \frac{\Pi_i\rho\Pi_i}{\mathrm{tr}(\Pi_i\rho)}$ *perfectly* and with *certainty* if the measurement device would not be perfectly efficient because of some structural defect. In other words, an imperfect device sometimes does not detect events even if the events indeed have been occurred.

The physical world is noisy and many side effects intertwine with our measurements, which together imply that our conclusions contain uncertainties. To give an example, let us assume that we have an imperfect photo-detector detecting an incoming photon with probability q. Our photo-detector have, in this scenario, limited photon number resolution. Additionally, we note that the photo-detector might click even if there is no photon. If the number of incoming photons are m in total, the probability $p(n \mid m)$ to detect n of the photons depends certainly on the photo-detector, which can be estimated as $p(n \mid m) = \binom{m}{n} q^n (1-q)^{m-n}$. If p_m is the probability provided that the incoming light has m photons, the overall probability to detect n photons is equal to $p(n) = \sum_{m=n}^{\infty} p(n \mid m) p_m = \sum_{m=n}^{\infty} \binom{m}{n} q^n (1-q)^{m-n} p_m$. Now, if we wish to write the detection probabilities in the form $\mathrm{tr}(\Pi_n \rho)$, where ρ is to be the quantum state of the incoming light, and Π_n is the measurement operator (or element) corresponding to n photon detection, then the measurement elements $\{\Pi_n = \sum_{m=n}^{\infty} \binom{m}{n} q^n (1-q)^{m-n} |m\rangle \langle m|\}_{n=0}^{\infty}$ do the job plausibly, which in contrast to projective measurements are not orthogonal projectors and rank-1 in general. In addition, they are complete, that is, they sum to the identity matrix.

A source of noise also leads to a non-projective measurement. For instance, let us assume that we have a device for determining whether a qubit is in the state $|0\rangle$ or the state $|1\rangle$ such that it detects the state wrongly with probability q due to the noise. If there were no noise, the device would act ideally and the measurement elements would be $\Pi_0 = |0\rangle \langle 0|$ and $\Pi_1 = |1\rangle \langle 1|$ respectively. Accordingly, if the prior quantum state is $|0\rangle$, the device detects it correctly with probability $1-q$ and as if it is $|1\rangle$ with probability q. For a quantum state ρ, the probabilities for each of two measurement outcomes are $p(0) = (1-q)\,\mathrm{tr}(\Pi_0 \rho) + q\,\mathrm{tr}(\Pi_1 \rho)$ and $p(1) = (1-q)\,\mathrm{tr}(\Pi_1 \rho) + q\,\mathrm{tr}(\Pi_i \rho)$, from which one can express the measurement elements as $\pi_0 = (1-q)\Pi_0 + q\Pi_1$ and $\pi_1 = q\Pi_0 + (1-q)\Pi_1$ respectively so that $p(0) = \mathrm{tr}(\pi_0 \rho)$ and $p(1) = \mathrm{tr}(\pi_1 \rho)$. As can be seen easily, these measurement elements do not satisfy the third property of projective measurements, i.e., $\pi_0 \pi_1 = q(1-q)I$. This is another example of POVMs.

We conclude that every projective measurement is a POVM but the converse is not true, as was just clarified by the given examples. We also note that any POVM $\{F_i\}_{i=1}^{N \geq d}$ in a Hilbert space of d-dimension can be realized as a projective measurement in an extended N-dimensional Hilbert space (Barnett, 2009, pp. 92-97).

Given two Hilbert spaces $\mathcal{H}^{d_1}$ and $\mathcal{H}^{d_2}$ of d_1 and d_2 dimensions, we take $\mathcal{L}(\mathcal{H}^{d_1}, \mathcal{H}^{d_2})$ as the set of all linear transformations of the form $A : \mathcal{H}^{d_1} \to \mathcal{H}^{d_2}$ throughout the thesis. If $d_1 = d_2 = d$, we simply write $\mathcal{L}(\mathcal{H}^d)$ instead of $\mathcal{L}(\mathcal{H}^d, \mathcal{H}^d)$. The set of all quantum states, $\mathcal{D}(\mathcal{H}^d)$, on Hilbert space $\mathcal{H}^d$ is a subset of $\mathcal{L}(\mathcal{H}^d)$, and is a convex set. This implies that quantum theory can be regarded purely as a statistical theory from a mathematical perspective. Furthermore, a useful fact is that the space of linear operators $\mathcal{L}(\mathcal{H}^d)$ on d-dimensional Hilbert space $\mathcal{H}^d$ can be spanned by a set of quantum states. An immediate implication of this fact is that every linear transformation having the form $T : \mathcal{L}(\mathcal{H}_d) \to \mathbb{C}$ can be characterized uniquely by its action on the elements of a basis of quantum states. For example, trace operation is such a linear transformation. A general quantum state ρ in a d-dimensional Hilbert space consists of d^2 real parameters if the normalization condition is ignored. This fact enables one to write a basis of d^2 elements for the set of all quantum states, which is also a basis for the vector space of hermitian operators on the real numbers. For any two Hilbert spaces $\mathcal{H}^{d_1}$ and $\mathcal{H}^{d_2}$ and for each choice of symbols $i \in \Gamma = \{1, 2, \ldots, d_1\}$ and $j \in \Lambda = \{1, 2, \ldots, d_2\}$, the operator $E_{i,j} \in \mathcal{L}(\mathcal{H}^{d_1}, \mathcal{H}^{d_2})$ is defined as $E_{i,j} = |i\rangle\langle j|$, where $|i\rangle = (0, 0, \ldots, 1, 0, \ldots, 0)$ is the standard basis vector having number 1 in the i^{th} position for its entries. The collection $\{E_{i,j} : i \in \Gamma, j \in \Lambda\}$ forms a basis for $\mathcal{L}(\mathcal{H}^{d_1}, \mathcal{H}^{d_2})$, which is known as the *standard basis*. A basis for the set of hermitian operators, $Herm(\mathcal{H}^d) \subset \mathcal{L}(\mathcal{H}^d)$, then can be constructed from a set of quantum states, which are to be expressed in terms of the elements of standard basis as

$$\sigma_{ij} = \begin{cases} |i\rangle\langle i| & \text{if} \quad i = j \\ \frac{1}{2}(|i\rangle + |j\rangle)(\langle i| + \langle j|) & \text{if} \quad i < j \\ \frac{1}{2}(|i\rangle + \mathrm{i}\,|j\rangle)(\langle i| - \mathrm{i}\,\langle j|) & \text{if} \quad i > j. \end{cases} \tag{2.1}$$

Any quantum state can be expressed as a linear combination of such a basis on real numbers $\mathbb{R}$, that is, a quantum state ρ is written as $\rho = \sum_{i,j} \lambda_{ij}\sigma_{ij}$ with all λ_{ij} being real numbers. The coefficients λ_{ij} are strictly connected with the probabilities, $p_{ij} = \mathrm{tr}(\sigma_{ij}\rho)$, if σ_{ij}'s can be regarded as measurement elements of some measurement. This fact inspires us the following question: Can a quantum state be reconstructed by performing a measurement of some basis or a set of measurements that are known as *informationally complete measurements*? We will see later that symmetrical informationally complete positive operator-valued measure and a set of $d+1$ projective measurements that are constructed from a set of $d+1$ mutually unbiased bases are two types of such measurement.

In the scope of quantum information theory, the measurement phenomenon is just a special case of quantum operations. Let us present some useful definitions as a preliminary for stating the formal definition of a quantum operation. We consider a linear map $\Phi : \mathcal{L}(\mathcal{H}^{d_1}) \to \mathcal{L}(\mathcal{H}^{d_2})$. The map Φ is *positive* if it transforms every positive semi-definite operator to another positive semi-definite operator. To express it formally, let us assume that $Pos(\mathcal{H}^d) = \{P \in \mathcal{L}(\mathcal{H}^d) : P \geq 0\}$ be the set of all positive semi-definite operators, then Φ is a positive map if $\Phi(P) \in Pos(\mathcal{H}^{d_2})$ for all $P \in Pos(\mathcal{H}^{d_1})$. The map Φ is *completely positive* if the tensor product $\Phi \otimes I_{\mathcal{L}(\mathcal{H}^m)}$ is also positive, where $I_{\mathcal{L}(\mathcal{H}^m)}$ is the identity transformation on $\mathcal{L}(\mathcal{H}^m)$. The map Φ is *trace-preserving* if $\mathrm{tr}(\Phi(A)) = \mathrm{tr}(A)$ for all $A \in \mathcal{L}(\mathcal{H}^{d_1})$. A *quantum operation* then is a linear map $\Phi : \mathcal{L}(H^{d_1}) \to \mathcal{L}(H^{d_2})$ satisfying the following properties:

i. Φ is completely positive.

ii. Φ is trace-preserving.

If $d_1 \neq d_2$, the system that undergoes the quantum operation changes to another system whose Hilbert space is different in general. For example, let $\mathcal{H}^{d_1}$ and $\mathcal{H}^{d_2}$ be two Hilbert spaces, and $\sigma \in \mathcal{L}(\mathcal{H}^{d_2})$ be a fixed quantum state, and let us consider the map $\Phi : \mathcal{L}(H^{d_1}) \to \mathcal{L}(H^{d_2})$ that is defined as $\Phi(X) = \mathrm{tr}(X)\sigma$ for all $X \in \mathcal{H}^{d_1}$. Accordingly, Φ is a quantum operation which is called replacement channel: It effectively discards its inputs and replace them with the fixed quantum state σ. If we take σ equal to identity matrix I, Φ becomes completely depolarizing channel: It takes every quantum state into a completely mixed state in $\mathcal{L}(H^{d_2})$. In this example, we lose information about the initial state. In the cases $d_1 \neq d_2$, we either lose information or combine some information in general.

Complete positivity guarantees the positivity of the resultant quantum state because, for instance, the transpose operation that is positive could yield a state which is not positive if the initial quantum state of the system would be entangled with the state of the environment (Nielsen & Chuang, 2010, p. 369). Any quantum operation $\Phi : \mathcal{L}(\mathcal{H}^{d_1}) \to \mathcal{L}(\mathcal{H}^{d_2})$ can be expressed in terms of a collection of operators $\{A_k \in \mathcal{L}(\mathcal{H}^{d_1}, \mathcal{H}^{d_2})\}_{k=1}^m$ as $\Phi(X) = \sum_{k=1}^m A_k X A_k^\dagger$, which is known as *Kraus representation* and A_k's are called *Kraus operators* (Watrous, 2018, pp. 77-91). Due to the trace-preserving condition of quantum operations, it holds that $\sum_{k=1}^m A_k^\dagger A_k = I_{d_1}$, where I_{d_1} is the identity matrix of the Hilbert space $\mathcal{H}^{d_1}$. The set $\{F_k = A_k^\dagger A_k\}_{k=1}^m$ can be considered as a POVM. It is possible to give a physical interpretation to Kraus representation: the action of the quantum operation on the quantum state ρ is equivalent to randomly applying the transformation $\rho \mapsto \rho_k = \frac{A_k \rho A_k^\dagger}{\mathrm{tr}\left(A_k \rho A_k^\dagger\right)}$ with probability $p(k) = \mathrm{tr}\left(A_k \rho A_k^\dagger\right)$. In accordance with this interpretation, we have

$$\rho \mapsto \Phi(\rho) = \sum_{k=1}^m p(k)\rho_k = \sum_{k=1}^m A_k \rho A_k^\dagger. \tag{2.2}$$

Quantum operation formalism is fundamentally the generalization of any physically accessible operation, including time evolution of quantum systems and the measurement procedure. Quantum operation formalism given above describes the evolution of the system of inquiry without having to explicitly know the properties of the external agent interacting with the system; all that we need to know is encapsulated into the operators A_k's, which act on the system of inquiry alone. The most advantageous feature of quantum operation formalism is that it enables us to handle open quantum systems. Any external effect on the principle system can be regarded as a part of the closed composite system, which includes the principle system and the effect. We generally refer to the effects as the environment. Accordingly, the dynamics of the composite system can be expressed by a unitary operation. Let us assume that the composite system at the beginning has the quantum state ρ^{se} such that the letters s and e refer to the principle system and the environment respectively. Then, the dynamics, including any interaction occurring between the principle system and environment, can be determined by a suitable unitary evolution U after which the environment is discarded by taking partial trace over it:

$$\rho_i^{se} \mapsto \left(U\rho_i^{se}U^\dagger\right) \mapsto \mathrm{tr}_e\left(U\rho_i^{se}U^\dagger\right) = \Phi(\rho_i^s) = \sum_{k=1}^{m} A_k\rho_i^s A_k^\dagger = \rho_f^s, \tag{2.3}$$

where ρ_i^{se} is the initial state of the composite system, ρ_f^s the final state of the principle system, and $\{A_k = (I\otimes\langle k|)U\rho_i^{se}U^\dagger(|k\rangle\otimes I)\}_{k=1}^m$ are the Kraus operators. The approach to the dynamics of the principle system by invoking the composite system and a suitable unitary evolution as given above is called *natural representation.* If the principle system has a Hilbert space of d dimension, it is possible to model the environment as residing in a Hilbert space of no more than d^2 dimension even if the environment has infinite degrees of freedom.

To give an example for composite system approach, let us assume that the unitary operation on the composite system of a single qubit system and a single qubit environment be $U = |0\rangle\langle 0|\otimes I + |1\rangle\langle 1|\otimes X$, where X is the usual Pauli matrix, and let the initial state of the composite system be a product state as $\rho_i^{se} = \rho\otimes|0\rangle\langle 0|$. The unitary evolution then yields

$$\begin{aligned}\rho\otimes|0\rangle\langle 0| \mapsto U(\rho\otimes|0\rangle\langle 0|)U^\dagger = P_0\rho P_0\otimes|0\rangle\langle 0| + P_1\rho P_0\otimes X|0\rangle\langle 0|\\ + P_0\rho P_1\otimes|0\rangle\langle 0|X + P_1\rho P_1\otimes X|0\rangle\langle 0|X,\end{aligned} \tag{2.4}$$

where $P_0 = |0\rangle\langle 0|$ and $P_1 = |1\rangle\langle 1|$. Taking the partial trace over the environment gives the quantum operation on the system

$$\mathrm{tr}_e(U(\rho\otimes|0\rangle\langle 0|)U^\dagger) = P_0\rho P_0 + P_1\rho P_1 = \Phi(\rho), \tag{2.5}$$

in which case the Kraus operators are $A_1 = P_0$ and $A_2 = P_1$. This is the quantum operation formalism of controlled-NOT gate. As an alternative approach, especially when a continuous-time evolution of the effects in differential equation form comes to the word, *Lindblad master equation* is often used, which has the form for a quantum state ρ

$$\frac{d\rho}{dt} = -\frac{\mathrm{i}}{\hbar}[H,\rho] + \sum_j \left(2L_j\rho L_j - \{L_j^\dagger L_j,\rho\}\right), \tag{2.6}$$

where the form $\{x,y\}$ is equal to anti-commutator $xy+yx$, H is the hamiltonian of the system own, and L_j's are the *Lindblad operators* representing the interaction between the system and its environment. Lindblad master equation takes on the above form in order that the evolution is completely positive and trace-preserving. It is also assumed generally that the initial state of the composite system is a product state in this formalism. Furthermore, in order to derive Lindblad master equation for a process, one usually starts with the hamiltonian of the composite system from which one obtains Lindblad operators by employing Born and Markov approximations (Moy et al., 1999). Every Lindblad master equation can be put into Kraus representation.

We examine measurement phenomenon as a quantum operation in the sphere of quantum information theory. Given the unitary operation U of a composite system, we are able to write the Kraus representation for the dynamics of the principle system by having the expression $A_k = I\otimes\langle k|\left(U\rho^{se}U^\dagger\right) I\otimes|k\rangle$ for the Kraus operators. One can extend this result a step further by including the possibility that a joint measurement is performed on the composite system after the unitary evolution U, allowing an information gain about the quantum state of the principle system. We recall that information gain about an observable in quantum information theory is possible by a measurement of the observable on the system of inquiry. It turns out that the aforementioned extension gives rise to a non-trace-preserving quantum operation (Nielsen & Chuang, 2010, p. 363). Let us suppose that, at the beginning, the composite system has the separated state $\rho_i^{se} = \rho_i^s\otimes\sigma_i^e$. After the unitary interaction U, we perform a projective joint measurement $\mathcal{M} = \{\Pi_m\}_{m=1}^n$ on the composite system whose outcomes gives us information by which we determine the final state of the principle system. The final quantum state ρ_f^{se} of the composite system is given by

$$\rho_i^s\otimes\sigma_i^e \mapsto U(\rho_i^s\otimes\sigma_i^e)U^\dagger \mapsto \frac{\Pi_m U(\rho_i^s\otimes\sigma_i^e)U^\dagger\Pi_m}{\mathrm{tr}\left(\Pi_m U(\rho_i^s\otimes\sigma_i^e)U^\dagger\Pi_m\right)} = \rho_f^{se} \tag{2.7}$$

provided that the outcome m has been disclosed while performing the measurement. The final quantum state of the principle system is then obtained by simply discarding the environment from the final state of the composite system

$$\rho_f^{se} \mapsto \mathrm{tr}_e(\rho_f^{se}) = \frac{\mathrm{tr}_e\Big(\Pi_m U(\rho_i^s \otimes \sigma_i^e)U^\dagger \Pi_m\Big)}{\mathrm{tr}\Big(\Pi_m U(\rho_i^s \otimes \sigma_i^e)U^\dagger \Pi_m\Big)} = \rho_f^s. \tag{2.8}$$

By defining a quantum operation $\Phi_m(\rho_i^s) = \mathrm{tr}_e\Big(\Pi_m U(\rho_i^s \otimes \sigma_i^e)U^\dagger \Pi_m\Big)$, the final state of the principle system can be described as $\rho_f^s = \frac{\Phi_m(\rho_i^s)}{\mathrm{tr}\left(\Phi_m(\rho_i^s)\right)}$. Let $\sigma_i^e = \sum_j \lambda_j |j\rangle\langle j|$ be the spectral decomposition of σ_i^e and $\{|e_k\rangle\}_{k=1}^l$ be an orthonormal basis for the environment. We have then

$$\begin{aligned} \Phi_m(\rho_i^s) &= \sum_{j,k} \lambda_j I \otimes \langle e_k| \Big(\Pi_m U(\rho_i^s \otimes |j\rangle\langle j|)U^\dagger \Pi_m\Big) I \otimes |e_k\rangle \\ &= \sum_{j,k} A_{kj}\rho_i^s A_{kj}^\dagger, \end{aligned} \tag{2.9}$$

where $A_{kj} = \sqrt{\lambda_j}\langle e_k|\Pi_m U|j\rangle$. Equation (2.9) is the generalization of quantum operation formalism to non-trace-preserving operations. Indeed, the $\Phi_m(\cdot)$ in equation (2.9) is not trace-preserving, which arises from taking into account the measurement phenomenon.

One of the fundamental problems of the quantum operations from the perspective of their realization is understanding how they can be specified in laboratory, i.e., in the physical world. In reality, we do experiments and perform some measurements whose outcomes are just numbers, not operators. Any quantum operation can be characterized physically by means of performing some experiment, which is a procedure that is known as *quantum process tomography.* Determining an unknown quantum state so-called *quantum state tomography* (QST) is a part of quantum process tomography. The number of the copies of the unknown quantum state is a sort of resource. If only one copy of an unknown quantum state ρ is given, it is impossible to determine the state with certainty due to the fact that two non-orthogonal quantum states cannot be distinguished by any measurement procedure with certainty. Therefore, in order to characterize ρ, we need an ensemble of the copies of ρ. In this case, many copies of ρ are prepared in an experiment and they are subjected to some measurements. But, how many copies do we need? Our desire is to keep the number of copies as fewest as possible. For example, let us assume that ρ is the quantum state of a qubit system. The set $\{I/\sqrt{2}, X/\sqrt{2}, Y/\sqrt{2}, Z/\sqrt{2}\}$ forms an orthonormal basis for the state space of qubits with respect to Hilbert-Schmidt inner product so that

ρ may be expressed as

$$\rho = \frac{\mathrm{tr}(\rho)}{2} I + \frac{\mathrm{tr}(X\rho)}{2} X + \frac{\mathrm{tr}(Y\rho)}{2} Y + \frac{\mathrm{tr}(Z\rho)}{2} Z. \tag{2.10}$$

Recalling that $\mathrm{tr}(X\rho)$ is the expectation value of the operator X, the above expression tells us that one can reconstruct the unknown quantum state ρ of a qubit system if one performs the measurement of the basis elements, or in general an informationally complete measurement, on its ensemble. Since quantum operations are linear, it is enough to characterize their effect on a basis. Let us assume that $\Phi : \mathcal{L}(\mathcal{H}^d) \to \mathcal{L}(\mathcal{H}^d)$ be the quantum operation under consideration such that $\Phi(X) = \sum_k A_k X A_k^\dagger$. We choose a basis $\{E_l\}_{l=1}^{d^2}$ for $\mathcal{L}(\mathcal{H}^d)$ by which one can write $A_k = \sum_l t_{kl} E_l$. Eventually,

$$\Phi(X) = \sum_{k,l,p} t_{kl} t_{kp}^\star E_l X E_p^\dagger = \sum_{l,p} \gamma_{lp} E_l X E_p^\dagger, \tag{2.11}$$

where the numbers $\gamma_{lp} = \sum_k t_{kl} t_{kp}^\star$ forms a positive semi-definite matrix. Therefore, if we are able to determine the numbers γ_{lp} by means of an experiment, we will be determining the Kraus operators of the quantum operation $\Phi(\cdot)$. We enlist the main rules of the procedure of how to determine a quantum operation as follows:

1.1 Choose a set of vectors $\{|\psi_i\rangle\}_{i=1}^{d^2}$ whose projections $\{|\psi_i\rangle\langle\psi_i|\}_{i=1}^{d^2}$ forms a basis for the Hilbert space of d-dimension. A basic choice is the standard basis $\{E_{ij} = |i\rangle\langle j|\}$ for the Hilbert space of d-dimension.

1.2 For each $|\psi_i\rangle$, prepare the principle system in that state and subject it to the quantum operation Φ, which is to be characterized. Assuming the standard basis again, one needs only to prepare the principle system in the states $|i\rangle$, $|j\rangle$, $|x_{ij}\rangle = \frac{1}{\sqrt{2}}(|i\rangle + |j\rangle)$ and $|y_{ij}\rangle = \frac{1}{\sqrt{2}}(|i\rangle + \mathrm{i}|j\rangle)$. These are quantum states that can be physically constructible in the laboratory. The effect of the quantum operation on the elements $|i\rangle\langle j|$ can be expressed in terms of its effect on these quantum states: $\Phi(|i\rangle\langle j|) = \Phi(|x_{ij}\rangle\langle x_{ij}|) + \mathrm{i}\Phi(|y_{ij}\rangle\langle y_{ij}|) - \frac{1+\mathrm{i}}{2}\Phi(|i\rangle\langle i|) - \frac{1+\mathrm{i}}{2}\Phi(|j\rangle\langle j|)$.

1.3 After the quantum operation on the basis element $|\psi_i\rangle$ completes, apply a QST in order to determine the resultant quantum state: $\Phi(|\psi_i\rangle\langle\psi_i|) = \sum_{k=1}^{m} s_{ik}\rho_k$, where ρ_k's are the measurement elements for QST. In this step, the numbers s_{ik} are specified by the measurement for QST. In addition, choosing the type of measurement $\{\rho_k\}_{k=1}^{m}$ in principle depends on the experimenter's choice. Repeat step 1.3 for every $|\psi_i\rangle$ and for the fixed informationally complete measurement $\{\rho_k\}_{k=1}^{m}$ for QST.

We schematize the above rules in Figure 2.1.

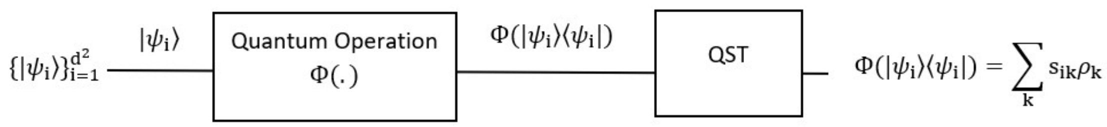


Figure 2.1 Schematic picture of quantum process tomography for a quantum operation $\Phi(\cdot)$. After operating the quantum operation $\Phi(\cdot)$ on the basis elements $|\psi_i\rangle$, the resultant quantum state $\Phi(|\psi_i\rangle\langle\psi_i|)$ is subjected to a quantum state tomography (QST) by means of an informationally complete measurement $\{\rho_k\}_{k=1}^m$. After all, the quantum operation $\Phi(\cdot)$ has been characterized by its effect on the basis elements.

Having known $\Phi(|\psi_i\rangle\langle\psi_i|)$, one can write

$$\Phi(|\psi_i\rangle\langle\psi_i|) = \sum_k \lambda_{ik} |\psi_k\rangle\langle\psi_k|, \tag{2.12}$$

where we can write $\lambda_{ik} = \mathrm{tr}(|\psi_k\rangle\langle\psi_k| \Phi(|\psi_i\rangle\langle\psi_i|))$ for simplicity without loss of generality. To proceed, we have the opportunity

$$E_l |\psi_i\rangle\langle\psi_i| E_p^\dagger = \sum_k \beta_{ik}^{lp} |\psi_k\rangle\langle\psi_k| \tag{2.13}$$

with $\beta_{ik}^{lp} = \mathrm{tr}\Big(|\psi_k\rangle\langle\psi_k| E_l |\psi_i\rangle\langle\psi_i| E_p^\dagger\Big)$ since $\{|\psi_i\rangle\langle\psi_i|\}_{i=1}^{d^2}$ is a basis. Using equation (2.13) in equation (2.11) and makes it equal to equation (2.13), we obtain finally the main result

$$\sum_{k,l,p} \gamma_{lp}\beta_{ik}^{lp} |\psi_k\rangle\langle\psi_k| = \sum_k \lambda_{ik} |\psi_k\rangle\langle\psi_k| \tag{2.14}$$

from which we achieve the equation $\sum_{l,p} \beta_{ik}^{lp}\gamma_{lp} = \lambda_{ik}$. From this equality we can form the matrix $\Gamma = (\gamma_{lp})$. Γ can be diagonalized by some unitary matrix U such that $U\Gamma U^\dagger = R = (r_i\delta_{ij})$ which leads to $\gamma_{lp} = \sum_{k,j} U_{kl}^\star r_k \delta_{kj} U_{jp} = \sum_k t_{kl} t_{kp}^\star$. In conclusion, we obtain $t_{kl} = \sqrt{r_k} U_{lk}$, and so thus, $A_k = \sqrt{r_k}\sum_j U_{jk} E_j$.

In the context of this thesis, we study two types of measurement in regard to entropic uncertainties. One of them is a set of projective measurements that are constructed from MUBs corresponding to a set of some specific observables. The other type of measurement is *symmetric, informationally complete, positive operator-valued measure* (SIC-POVM). In contrast to the projective measurements based on MUBs, a

SIC-POVM in a d-dimensional Hilbert space is an informationally complete measurement whose elements are d^2 projective operators up to a multiplication. Therefore, it is not a set of measurements but just one measurement.

2.1.1 Mutually unbiased bases

The importance of MUBs appears on the stage if the issue is related to QST. One of the fundamental tasks in the quantum information theory is how to extract the complete information of the quantum state of a system. To this aim, an informationally complete set of measurement elements with rank-1 is performed so that it is a maximally efficient measurement. Mutually Unbiased Bases (MUBs) (Wootters & Fields, 1989) provide such a measurement. In addition to their importance in the theoretical view (Durt et al., 2010), they have found room in diverse application areas such as quantum error correction (Calderbank et al., 1997), quantum cryptography (Mafu et al., 2013), entanglement detection (Spengler et al., 2012; Wang & Zheng, 2021), quantum key distribution (Cerf et al., 2002) and quantum state tomography (Ivanovic, 1981; Wootters & Fields, 1989).

Ivanovic (Ivanovic, 1981) first examined the importance of MUBs for QST. A general quantum state in d-dimensional Hilbert space has d^2-1 real parameters. Assuming that we perform a projective measurement in the space, we can specify d real parameters by means of the probabilities of the applied projective measurement. One of the probabilities can be determined in terms of the others because of the completeness relation, that is, the summation of the probabilities is equal to 1. Therefore, $d-1$ parameter can be determined by a projective measurement. Eventually, we need $d+1$ projective measurements to determine a quantum state in general that is deduced from counting the required projective measurements: $\frac{d^2-1}{d-1} = \frac{(d+1)(d-1)}{d-1} = d+1$.

One can construct a projective measurement from a basis. For example, let us assume that $\{|e_i\rangle\}_{i=1}^{d}$ be an orthonormal basis for d-dimensional Hilbert space. Then, $\{P_i = |e_i\rangle\langle e_i|\}_{i=1}^{d}$ forms a projective measurement. MUBs are orthonormal bases. Therefore, one can construct projective measurements from MUBs. For this reason, we say that MUBs represent projective measurements. The quest of the existence of $d+1$ MUBs enforces itself to us naturally. We call a set of $d+1$ MUBs as an informationally complete set of MUBs since a set of $d+1$ projective measurements constructing from MUBs forms a set that is informationally complete. Labeling the bases in a set of MUBs with the integer n and the basis element with the integer k, we can write an informationally complete set of MUBs as $\{|nk\rangle, n=0,1,\ldots,d, k=0,1,\ldots,d-1\}$ such that one reads $|nk\rangle$ as the k^{th} basis element of the n^{th} basis. As was mentioned above the eigenvectors of Pauli matrices are MUBs and since they are $2+1=3$, so do they form an informationally complete set of MUBs. In this case, we write $\{|nk\rangle, n=x,y,z; k=0,1\}$ so that $|x0\rangle = |x_0\rangle, |y1\rangle = |y_1\rangle$ and so on in equation (2.23). We are to construct projective measurements from MUBs trivially as $\{\Pi_{nk} = |nk\rangle\langle nk|\}_{n,k=0,0}^{d,d-1}$ such that they satisfy the completeness condition of a measurement: $\sum_{k=0}^{d-1} \Pi_{nk} = I$ for all n. Therefore, we revise the basic relation between the basis elements of a pair of MUBs in equation (2.20) in favor of the respective projection operators as

$$\operatorname{tr}(\Pi_{nk}\Pi_{ml}) = \frac{1+(d\delta_{kl}-1)\delta_{nm}}{d}. \tag{2.15}$$

Having these informationally complete set of MUBs and defining the probabilities $p_{nk} = \operatorname{tr}(\Pi_{nk}\rho)$, the quantum state ρ in d-dimensional Hilbert space is reconstructed as

$$\rho = \sum_{n,k=0,0}^{d,d-1} p_{nk}\Pi_{nk} - I. \tag{2.16}$$

From now on, whenever we mention the measurements of MUBs, we actually means the projective measurements that are constructed from MUBs. This is just a convention to economize writing.

As we noted above, the importance of an informationally complete set of MUBs arises when we performs a QST. In that case, we of course wish to choose an informationally complete measurement which minimizes our statistical errors in the sense that one estimates the probabilities as precisely as it is possible with the optimal usage of resources. We wish to give an example for this point to concrete the privilege status of MUBs over the other sets of $d+1$ projective measurements.

Let us assume that we are to ascertain the unknown quantum state of a qubit system by means of performing a set of measurements on the ensemble of N copies of the quantum state. Since we have a finite number N of the copies, we are surely not going to estimate the probabilities precisely. We seek for 3 measurements that minimizes our statistical error while fixing the number of the copies. The quantum state of a qubit system is expressed as $\rho = \frac{1}{2}(I + \mathbf{r}\cdot\sigma)$ such that $\mathbf{r} = (r_1, r_2, r_3) \in \mathbb{R}^3$ with $|\mathbf{r}| \leq 1$ is the *Bloch vector* and $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ is the vector representation of Pauli matrices. We perform the measurement of some operators having the form $A = \mathbf{a}\cdot\sigma$ (Durt et al., 2010) since any operator on 2-dimensional Hilbert space can be written as a linear combination of Pauli operators and the identity, which we ignore without lost of generality. It is enough to measure three operators like A whose Bloch vectors $\mathbf{a}$ do not rely on the same plane. Each measurement restricts the Bloch vector to a specific plane and the intersection of the planes characterizes $\mathbf{a}$ with certainty. However, the absent of precisely defining the probabilities precludes us to restrict the Bloch vector to a plane but rather to, as it were, a plane having some thickness. All measurements together thus leave the Bloch vector in a volume. Our concern then is that what kind of three measurements minimizes that volume, which is accounted for the statistical error (Wootters & Fields, 1989). The volume is a parallelepiped in general and it would be minimum if it takes a cube shape which is formed by three orthogonal planes. This implies that if the Bloch vectors characterizing the planes are orthogonal, the statistical error is minimum. Surprisingly, two operators $A = \mathbf{a}\cdot\sigma$ and $B = \mathbf{b}\cdot\sigma$ are mutually unbiased, or complementary, if their Bloch vectors are orthogonal, i.e., $\mathbf{a}\cdot\mathbf{b} = 0$ (Durt et al., 2010). In our example, the simple choices for Bloch vectors are $\mathbf{a}_1 = (1,0,0), \mathbf{a}_2 = (0,1,0)$ and $\mathbf{a}_3 = (0,0,1)$, which are corresponding to Pauli operators. This geometrical reasoning can be extended straightforwardly to higher dimension by means of the generalized Pauli matrices. As is well known, Pauli matrices are the generators of $SU(2)$ group. Any quantum state ρ in d-dimensional Hilbert space can be expanded in terms of the generators $\{\Lambda_j\}_{j=1}^{d^2-1}$ of $SU(d)$ as

$$\rho = \frac{1}{d}I + \frac{1}{2}\sum_{j=1}^{d^2-1}\lambda_j\Lambda_j, \tag{2.17}$$

where $\lambda_j = \mathrm{tr}(\rho\Lambda_j)$ is the expectation value. Then, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_{d^2-1}) \in \mathbb{R}^{d^2-1}$ can be considered as Bloch vector in d-dimensional Hilbert space (Kimura, 2003). Based on this extension and the above reasoning, we can argue directly that there are not more than $d+1$ MUBs, that is indeed the case (Ivanovic, 1981; Wootters & Fields, 1989). In conclusion, we call $d+1$ MUBs optimal, or minimizing the statistical error, in the above sense.

In his famous paper (Heisenberg, 1927), Heisenberg introduced the celebrated uncertainty principle of position and momentum operators. Without dealing with its technical expression, its most important conclusion was that if one has the full information of one of these operators, the other operator becomes fully uncertain in the sense that there is no way to acquire any piece of information about it while respecting the rules of quantum mechanics. This maximal trade-off relation between position and momentum operators is called complementarity: our certainty of one of them leaves the other fully uncertain. Speaking of any finite system, we naturally seek for operators that have complementarity relation. The formal expression and importance of complementary operators, especially in finite dimensions, was first highlighted by Schwinger (Schwinger, 1960) in the sense that the bases corresponding to complementary operators represent measurements that are maximally unbiased. In quantum information theory, two complementary operators are known as *mutually unbiased*. Let us assume that we have two non-degenerate cyclic operators X and Z with period d, that is, $X^d = I$ and $Z^d = I$, where I is the identity and any power less than d of them does not give identity. The eigenvalues of X and Z are then going to be d roots of unity,

$$X|x_k\rangle = w_d^k |x_k\rangle\,; \quad Z|z_k\rangle = w_d^k |z_k\rangle\,, \quad \text{with} \quad w_d = e^{\frac{i2\pi}{d}}. \tag{2.18}$$

That the cyclic operators X and Z are mutually unbiased can be stated as

$$\frac{1}{d}\operatorname{tr}(X^s Z^p) = \delta_{s0}\delta_{p0} \quad \text{for all} \quad s,p = 0,1,\ldots,d-1. \tag{2.19}$$

If two operators X and Z on a Hilbert space of d-dimension are mutually unbiased, so do aX and bZ with $ab \neq 0$. In addition, if a unitary transformation takes X to X' and Z to Z', the pair (X', Z') is also mutually unbiased. Therefore, we can anytime turn our attention from the operators X and Z to their corresponding set of eigenvectors $\{|x_k\rangle\}_{k=1}^d$ and $\{|z_k\rangle\}_{k=1}^d$ each of which forms a basis for the Hilbert space of d-dimension provided that the operators are non-degenerate. From now on, we will focus on the bases of mutually unbiased operators which are called mutually unbiased bases (MUBs). Formally, two orthonormal bases $\{|e_k\rangle\}_{k=0}^{d-1}$ and $\{|f_l\rangle\}_{l=0}^{d-1}$ of a d-dimensional Hilbert space are called mutually unbiased bases if

$$|\langle e_k|f_l\rangle| = \frac{1}{\sqrt{d}} \quad \text{for all} \quad k,l \in \{0,1,\ldots,d-1\}. \tag{2.20}$$

Equation (2.19) and equation (2.20) result in each other, which means that they are equivalently the statement of mutually unbiasedness. It is always easy to construct two MUBs by Fourier transformations. To be more explicit, if $\{|e_j\rangle\}_{j=0}^{d-1}$ is the standard basis of d-dimensional Hilbert space, its Fourier transformation

$$|f_k\rangle = \frac{1}{\sqrt{d}}\sum_{j=0}^{d-1} w_d^{-kj}|e_j\rangle \tag{2.21}$$

gives another basis $\{|f_k\rangle\}_{k=0}^{d-1}$, which is mutually unbiased with respect to it, that is, $\langle f_k|e_j\rangle = \frac{1}{\sqrt{d}}w_d^{kj}$ for all $j,k \in \{0,1,\ldots,d-1\}$. Here, $w_d = e^{\frac{2\pi i}{d}}$. The most famous and familiar MUBs are the set of eigenvectors of Pauli matrices $\{\sigma_x, \sigma_y, \sigma_z\}$. Indeed, we have

$$\begin{aligned} \sigma_x|x_j\rangle =& w_2^j|x_j\rangle\,, \quad \sigma_x^2 = I; \quad \sigma_y|y_j\rangle = w_2^j|y_j\rangle\,, \quad \sigma_y^2 = I; \\ \sigma_z|z_j\rangle =& w_2^j|z_j\rangle\,, \quad \sigma_z^2 = I, \quad j = 0,1 \end{aligned} \tag{2.22}$$

with $\frac{1}{2}\operatorname{tr}\left(\sigma_x^j\sigma_y^k\right) = \frac{1}{2}\operatorname{tr}\left(\sigma_x^j\sigma_z^k\right) = \frac{1}{2}\operatorname{tr}\left(\sigma_y^j\sigma_z^k\right) = \delta_{j0}\delta_{k0}$ and $|\langle x_j|y_k\rangle| = |\langle x_j|z_k\rangle| = |\langle y_j|z_k\rangle| = \frac{1}{\sqrt{2}}$ for $j,k = 0,1$. The bases are

$$\begin{aligned}&\left\{|x_0\rangle = \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix}, |x_1\rangle = \begin{pmatrix}\frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}}\end{pmatrix}\right\}, \left\{|y_0\rangle = \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{\mathrm{i}}{\sqrt{2}}\end{pmatrix}, |y_1\rangle = \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{-\mathrm{i}}{\sqrt{2}}\end{pmatrix}\right\},\\ &\left\{|z_0\rangle = \begin{pmatrix}1\\0\end{pmatrix}, |z_1\rangle = \begin{pmatrix}0\\1\end{pmatrix}\right\}.\end{aligned} \tag{2.23}$$

It is easy to construct $d+1$ MUBs by using *Weyl-Heisenberg group* of unitary operators if the finite dimension we study on is a prime number. Recalling the operators in equation (2.18), we express them as

$$\begin{aligned}X &= \sum_{k=0}^{d-1} |k+1\rangle\langle k|,\\ Z &= \sum_{k=0}^{d-1} w_d^k |k\rangle\langle k|, \quad w_d = e^{\frac{2\pi\mathrm{i}}{d}},\end{aligned} \tag{2.24}$$

on the field $\mathbb{Z}_d = \{0,1,\ldots,d-1\}$, whose addition and multiplication operations are performed according to modulo(d). The group of the unitary operators generating from these two unitary operators is

$$WH = \left\{X, Z, XZ, XZ^2, XZ^3, \ldots, XZ^{d-1}\right\}, \tag{2.25}$$

which is known as Weyl-Heisenberg group. The basic and general relation of the elements is that $X^m Z^n = w_d^{-mn} Z^n X^n$. The eigenbases of the elements of Weyl-Heisenberg group provides an informationally complete set of MUBs, i.e., $d+1$ MUBs. In addition to the eigenbasis of Z, $\{|0k\rangle\}_{k=0}^{d-1}$, which is considered as computational basis, the eigenbases of the elements XZ^{m-1} are $\{|mk\rangle\}_{m,k=1,0}^{d,d-1}$, with the explicit expression (Bandyopadhyay et al., 2002)

$$|mk\rangle = \frac{1}{\sqrt{d}}\sum_{j=0}^{d-1} w_d^{-jk-(m-1)s_j}|j\rangle, \quad \text{for} \quad m = 1,2,\ldots,d \tag{2.26}$$

such that $|j\rangle = |0j\rangle$, $s_j = j + j + 1 + \ldots + d - 1$ and $XZ^{m-1}|mk\rangle = w_d^k|mk\rangle$. We also note that $XZ^{l-1}|mk\rangle = w_d^{k+m-l}|m(k+m-l)\rangle$ for $l = 1,2,\ldots,d$. The set $\{|mk\rangle\}_{m,k=0,0}^{d,d-1}$ then forms an informationally complete set of MUBs. For instance, in the light of above methods, we have in dimension $d = 3$ the unitary operators

$\{Z, X, XZ, XZ^2\}$ with the respective following expressions

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & w_3^2 \\ 1 & 0 & 0 \\ 0 & w_3 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & w_3 \\ 1 & 0 & 0 \\ 0 & w_3^2 & 0 \end{pmatrix} \tag{2.27}$$

whose corresponding bases

$$\begin{gathered} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{w_3^2}{\sqrt{3}} \\ \frac{w_3}{\sqrt{3}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{w_3}{\sqrt{3}} \\ \frac{w_3^2}{\sqrt{3}} \end{pmatrix} \right\} \\ \left\{ \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{w_3}{\sqrt{3}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{w_3^2}{\sqrt{3}} \\ \frac{w_3^2}{\sqrt{3}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{w_3}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \right\}, \left\{ \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{w_3^2}{\sqrt{3}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{w_3^2}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{w_3}{\sqrt{3}} \\ \frac{w_3}{\sqrt{3}} \end{pmatrix} \right\} \end{gathered} \tag{2.28}$$

forms an informationally complete set of MUBs.

We are still able to construct $d+1$ MUBs if the dimension d is a power of an odd prime number, that is, if $d = p^r$, where p is an odd prime number and r is a positive integer. The construction of $d+1$ MUBs in this case is based on Galois fields $\mathbb{GF}(p^r)$, the finite fields with characteristic p and cardinality $d = p^r$ (Wootters & Fields, 1989; Bandyopadhyay et al., 2002). The idea substantially consists on the case of when the dimension is a prime number. We always refer to the number p as an odd prime number in the following unless otherwise stated. Additionally, the following prescription of constructing MUBs is a summary of the construction given in (Wootters & Fields, 1989) together with a little revision of the notation that was used there so as to stay in coherent with the notation we use in the thesis.

As a consequence of a theorem from abstract algebra, the number of the elements of any finite field is equal to a power of a prime number. If the dimension is a prime number p, the two binary operations of a finite field, that is addition and multiplication, are simply defined according to modulo(p). It is always possible to find an n-degree polynomial on a finite field of cardinality p which is not solvable, that is, that polynomial does not have solution in the field. From such a polynomial, one can always construct a field of cardinality p^n and of characteristic p. To give an example, the set of complex numbers is an extension of real numbers with the polynomial $x^2 + 1 = 0$, which is not solvable in real numbers. Let us assume that $\mathbb{F}_p = \{0, 1, \ldots, p-1\}$ be a finite field of cardinality p. We first consider an insoluble, n^{th}-degree polynomial on $\mathbb{F}_p$: $\beta_0 + \beta_1 x + \ldots + \beta_n x^n = 0$ such that $\beta_0, \beta_1, \ldots, \beta_n \in \mathbb{F}_p$. We then introduce a solution α, like the imaginary number $\mathrm{i} = \sqrt{-1}$, and for the closure of the multiplication operation we need to also introduce the powers of α: $\alpha^2, \alpha^3, \ldots, \alpha^{n-1}$. We then take the linear combination of the powers of α on $\mathbb{F}_p$ such that the set $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$ plays the role of a basis for a vector space whatsoever of dimension n. For example, $\{1, \mathrm{i}\}$ can be considered as a basis in the case of complex numbers, and such that any complex number can be expressed as a linear combination like $a + ib$ on real numbers $\mathbb{R}$, i.e., $a, b \in \mathbb{R}$. The set of all linear combinations

$$\mathbb{F}_{p^n} = \{a_0 + a_1\alpha + \ldots + a_n\alpha^n \mid a_0, a_1, \ldots, a_n \in \mathbb{F}_p\} \tag{2.29}$$

then forms a field with the fact that addition and multiplication operations are performed based on modulo$(\beta_0 + \beta_1\alpha + \ldots + \beta_n\alpha^n)$. For example, let us assume that $p^2 = 3^2 = 9$, and as is known that $\mathbb{F}_3 = \{0, 1, 2\}$. An irreducible, that is, unsolvable, second-order polynomial on $\mathbb{F}_3$ is $x^2 + 1 = 0$. Introducing α to be a root of this equation, we obtain $\mathbb{F}_9 = \{0, 1, 2, \alpha, 1+\alpha, 2+\alpha, 2\alpha, 1+2\alpha, 2+2\alpha\}$. Then, One can construct the addition and multiplication tables for $\mathbb{F}_9$ bearing in mind the fact that $\alpha^2 + 1 = 0$.

We also define a trace operation over $\mathbb{F}_{p^n}$: $\mathrm{Tr} : \mathbb{F}_{p^n} \to \mathbb{F}$ such that $\mathrm{Tr}(a) := a + a^p + a^{p^2} + \ldots + a^{p^{n-1}}$, for all $a \in \mathbb{F}_{p^n}$. The trace function has the following properties:

i. $\mathrm{Tr}(a) \in \mathbb{F}_p$ for all $a \in \mathbb{F}_{p^n}$.

ii. Tr is linear in $\mathbb{F}_{p^n}$ where the coefficients are from $\mathbb{F}$, that is, $\mathrm{Tr}(c_1 a + c_2 b) = c_1 \mathrm{Tr}(a) + c_2 \mathrm{Tr}(b)$ for all $a, b \in \mathbb{F}_{p^n}$ and for all $c_1, c_2 \in \mathbb{F}_p$.

iii. Linear mappings from $\mathbb{F}_{p^n}$ to $\mathbb{F}_p$ have the form $a \mapsto \mathrm{Tr}(ba)$ for $b \in \mathbb{F}_{p^n}$.

Having now the theoretical background of $\mathbb{F}_{p^n}$, we introduce the MUBs for dimension $d = p^n$. We still use the integers like i, j, k, l, r, etc. as indices meanwhile also for representing the field numbers. As with the case of the dimension $d = p$, we start with the computational basis again, by taking it as $\{|0k\rangle\}_{k=0}^{p^n-1}$. The remaining p^n MUBs are given by the explicit expression of their elements as follows

$$|nk\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{p^r-1} w_p^{\mathrm{Tr}\left(nj^2+kj\right)} |0j\rangle\,; n, k, l \in \mathbb{F}(p^r), n \neq 0 \tag{2.30}$$

Mutually unbiasedness of the constructed bases above is guaranteed by the equality

$$\left| \sum_{j \in \mathbb{F}_{p^n}} e^{2\pi \mathrm{i}/p\, \mathrm{Tr}\left(mj^2+nj\right)} \right| = \sqrt{p^n}, \quad (m \neq 0, \quad \text{p is an odd prime}). \tag{2.31}$$

The above treatment does not work for dimensions $d = 2^r$ since left side of equation (2.31) is zero any more. In that case, the following construction for $d = 2^r$ was proposed in (Wootters & Fields, 1989). Since any field $\mathbb{F}_{p^r}$ can be considered as a vector space, any element $a \in \mathbb{F}_{p^r}$ can be written as a linear combination of a basis $\{v_i\}_{i=1}^r$: $a = \sum_{i=1}^r a_i v_i$, where $a_i \in \mathbb{F}_p$. We define the product of the basis elements as

$$v_j v_k = \sum_{s=1}^{n} \alpha_{jk}^{(s)} v_s, \tag{2.32}$$

from which we define the matrices $\alpha^{(i)} = (\alpha_{ij}^{(i)})_{r\times r}$. Bearing in mind these matrices, we first consider the standard basis $\{|0k\rangle\}_{k=0}^{d-1}$ again, and the remained d MUBs are

$$|nk\rangle = \frac{1}{\sqrt{2^r}} \sum_{j=0}^{2^r-1} \mathrm{i}^{j^T(n\cdot\alpha)j} (-1)^{k\cdot j} |j\rangle, \quad n,k,j \in \mathbb{F}_{2^r}, r \neq 0. \tag{2.33}$$

Here the indices n,k,j are r-component vectors with elements in the set $\{0,1\}$, and $\alpha = (\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(r)})$. In addition, i is the imaginary number, and the multiplication and addition in the exponent of i is modulo (4). For further information, we refer to (Wootters & Fields, 1989).

However, the treatment ensured by Galois fields does not work for the composite dimension d, such as $d = 6, 10, 15$, etc. because the cardinality of a finite field is always a power of a prime number.

The existence of an informationally complete set of MUBs is still a conundrum if the dimension is a composite number. Nevertheless, there exist at least 3 MUBs in any d-dimensional Hilbert space. Therefore, the first step of an approach to the existence of MUBs in the composite dimensions is to study the fourth MUB. In line with this task, dimension six is the lowest composite dimension, which occupies a central importance in searching the existence of more than three MUBs in composite dimension. A proof of the existence of more than three MUBs in dimension six has not still been given, although lots of works have been dedicated to the understanding of them (Brierley & Weigert, 2008; Jaming et al., 2010; Goyeneche, 2013; D'Ambrosio et al., 2013; Chen & Yu, 2017, 2018; Liang et al., 2021). There are also some work that have tried to understand the existence of MUBs by looking at their relations with SIC-POVMs (Wootters, 2006; Albouy & Kibler, 2007; Appleby, 2009). In the sphere of this thesis, we shall use MUBs when studying the entropic uncertainty.

2.1.2 An informationally complete measurement: SIC-POVM

Another celebrated, informationally complete measurement is SIC-POVM, which has been occupying a central position in quantum information theory since the work by Zauner in 1999 (Zauner, 2011). Zauner showed that a set of equiangular vectors can be constructed using the Weyl-Heisenberg group up to dimension 7 by means of an appropriate complex vector known as a fiducial vector, and conjectured that fiducial vectors exist in every dimension. The existence problem of the fiducial vector in every dimension is known as Zauner's conjecture. Currently, SIC-POVMs are studied numerically and analytically. The numerical solutions have shown that SIC-POVM are related to deeper questions in number theory, and have allowed the formulation of exact solutions from the numeric ones (Scott, 2006; Appleby et al., 2017, 2018; Appleby & Bengtsson, 2019). Some connection between the solutions of different dimensions has been revealed by the analytical studies of the SIC-POVMs (Appleby et al., 2017; Fuchs et al., 2017). Currently, the solutions based on Weyl-Heisenberg group have been constructed both numerically and analytically in all dimensions up to 151 and in special dimensions as high as 844 (Grassl & Scott, 2017). Focusing on the analytical research, some remarkable algebraic properties of SIC-POVMs that are covariant under the action of Weyl-Heisenberg group have been explored by focusing on the extended Clifford group, which is a group of unitary and anti-unitary operators, and is the normalizer of Weyl-Heisenberg group (Appleby, 2005; Zhu, 2010; Hughston & Salamon, 2016).

By definition, a SIC-POVM in a d-dimensional Hilbert space is constructed from a set of d^2 vectors, $\{|\psi_i\rangle\}_{i=1}^{d^2}$, where the norm of the inner product of any two vectors is constant: $|\langle\psi_i|\psi_j\rangle|^2 = \frac{d\delta_{ij}+1}{d+1}$. This means that the vectors are equiangular to each other. The measurement elements of the SIC-POVM is the set of rank one operators $\{\Pi_j = \frac{1}{d}|\psi_i\rangle\langle\psi_i|\}_{i=1}^{d^2}$, which are satisfying the completeness condition: $\sum_{i=1}^{d^2}\Pi_i = I$. Defining the probabilities of detecting the outcomes corresponding to the measurement elements as $p_i := \mathrm{tr}(\Pi_i\rho)$, where ρ is the quantum state under consideration, we can reconstruct ρ from the SIC-POVM as

$$\rho = (d+1)d\sum_{i=1}^{d^2} p_i\Pi_i - I. \tag{2.34}$$

SIC-POVMs are optimal measurement for QST (Renes et al., 2004; Scott, 2006; Fuchs & Schack, 2013). In addition to their use for QST, they are also used in the foundational study of quantum mechanics, where a specific SIC-POVM is chosen primarily as the underlying measurement for reformulating quantum mechanics completely based on probabilities (Fuchs & Schack, 2013; Yashin et al., 2020). They have been used also for quantum cryptography (Renes, 2005), quantum communication (Oreshkov et al., 2011; Szymusiak & Słomczyński, 2016) and entanglement detection (Chen et al., 2015).

All known SIC-POVMs with the exception of the Hoggar SIC-POVM in dimension 8 (Hoggar, 1998), are constructed by using Weyl-Heisenberg group (Bengtsson, 2017; Fuchs et al., 2017). For any dimension $d \in \mathbb{N}$, let $\{|k\rangle\}_{k=0}^{d-1}$ be an orthonormal basis for $\mathbb{C}^d$, and we define the following operators

$$w = e^{\frac{2\pi i}{d}}, \quad D_{jk} = w^{jk} \sum_{m=0}^{d-1} w^{jm} |k \oplus m\rangle \langle m|, \tag{2.35}$$

where $\oplus$ denotes addition modulo d. Then, it was conjectured that there exists a normalized fiducial vector $|\phi\rangle \in \mathbb{C}^d$, such that the set $\{D_{jk}|\phi\rangle\}_{j,k=1}^{d}$ is a SIC-POVM. The existence of a fiducial vector is still an open problem, which in turn means that the existence of SIC-POVMs for every dimension has not been proved yet (Kopp, 2021; Appleby et al., 2021).

A fiducial vector in dimension $d = 8$ was proposed by Jedwab and Wiebe such that the SIC-POVM is obtained by acting Weyl-Heisenberg group elements on it (Jedwab & Wiebe, 2016). Weyl-Heisenberg group generators in dimension $d = 2$ are

$$Z = |0\rangle\langle 0| - |1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad X = |1\rangle\langle 0| + |0\rangle\langle 1| = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{2.36}$$

from which we obtain the group $\{I, U, V, UV\}$. We consider threefold tensor product G of this set: $G = \{A_1 \otimes A_2 \otimes A_3 \mid A_i \in \{I, U, V, UV\}\}$. Having the proposed fiducial vector $|\phi\rangle = \left(0, 0, \frac{1+i}{\sqrt{2}}, \frac{1-i}{\sqrt{2}}, \frac{1+i}{\sqrt{2}}, -\frac{1+i}{\sqrt{2}}, 0, \sqrt{2}\right)^T$, the set $\{A|\phi\rangle \mid A \in G\}$ is a SIC-POVM.

2.2 An outlook of entropy

We studied mainly the entropic uncertainties of MUBs and SIC-POVM as two important quantum measurements. We now wish to present some important entropies and their characteristic properties. It would better to give a clear reason for the structure of this subsection:

"The literature devoted to analysing the second law of thermodynamics

> in the context of statistical mechanics starts with the development of statistical mechanics in the late 19th century. Considerable confusion has arisen simply due to the fact that as the subject has developed, *the meaning of key terms have changed or become ambiguous.* When one paper speaks of the second law being violated or of entropy decreasing and another of it being saved or being non-decreasing, *it is not necessarily the case that they are referring to the same things"* (Maroney, 2009) [emphasis added].

There are many entropy functions that are used in diverse but relevant fields successfully. Even this bare fact alone suggests us that entropy is not as fundamentally clear as energy is. Based on this suggestion and the given quotation, we wish first to draw a clear outline of some relevant entropies and the reasons behind their introduction.

The fundamental problem of thermodynamics is to find the equilibrium state of an overall isolated composite system accompanied by relaxing some internal constraints. Aiming at this goal, thermodynamics states that: *"All physical systems in thermal equilibrium can be characterized by a quantity called entropy, and that this entropy cannot decrease in any process in which the system remain adiabatically isolated, i.e., shielded from heat exchange with its environment"* (Uffink, 2001). In the frame of thermodynamics, entropy is characterized by the following properties: (1) The minimality of energy in an equilibrium state corresponds to maximality of entropy; (2) entropy is a continuous, differential and monotonically increasing function of energy; (3) entropy of a composite system of two independent subsystems is the sum of the entropies of the subsystems, i.e., entropy is an additive function. The first property yields to concavity of entropy; the third property requires that entropy is a first-order homogeneous function of extensive parameters, while the second property implies that entropy is invertible with respect to energy, which first and foremost suggests that entropy has a fundamental and determinative relation with energy. Based on these properties and *asserting that the state of a thermodynamics system can be fully characterized by energy (U), number of molecules (N) and the volume of the considered system (V)*, thermodynamic entropy, $S(U,V,N)$, can be expressed as

$$S(U,V,N) = \partial_U S(U,V,N)U + \partial_V S(U,V,N)V + \partial_N S(U,V,N)N \tag{2.37}$$

First and second properties imply that $\partial_U S(U,V,N) > 0$, which enables us to rewrite equation (2.37) in the form of energy expansion

$$
\begin{aligned}
U(S,V,N) &= \frac{1}{\partial_U S(U,V,N)} S - \frac{\partial_V S(U,V,N)}{\partial_U S(U,V,N)} V - \frac{\partial_N S(U,V,N)}{\partial_U S(U,V,N)} N \\
&= T(S,V,N)S - P(S,V,N)V + \mu(S,V,N)N
\end{aligned} \tag{2.38}
$$

from which one can deduce that $\partial_U S(U,V,N) = \frac{1}{T}(U,V,N)$, $\partial_V S(U,V,N) = \frac{P}{T}(U,V,N)$ and $\partial_N S(U,V,N) = -\frac{\mu}{T}(U,V,N)$. In the context of this thesis, the hallmark of thermodynamic entropy can be characterized by two such properties: 1) it is a measure of the uncertainty (or disorder) inherent to the system; 2) there is a one-to-one and determinative correspondence between entropy and energy, and 3) it is defined for the equilibrium states, i.e., non-equilibrium states are out of the context of thermodynamic entropy. Indeed, the differentiability in the second property of entropy is inferred from a detailed analysis of Carnot's cycle (Kondepudi & Prigogine, 2015, pp. 89-104), which is a reversible process and directly yields that the integral of ${đQ}/{T}$ over a closed a path $\mathcal{C}$ representing a reversible process is

$$
\oint_{\mathcal{C}} \frac{đQ}{T} = 0, \tag{2.39}
$$

which suggested Clausius to define thermodynamic entropy, $dS := {đQ}/{T}$. If one wishes to generalize this definition to any process, it must be assumed that any two thermodynamics states can be connected by a reversible transformation or process (Kondepudi & Prigogine, 2015, p. 101). The concept of thermodynamic entropy as a state function is purely macroscopic and it is stated that, as the second law of thermodynamics, thermodynamic entropy is a non-decreasing state function, whose validity is based on the irreversible processes. However, in contrast to the irreversibility of processes, the law of both classical and quantum mechanics are reversible. The classical and quantum laws of motion admit the time reversal of any evolution from a state A to a state B. For example, the flow of gas molecules confined in one half of a container to the whole container and its reverse (which violates the second law) are admissible to classical and quantum mechanics. Processes that are not allowed by the second law do not violate the laws of mechanics. Fundamentally, bearing in mind that all irreversible processes are the consequences of the motion of molecules ruled by the laws of mechanics, some of questions arise naturally: 1) How can irreversible processes emerge from the reversible processes of of molecules?; 2) What connection can be derived between entropy as a macroscopic property of a system and the microscopic constituents of the system?; 3) If there is a determinative relation between entropy and energy, and if energy of a system is the sum of the energies of its microscopic constituents, can one determine, if possible and reasonable, entropy in terms of microscopic properties of the system? Addressing to such sort of questions, Boltzmann obtained (thermodynamics) entropy as the logarithm of the number of microscopic states W corresponding to a macroscopic state:

$$H_{BG} = k_B \ln(W), \tag{2.40}$$

where k_B is the Boltzmann constant. This expression corresponds to the entropy of an isolated system. This is the first step that links entropy of a system to the statistical properties of the system. In the context of statistical mechanics, thermodynamic entropy is also known as *Boltzmann-Gibbs entropy*, which we use hereafter. When the system of inquiry is in contact with a heat reservoir, Boltzmann-Gibbs entropy takes the form $H_{BG} = -k_B \sum_i p_i \ln(p_i)$ such that p_i is the probability of the state having energy ϵ_i that the system to be in.

To summarize, Boltzmann-Gibbs entropy with its properties given above is a consequence of reversible processes, which are considered to be in thermodynamic equilibrium at every stage of their evolution.

On the other hand, according to Shannon, the fundamental problem of communication theory is to reproduce at some point at a possible desired accuracy a message, which is delivered from another point. As is well known, messages have meaning, they *"refer to or are correlated according to some system with certain physical or conceptual entities."* (Shannon, 1948, p. 1). However, according to Shannon, it is not this semantic aspect of messages that is *relevant* to the problem of engineering but is the fact that a message is a selection from a set of messages. In other words, the significant aspect of a message is its distinction from (or its distinguishability among) a set of messages. Immediately after putting forward this statement, a very natural question imposes itself at the outset upon the reason to be answered urgently: How can one quantify the information content of the set of considered messages? The answer was given by Shannon to be entropy based on some requirements that were set down as the pillars of the definition of entropy. These requirements are as follows (Shannon, 1948):

i. Entropy function H is a continuous function of probabilities, $\{p_i, i = 1, 2, \ldots, n\}$, of the messages released by the information source.

ii. If all messages have equal probabilities, $p_i = \frac{1}{n}$, entropy function $H(p_1, p_2, \ldots, p_n)$ is a monotonically increasing function of n.

iii. If a message is broken into two successive messages, the original entropy should be the weighted sum of the individual entropies of the new messages sample.

Later, Khinchin reformulated the requirements in a rigorous mathematical manner which determine the entropy proposed by Shannon uniquely (Khinchin, 1957, pp. 9-13). Today, these requirements are known as Shannon-Khinchin (SK) axioms, which can be enlisted as follows:

2.1 *Continuity* (SK1): Entropy is a continuous and non-negative function of probabilities of the occurrence of messages.

2.2 *Principle of Maximum Value* (SK2): Having n messages with corresponding probability distribution $\{p_i, i = 1,2,\ldots,n\}$, entropy function $H(p) := H(p_1,p_2,\ldots,p_n)$ takes its largest value when $p_k = \frac{1}{n}$ $(k = 1,2,\ldots,n)$.

2.3 *Expansibility* (SK3): Adding impossible messages to the set of messages does not change the entropy of a scheme, i.e.,
$H(p_1,p_2,\ldots,p_n,0) = H(p_1,p_2,\ldots,p_n)$.

2.4 *Additivity* (SK4): The joint entropy of two independent subsystems A and B is equal to the sum of the individual entropies, i.e., $H(A \cup B) = H(A) + H(B)$.

Based on SK axioms, entropy for measuring the information content of a set of messages is defined uniquely, $H_S(p) := -\sum_i p_i \ln(p_i)$, which is known as *Shannon entropy.* The first three axioms correspond to the requirements respectively that was stated by Shannon. Axiom SK1 was reformulated by B. Lesche in a mathematical syntax (Lesche, 1982):

Definition 2.2.1 (Lesche Stability) *Having a system which takes countable energy values together with a corresponding probability distribution of microstates* $\{p_i, i = 1,2,\ldots,n \mid n \in \mathbb{N}\}$*, and* l_1*-norm,* $\|p' - p\|_1 = \sum_{i=1}^{n} |p'_i - p_i|$*, the entropy of the system is a continuous measurable quantity if and only if there is at least a* $\delta > 0$ *such that for all* $\varepsilon > 0$*, if* $\|p' - p\|_1 < \delta$ *then* $\frac{|H_S(p') - H_S(p)|}{H_S^{max}(p)} < \varepsilon$.

The reason behind Lesche stability so as to consider it as a fundamental requirement for the validity of an entropy is that any physically observable quantity must change infinitesimally if an infinitesimal change happens to probability distribution, otherwise the experimental observations of the quantity cannot be reproducible (Abe, 2002). Apart from everything, SK axioms fundamentally represent a distillation of the requirements underlying the physical reasoning on the problem of quantifying information content of a set of messages. However, a question arises: Do SK axioms comprise exhaustively all sort of information content of all systems, for example, the systems having long range interaction? Since SK axioms are based on the nature of information, and since, according to Shannon, that *the messages [states of a system] are correlated according to some system with certain physical entities are irrelevant to the concept of entropy* (Shannon, 1948), one cannot put forward an affirmative answer to the question with certainty. We shall give an example to this fact in the context of the existence of MUBs in section 3.6. This fact implies that there is not an equivalence relation between the comprehensive requirements for information quantification and the formal SK axioms. In other words, both requirements for information quantification and SK axioms cannot be biconditionally related to each other due to the logical connection that links them. Bearing in mind this fact, any critic or attempt that aims to generalize Shannon entropy should pay attention not on SK axioms, but on the requirements that are backbone of SK axioms, and on the constraints imposed on the constituents of the relevant system. Unlike Shannon entropy, Boltzmann-Gibbs entropy has a determinative relation with the energy of the system, which enables one to examine Boltzmann-Gibbs entropy on the ground of physical considerations rather than in the abstract frame of some axioms. Another but maybe the most important difference between Boltzmann-Gibbs entropy and Shannon entropy is that while the former in physics quantifies the number of microstates corresponding to a particular macrostate, the latter in information theory serves to infer the most attainable information under some given constraints (Jizba & Korbel, 2019).

However, regarding SK axioms comprehensive so that Shannon entropy is general enough to comprise all kind of information content of systems, a very fundamental question is exist that one must deal with in the context of engineering: having a priori probability distribution $\{p_i\}_{i=1}^n$ of a set of messages or events, how can one obtain the best estimate of a posterior probability distribution $\{p'_i\}_{i=1}^n$ after learning some constraint on the events in terms of the expectation value of some certain function, $\sum_i p_i f(x_i)$, or in terms of some certain bounds on the values of these functions? Addressing to this question, Jaynes suggested *principle of maximum entropy* by proposing that it is uniquely correct method for the best estimate of a posterior probability distribution after learning some constraints in terms of the expectation values of some certain functions (Jaynes, 1957a,b). *Principle of maximum entropy states that among all probability distributions that satisfy the constraints one must choose the one for which Shannon entropy $H_S(p)$ takes its largest value.* To put it in a formal expression, assuming that $\sum_i p_i f(x_i) = E_f$ is the given constraint, one maximizes the expression

$$H_S(p) - \lambda\Big(\sum_i p_i - 1\Big) - \beta\Big(\sum_i p_i f(x_i) - E_f\Big), \tag{2.41}$$

which is known as *maximization procedure*. It is fundamentally a formal extension of the principle of insufficient reason [1] that was used as a guiding principle for

---

[1] "*The principle of insufficient reason states that probability assignments are based on a symmetry in our*

inference at the birth time of probability theory (Uffink, 1995). Jaynes approach is based on the SK axioms that determine entropy uniquely as information measure. Indeed, since the requirements stated by Shannon are specified according to the quantification of the information content of the messages, which is entropy, Shore and Johnson explore some new axioms, for which it was claimed that they do not have in themselves any reference to a particular information measure. According to Shore and Johnson, all reasonable methods for revising a probability distribution based on some posterior evidence must lead to a consistent result when different ways of taking the posterior evidence into account exist. This intuitively reasonable dictum was stated as four consistency axioms, which are as follows (Shore & Johnson, 1980):

3.1 *Uniqueness* (SJ1): The result must be unique.

3.2 *Invariance* (SJ2): The choice of coordinate system must not change the result.

3.3 *System Independence* (SJ3): If two systems are independent, individual information measures of them are the same for marginal and joint probability distributions of the systems.

3.4 *Subset Independence* (SJ4): The treatment of any independent subset of the system states does not make any difference whether it is taken in terms of separated conditional probability distributions or in terms of joint probability distributions.

The fact that makes SJ axioms special is that they have precedence to the principle of maximum entropy because they focus on the consistency and unbiasedness such that the maximization procedure of any information measure is an issue coming afterward. SJ axioms serve to derive a functional for information measure which in turn is to be maximized. Ignoring this vital step implies the possible failure of the principle of maximum entropy and some posterior arguments about the formal structure of entropy function such as additivity. Furthermore, if a given constraint is on the mean value of an observable quantity, SJ axioms yield Shannon entropy as the unique functional for information measure (Shore & Johnson, 1980). Therefore, the crux of finding a consistent functional for information measure is not only the issue of axioms but also of the relevant constraints.

It is worth noting that it was recently shown that SK and SJ axioms give rise to the same functional form for information measure if $SK4$ is generalized so as to capture systems and subsystems independence which implies that SK and SJ axioms plays

---

judgement, i.e., on the absence of our knowledge that would favour the occurrence of one event above the others" (Uffink, 1995).

the same role on the stage (Jizba & Korbel, 2020). To this aim, the elementary algebraic operations, that is, addition, multiplication, subtraction and division, are firstly extended by means of a strictly increasing bijection $f : M \mapsto N \subset \mathbb{R}$. Having such a bijection, one can extend the elementary operations in M respectively as follows:

$$\begin{aligned} x \oplus y = f(f^{-1}(x) + f^{-1}(y)); x \otimes y = f(f^{-1}(x) f^{-1}(y)) \\ x \ominus y = f(f^{-1}(x) - f^{-1}(y)); x \oslash y = f(f^{-1}(x) / f^{-1}(y)). \end{aligned} \tag{2.42}$$

Generalized arithmetic of real multivariate functions naturally arises in the light of generalizing the elementary operations. For example, a function of two variables, $G(x,y)$, can be extended as $G_f(x,y) := f(G(f^{-1}(x), f^{-1}(y)))$. After these generalizations, $SK4$ axiom can be revised as follows (Jizba & Korbel, 2020):

Definition 2.2.2 (Composability (SK_g4)) *Joint entropy of two physically observables A and B can be expressed as $H(A,B) = H(A \mid B) \otimes_f H(B)$ such that $H(A,B)$ and the conditional entropy $H(A \mid B)$ should satisfy the following requirements for consistency:*

i. For two independent physical observables, or random variables, A and B, the joint entropy should be composed of the individual entropies $H(A)$ and $H(B)$, i.e., $H(A,B) = F(H(A), H(B))$.

ii. Conditional entropy can be decomposed into entropies of conditional distributions, i.e., $H(A \mid B) = G(\{p_i^A, H(B \mid A = a_i)\}_{i=1}^n)$.

It was shown that SK axioms with the generalized form of SK_g4 give rise to the same function for information measure as SJ axioms such that the function has the form $H_q^f(p) = f(\left(\sum_i p_i^q\right)^{1/(1-q)})$. That SK and SJ axioms yield the same functional form is indeed true if one adds the maximality condition to SJ axioms which states that the uniform distribution is the solution for posterior distribution in the case of having no prior information.

On the other hand, $SJ3$ and $SJ4$ axioms are disputable. Because they are about the independence of systems which are special cases reduced from the general systems that could have interrelation. In addition, they are about the state of the observer's knowledge based on the experimental data of the systems that can be considered independent of the real relation between the systems (Uffink, 1995). Indeed, assuming that subsystems or two systems are not independent, could one still claim that the functional form of information measure is again the same as that of independent subsystems or systems? For example, as was stated before, Boltzmann-Gibbs entropy is restricted to reversible processes. Accordingly, how could one extend it to irreversible processes? A successful and celebrated attempt during the second half of $20th$ century was taken by adding an extra term to the orthodox expression, that is, the general entropy was stated as $dS = d_eS + d_iS$, where $d_eS = {đQ}/{T}$ is the orthodox expression of entropy and $d_iS = \sum_k F_k J_k$ is the entropy change caused by the irreversible processes within the relevant system, and F_k and J_k are thermodynamic forces and thermodynamic flows respectively (Kondepudi & Prigogine, 2015). If the independence of subsystems or systems is assumed, $dS = d_eS + d_iS$ reduces to $dS = d_eS$. Having confined with this reduction, it is dubious in principle to claim that the characteristic properties of the reduced entropy, $dS = d_eS$, encompass the characteristic properties of the general entropy, $dS = d_eS + d_iS$. This suggests that the extension of the entropy limited to thermal equilibrium regimes or independent systems to the general scheme is not a trivial matter, even if it is simply by changing the axioms that determine the characteristic properties of entropy. In addition to these sort of questions, it becomes an extra problem how to write the conditional entropy of the subsystems and how to take into account the stochastic effect occurring in the system of inquiry. Addressing to these questions and the existence of some experimental events such as anomalous diffusion (Abe & Thurner, 2005) whose statistics are not explained by H_{BG} based on the maximization procedure in equation (2.41), some new entropy functions were proposed among which *Rényi entropy* ($I_\alpha(p)$) and *Tsallis entropy* ($H_\alpha(p)$) are the most wide known due to the fact that they satisfy many desired physical and informational properties.

Rényi entropy was proposed by Rényi (Reńyi, 1961, 1965) as a function for information measure based on the generalization of mean value estimation while preserving the additive property of information measure:

$$I_\alpha(p) := \frac{1}{1-\alpha}\log_2\left(\sum_i p_i^\alpha\right), \tag{2.43}$$

where $\{p_i\}_{i=1}^n$ is the probabilities of a set of n events of a random variable X, and α is the generalization parameter with the fact that it converges to Shannon entropy for the limit $\alpha \to 1$.

Mean value of a set of data $\{x_i, i = 1,2,\ldots,n\}$ weighted with a probability distribution $\{p_i\}_{i=1}^n$ can be measured in general by means of a continuous and strictly increasing function $f(x)$ as $M(x_1,x_2,\ldots,x_n) := f^{-1}\left(\sum_i p_i f(x_i)\right)$. In information theory, the set of data is the information gain, $\{-\log_2(p_i), i = 1,2,\ldots,n\}$. In contrast to Shannon entropy, mean value is measured exponentially in case of Rényi entropy, that is, with $g(\alpha,p_i) = \frac{p_i^\alpha}{\sum_i p_i^\alpha}$, $f(x) = {(2^{(1-\alpha)x}-1)}/{(1-\alpha)}$ and assuming two random variables X and Y with conditional probabilities $\{p_{i|j}\}_{i=1}^n$ and marginal probabilities

$\{q_j\}_{j=1}^m$ respectively, conditional entropy of random variable X given Y is given by

$$
\begin{aligned}
I_\alpha(X \mid Y) &:= f^{-1}\left(\sum_i g(\alpha,q_i) f\left(I_\alpha\left(X \mid Y = y_i\right)\right)\right) \\
&= \frac{1}{1-\alpha} log_2 \left(\frac{\sum_{i,j} q_i^\alpha p_{i|j}^\alpha}{\sum_k q_k^\alpha}\right) \\
&= I_\alpha(X,Y) - I_\alpha(Y),
\end{aligned}
\tag{2.44}
$$

which reduces to Shannon entropy if the limit $\alpha \to 1$ is taken. If $\alpha \to 1$, $g(1,p_i) = p_i$, which means that mean value is estimated linearly in case of Shannon entropy. Using the exponential average makes the role of probabilities questionable. Indeed, after using the exponential average, one acquires the right to query the role and meaning of the probability of the occurrence of the events; that is an inquest out of the frame of the mathematical formalism of probability theory (Uffink, 1995, 1996).

Rényi entropy satisfies SK axioms while, in contrast to Shannon entropy, it does not satisfy Lesche stability criterion (Lesche, 1982) and violates $SJ3,4$ axioms independent of whether the linear or exponential average is used (Oikonomou & Bagci, 2019). The most important fact is that Rényi entropy induces artificial biases which are not warranted by data.

Tsallis entropy was proposed by Tsallis (Tsallis, 1988) so as to generalize Shannon entropy with the generalization parameter α which has the form

$$
H_\alpha(p) = \frac{1}{1-\alpha}\left(\sum_k p_k^\alpha - 1\right). \tag{2.45}
$$

Rényi entropy and Tsallis entropy have close connection such that the former is a monotonic increasing function of the latter:

$$
I_\alpha(p) = \frac{1}{(1-\alpha)} \log_2(1 + (1-\alpha)H_\alpha). \tag{2.46}
$$

By means of this relation, Tsallis entropy for composite of two random variables X and Y reads $H_\alpha(X,Y) = H_\alpha(X) + H_\alpha(Y \mid X) + (1-\alpha)H_\alpha(X)H_\alpha(Y \mid X)$, which means that Tsallis entropy does not obey $SK4$ axiom. It satisfies Lesche Stability for $\alpha > 0$ (Abe, 2002). Like Rényi entropy, Tsallis entropy also gives rise to spurious biases that are not warranted by data (Pressé et al., 2013). It does also not obey $SJ3$ axiom, which requires that in the absence of coupling between two events x_i and y_j of random variables X and Y respectively, the posterior joint probability of the variables obtained through maximization procedure should satisfy multiplication rule, that is, $p(x_i,y_j) = p(x_i)p(y_j)$.

The stimulative reasons behind proposing the generalized entropies are the beliefs that Shannon entropy does not capture the systems whose constituents have long-range interaction (Plastino & Wedemann, 2020; Rodríguez et al., 2019), and to find a systematic way of how to infer power-law distributions that arise in some problems such as anomalous diffusion. However, first of all, the generalized entropies given above induce biased posterior probabilities that are obtained through maximization procedure, in which case the function of information measure is the relevant generalized entropy rather than Shannon entropy and the mean value constraint might be given by exponential average rather than linear average. Exponential average, which is also known as escort average, is invoked in order to preserve the concavity and additivity of generalized entropies. Secondly, invoking generalized entropies in maximization procedure results in probability distributions that do not have any power of prediction because of that the generalized parameter is specified by posterior data, which in turn is incompatible with Bayesian updating law (Pressé, 2014). Assuming that $P(\alpha, M)$ is a prior generalized probability and D is a set of data gained by experiment, Bayesian updating law then states that the posterior probability $P(\alpha, M \mid D)$ has the following proportionality: $P(\alpha, M \mid D) \propto P(D \mid \alpha, M)P(M \mid \alpha)P(\alpha)$. Here, $P(\alpha)$ has to be determined prior to experiment data D. However, generalized entropies, in general, require the generalized parameter to be informed of the experimental data. In conclusion, we are left to the fact that in order to model a non-exponential probability distribution, such as power laws, non-extensivity or the coupling risen from the interaction between events must be expressed in terms of constraints or by a prior probability distribution, not in terms of entropy. Indeed, if a constraint is imposed on the number of events, such as $\sum_n p_n \ln(n) = \chi$, one can obtain a power law through the maximization procedure in case of Shannon entropy (Visser, 2013).

Finally, as a candidate for information measure *von Neumann entropy* is also used, which is formally the counterpart of Shannon entropy in quantum information theory. Having the quantum state ρ of the relevant system, von Neumann entropy reads $S(\rho) = -\operatorname{tr}(\rho \ln(\rho))$.

In the scope of this thesis, we mainly studied Shannon entropy ($H_S(p)$) and von Neumann entropy ($S(\rho)$) to express entropic uncertainties, not the generalized entropies because of that, in contrast to the former entropies, the latter entropies do not satisfy fully either SK axioms or SJ axioms (Pressé et al., 2013; Oikonomou & Bagci, 2019), which serve to draw a reasonable and consistent frame for information measure.

3. ENTROPIC UNCERTAINTIES IN QUANTUM MEASUREMENTS

The notion of uncertainty is not special to physics. For example, in his famous work in the field of communication theory Shannon preferred to talk about the measure of uncertainty about the transmitted messages, which was expressed by the function of entropy (Shannon, 1948). Shannon's idea addresses to the state of humankind's knowledge about the system under consideration; that is an uncertainty which can be always remedied in principle by means of some observation whatsoever. However, Heisenberg uncertainty principle has radically changed our view of nature ever since it was stated by Heisenberg as a principle in quantum mechanics (Heisenberg, 1927).

Monumental improvements in physics have usually resulted from a philosophical perspective on its foundation. Interpreting the motion of physical systems in phase-space and the idea of (the principle of) action has given rise to analytical mechanics. A strict and robust examination of both measurement and observation led A. Einstein to a new comprehension of space-time, and thus, to the theory of special relativity; and in contrast to the prevalent view at that time, assuming the disintegration of energy discreetly led M. Planck to reach his formula of black body radiation, which is also the origin of quantum theory. In a similar manner, after Heisenberg and Schrödinger published their works (Heisenberg, 1927, 1925; Schrödinger, 1928) on the dynamics of quantum systems, a debate about the physical content of Schrödinger wave function arose among physicists, which again had a philosophical character and still maintains its vitality. Also related to this debate and after the field of quantum information and computation emerged, many studies have been dedicated to giving a reasonable and satisfactory explanation to quantum mechanics, especially to its fundamental features, such the quantum state of a quantum system (Pusey et al., 2012; Lewis et al., 2012; Fuchs et al., 2014; Combes et al., 2018; Frauchiger & Renner, 2018), the measurement process (Moreira et al., 2018), and most importantly the uncertainty principle (Bialynicki-Birula & Mycielski, 1975; Deutsch, 1983; Maassen & Uffink, 1988; Ozawa, 2003; Bush et al., 2013). Heisenberg's uncertainty principle is expressed in two version, one in terms of position and momentum observable and the other in terms of energy and time. It states that it is impossible to determine with certainty the position and momentum of a physical system simultaneously; or, to express in terms of energy and time, it is impossible to determine with certainty the energy of a physical system at an instant of time. In the literature, since the works on uncertainty principle have been placed from an operational perspective (Hilgevoord & Uffink, 2016), they have unavoidably given rise to either a measuremental interpretation (Heisenberg, 1927; Ozawa, 2003; Bush et al., 2013) or statistical interpretation of it (Wigner, 1963; Margenau, 1963).

The concept of uncertainty is presented in several different meanings in the physical literature. It may mean to a lack of knowledge of a physical quantity that is to be observed by an observer (Fuchs et al., 2014), or to the experimental inaccuracy with which the quantity is measured (Ozawa, 2003; Bush et al., 2013), or to a statistical spread in the preparation of an ensemble of a particular system (Margenau, 1963). In the operational perspective of Heisenberg, two complementary observables *"can be determined simultaneously only with a* ***characteristic indeterminacy****. This indeterminacy is the real basis for the occurrence of statistical relations in quantum mechanics"* (Wheeler & Zurek, 1983, p. 1) [emphasis added]. Operational perspective states that only those quantities that are in principle observable should play a role in the understanding of the considered theory. For example, it is principally meaningful to mention about the orbit of an atom in the classical realm, whereas, as was argued by Heisenberg (Heisenberg, 1927), it should be avoided from quantum mechanics since it is impossible to observe the position and momentum of an electron simultaneously which are required to define the orbit. To be more precise, Heisenberg argued that *"when one wants to be clear about what is to be understood by the words 'position of the object', for example of the electron (relative to a given frame of reference), then one must specify definite experiments with whose help one plans to measure the 'position of the electron'; otherwise this word has no meaning"* (Wheeler & Zurek, 1983, p. 64). After formulating the basic physical quantities, such as position and momentum, in the form of matrices (Heisenberg, 1925), Heisenberg postulated that the matrices X and P representing the canonical position and momentum quantities of a particle obey the so-called canonical commutation rule

$$[X,P] := XP - PX = \mathrm{i}\hbar, \tag{3.1}$$

where $\hbar = \frac{h}{2\pi}$, h is Planck's constant. A first attempt for giving a physical interpretation of this commutation relations was explored by Heisenberg (Heisenberg, 1927) after Schrödinger published his works (Schrödinger, 1928) on wave mechanics in which Schrödinger argued that his approach is more perceptual, or physical (anschaulichen), than matrix mechanics, which was developed primarily by Heisenberg

Born and Jordan (Born & Jordan, 1925; Born et al., 1926). For Schrödinger, the perceptual understanding of a phenomenon refers to having a space-time picture of the phenomenon since a complete physical description of a system can be provided by a continuous determination of position and momentum quantities, which give rise to a space-time picture. However, for Heisenberg, gaining a perceptual understanding of physical theory is succeeded if in all basic cases, we can grasp the experimental results qualitatively and be sure that the theory does not give rise to any contradiction. We wish to give his idea by his own word:

> "We believe we understand the physical content (anschaulichen inhalt) of a theory when we can see its qualitative experimental consequences in all simple cases and when at the same time we have checked that the application of the theory never contains inner contradictions. For example, we believe that we understand the physical content of Einstein's concept of a closed 3-dimensional space because we can *visualize consistently the experimental consequences of this concept*. Of course these consequences contradict our everyday physical concepts of space and time. However, we can convince ourselves that the possibility of employing usual space-time concepts at cosmological distances can be justified *neither by logic nor by observations*. The physical interpretation of quantum mechanics is still full of internal discrepancies, which show themselves in arguments about continuity versus discontinuity and particle versus wave" (Wheeler & Zurek, 1983, p. 1) [emphasis added]

Therefore, for Heisenberg, experiment is underlying the perceptual understanding of a theory. In accordance with his idea, Heisenberg interpreted the canonical commutation rule of position and momentum in terms of their deviations, $\Delta(X)\Delta(P) \geq h$, with the aid of a thought experiment he considered. The experiment is to measure the position of an electron by illuminating it with light and using a microscope. When a monochromatic light is sent on the electron, light reflects from the electron and comes toward the microscope, which has an aperture angle θ. We have depicted this event in Figure 3.1.

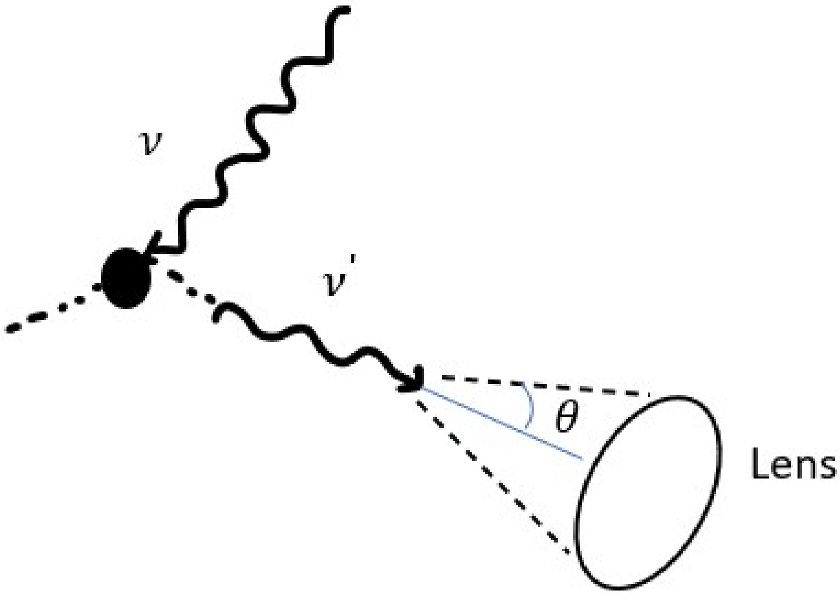


Figure 3.1 Schematic picture of electron-photon collision for determining the position of an electron. A photon with frequency ν hits the electron and scatters with frequency ν' toward the microscope that has an aperture angle θ, meanwhile the electron suffers a recoil. The higher frequency does the photon has, the more precise is the position of the electron determined. To this aim, γ-ray must be used, which in turn causes a discontinuous change in the momentum value of the electron, which is known as Compton effect, that cannot be ignored. Eventually, the more accurately is position ascertained, the more deviation occurs in momentum value.

At the moment when the position is determined, while at the same time the photon is scattered by the electron, the electron undergoes an discontinuous change in momentum. This change increases by decreasing the wavelength of the photon, which is required for a precise determination of the position. Conversely, if we do not want to causes a significant change in the momentum of the electron, we need to use a monochromatic light with relatively larger wavelength. The scattered light of wavelength λ enters the microscope of aperture angle θ. According to the law of optics, the accuracy of the microscope depends on both λ and θ, which is stated by Abe's criterion for its resolving power, that is, the size of the smallest deviation $\Delta(X)$ of the position is about $\Delta(X) \approx \frac{\lambda}{\sin(\theta)}$. On the other hand, the direction of the entering photon is unknown within the angle θ, rendering a change $\Delta(P)$ in the momentum of the electron by an amount $\Delta(P) \approx \frac{h \sin(\theta)}{\lambda}$. Therefore, $\Delta(X)\Delta(P) \approx h$, which was firstly formulated by Heisenberg (Heisenberg, 1927). We can state Heisenberg's idea of uncertainty as follows: "Essentially, experiments and only experiments can serve to provide a determination of the physical quantities, and they are subjected to irreducible indeterminacies[1]."

The first mathematical formulation of Heisenberg uncertainty principle was given by Kennard (Kennard, 1927). He proved that for a normalized quantum state $|\psi\rangle$, the following inequality holds:

$$\Delta(X)\Delta(P) \geq \frac{\hbar}{2}, \tag{3.2}$$

where $\Delta(X)$ and $\Delta(P)$ are deviations of position and momentum in state $|\psi\rangle$, that is,

$$(\Delta(X))^2 = \langle\psi| X^2 |\psi\rangle - (\langle\psi| X |\psi\rangle)^2 ; (\Delta(P))^2 = \langle\psi| P^2 |\psi\rangle - (\langle\psi| P |\psi\rangle)^2 . \tag{3.3}$$

The inequality (3.2) was later generalized by Robertson for all hermitian operators:

$$\Delta(A)\Delta(B) \geq \frac{1}{2} | \langle\psi|[A,B] |\psi\rangle|. \tag{3.4}$$

There are experimental results, some of which support the uncertainty (Nairz et al., 2002; Nikolic & Nesic, 2011; Qu et al., 2021), while some violate it (Sulyok et al., 2013; Rozema et al., 2012). Schrödinger improved the inequality further by considering the correlation between the observables (Schrödinger, 1930) as follows:

$$\Delta(A)\Delta(B) \geq \frac{1}{2}\sqrt{(\langle\psi|\{A,B\}|\psi\rangle - \langle\psi|A|\psi\rangle \langle\psi|B|\psi\rangle)^2 + | \langle\psi|[A,B] |\psi\rangle|}. \tag{3.5}$$

Position and momentum are two complementary observables, that is, they are two observables corresponding to two MUBs. Heisenberg uncertainty principle puts a limit on obtaining information content of a quantum system. The observables corresponding to MUBs cannot be determined simultaneously with certainty; the more precisely one of such observables is determined, the more uncertain the other must be.

In the formulation of Heisenberg uncertainty, there is no reference to the resolving power of the experimental apparatus and the effect of the measurement device on the conjugate variable. The variances given in equation (3.4) are based on the usage of many copies of the quantum state under consideration. However, in reality, we need to also take the imprecision and effect of the devices on the conjugate quantity into account for a reasonable and applicable purpose. To put it clearly, there is no measurement procedure in quantum mechanics, just as was claimed by Heisenberg,

[1] This idea is fraught with vicious circle: According to this idea, we should test an irreducible indeterminacy of an experiment by performing that experiment.

by which we can determine accurately one of the complementary observables without disturbing the other; that is, the inaccuracy in one of them and the disturbance in the other cannot together be arbitrarily small in a particular experiment. To this aim, Ozawa (Ozawa, 2003) and Bush *et. al.* (Bush et al., 2013) reformulated the uncertainty relation based on this consideration. In Ozawa's approach, we consider the measurement device and the system together which are to interact at the instant of measurement. The system with the quantum state $|\phi\rangle$ interacts with the device having the state $|E_0\rangle$. Their interaction is governed by a unitary operation U in accordance with the quantum operation formalism. The observable A of the system that are to be measured can be expressed as $A_{in} = A \otimes I$ on the joint Hilbert space of the system and the device. We express the action of reading of the device by a pointer observable M, and $M_{in} := I \otimes M$ on the joint space before measurement. After the interaction, the pointer change as $M_{in} \mapsto M_{out} = U^{\dagger}(I \otimes M_{in})U$ and $A_{in} \mapsto A_{out} = U^{\dagger}(A \otimes I)U$. Based on this notational conventions, the inaccuracies in the measurement of, for instance, position X was given as

$$\epsilon(X, |\phi\rangle) = \Big(\langle \phi \otimes E_0| \, (M_{out} - X_{in})^2 \, |E_0 \otimes \phi\rangle\Big)^{1/2}, \tag{3.6}$$

and the disturbance of momentum P that is caused by the device which has been used for measuring the position was proposed as (Ozawa, 2003)

$$\eta(P, |\phi\rangle, X) = \Big(\langle \phi \otimes E_0| \, (P_{out} - P_{in})^2 \, |E_0 \otimes \phi\rangle\Big)^{1/2}. \tag{3.7}$$

In order to show that this formulation implies the inequality (3.4), Ozawa considered the expressions $M_{out} := X_{in} + N(X)$ and $P_{out} := P_{in} + D(P)$, where $N(X)$ is the noise operator and $D(P)$ the disturbance operator such that $[M_{out}, P_{out}] = 0$ since they are observables in different systems. Hence, we have

$$[N(X), D(P)] + [N(X), P_{in}] + [X_{in}, D(P)] = -[X_{in}, P_{in}] = -[X, P], \tag{3.8}$$

from this last expression, Ozawa proposed the following inequality

$$\begin{aligned} \epsilon(X, |\phi\rangle)\eta(P, |\phi\rangle, X) + \frac{|\langle [N(X), P_{in}] \rangle + \langle [X_{in}, D(P)] \rangle|}{2} \\ \geq \frac{|\langle \phi| \, [X_{in}, P_{in}] \, |\phi\rangle|}{2} = \frac{|\langle \phi| \, [X, P] \, |\phi\rangle|}{2}, \end{aligned} \tag{3.9}$$

which reduces to the inequality (3.4) if second term on the left side vanishes. There are some experimental works (Erhart et al., 2012) that supports this inequality.

The approach of Bush *et. al.* is again to formulate the slogan "no measurement without disturbance", which is underlying the physical content of Heisenberg uncertainty principle, and to put forward the statement that Heisenberg uncertainty principle is still correct in spite of the conflicting result obtained in the reference (Rozema et al., 2012), where the violation of the uncertainty relation in equation (3.4) was argued experimentally. The authors have argued that Heisenberg's discussion of uncertainty presented in his work (Heisenberg, 1927) is not covered by the inequality in equation (3.4), because the momentum disturbance in Heisenberg's discussion apparently involves the comparison of the momentum before measurement and the momentum after measurement. In addition, they argued that the inequality (3.4) is just a quantitative measure of the fact that *"there are no dispersion-free quantum states"* (Bush et al., 2013) since the observables are measured in different experiment on the distinct copies of a particular quantum state. They have considered a measurement device which performs an joint unsharp measurement of both position and momentum. Using the formalism of quantum operations, they assume a set of POVMs $\mathcal{N} = \{N(x,p)\}$ for joint measurement such that

$$
\begin{aligned}
&N(x,p) \geq 0 \quad \text{and} \iint dxdpN(x,p) = I \\
&N_X(x) = \int dpN(x,p) \quad \text{and} \quad N_P(p) = \int dxN(x,p).
\end{aligned} \tag{3.10}
$$

For a system having a quantum state $|\phi\rangle$, the joint distribution for the joint outcomes (p,x) is

$$
\mathcal{P}(x,p) := \langle\phi| N(x,p) |\phi\rangle \tag{3.11}
$$

such that

$$
\begin{aligned}
\mu_1(x) &:= \int dp\mathcal{P}(x,p) = \langle\phi| N_X(x) |\phi\rangle \\
\nu_1(p) &:= \int dx\mathcal{P}(x,p) = \langle\phi| N_P(p) |\phi\rangle .
\end{aligned} \tag{3.12}
$$

The distributions defined in equation (3.12) represent a real scenario of measurement that are unsharp. On the other hand, the ideal measurements in the quantum formalism are represented by von Neumann measurement. The corresponding ideal distributions of the realistic distributions of equation(3.12) then are

$$
\mu_0(x) := |\langle x|\phi\rangle|^2 \quad \text{and} \quad \nu_0(p) := |\langle p|\phi\rangle|^2. \tag{3.13}
$$

Bush *et. al.* proposed a distance function D between the realistic probability distributions and the ideal. The distance they proposed is Wasserstein-2 distance, which can be stated as follows:

Definition 3.0.1 *Let $\mu_1(x)$ and $\mu_0(y)$ be two marginal probability distributions of* *a joint probability distribution $p(x,y)$. Then, Wasserstein-2 distance is defined as*

$$D(\mu_1,\mu_0) := \inf_{p(x,y)} \left(\iint (x-y)^2 p(x,y) dx dy \right)^{\frac{1}{2}}. \tag{3.14}$$

Applying this definition to above distributions for position and momentum, and writing the measured and ideal values of position and momentum as (X_1, X_0) and (P_1, P_0) respectively, they have defined the inaccuracy of position and momentum as the supremum over all possible input states $|\phi\rangle$:

$$\Delta(X_0,X_1) = \sup_{|\phi\rangle} D(\mu_0,\mu_1) \quad \text{and} \quad \Delta(P_0,P_1) = \sup_{|\phi\rangle} D(\nu_0,\nu_1). \tag{3.15}$$

They have obtained from this statements their final result as

$$\Delta(X_0,X_1)\Delta(P_0,P_1) \geq \frac{\hbar}{2}. \tag{3.16}$$

Inequality of (3.16) has been experimentally supported (Xiong et al., 2017). We emphasize that both statements of uncertainty in equations (3.9) and (3.16) suffer from a vicious circle. Because they are statements about the measurements devices which are to be tested by measurement devices. The expression of uncertainty principle in terms of deviations was also formulated as the sum of them (Mondal et al., 2017); from a quantum operational perspective (Renes et al., 2017) and based on the median of position and momentum (Bera et al., 2019). A detailed analysis of Heisenberg uncertainty principle has been explored in (Busch et al., 2007).

However, as firstly highlighted by Deutsch (Deutsch, 1983), these formulations of uncertainty based on deviation have some drawbacks; for example, lower bound of the uncertainty principle, $\Delta(A)\Delta(B) \geq \frac{1}{2}|\langle\psi|[A,B]|\psi\rangle|$, depends on the initial state $|\psi\rangle$, and thus, is not fixed such that it can vanish for some choices of $|\psi\rangle$ which do not have to be simultaneous eigenfunctions of the observables A and B. In addition, deviation-based uncertainty relations do not capture in general the physical content of the complementary aspect (Dammeier et al., 2015), and the spread of informational content (Bialynicki-Birula & Rudnicki, 2011), of the observables. Due to these reasons, which we shall explore in detail below, uncertainty was formulated in terms of entropy function.

Expressing uncertainty in terms of entropies of the observables was first set forth as a question by Everett (Everett, 1957). It was answered affirmatively by Hirschman (Hirschman, 1957) such that the sum of entropies of position and momentum observables satisfies an inequality. This *Entropic Uncertainty Relation* (EUR) was proved and improved respectively in Refs.(Beckner, 1975; Bialynicki-Birula & Mycielski, 1975) for the observables of having a continuous spectrum. The lower bound of the inequality is achieved when the state of the system is a Gaussian wave-packet. The extension of EUR to the observables in a finite dimensional Hilbert space was first presented by Deutsch (Deutsch, 1983), and improved later by Maassen and Uffink (Maassen & Uffink, 1988). The importance of EUR is that it does not have the aforementioned drawbacks of the uncertainty relations based on deviations. We wish to explore some drawbacks of deviation-based uncertainty principles in order to see the necessity of a new formulation of uncertainty principle.

3.1 Problematic nature of Heisenberg uncertainty principle

We note that at least one of the following problems are going to arise in the other formulations of uncertainty principle that are based on deviation of the observables such as those in the references (Ozawa, 2003; Bush et al., 2013; Mondal et al., 2017; Inoue & Ozawa, 2020).

Example 1. First of all, let us assume spin-1/2 observables, $S_x = \frac{\hbar}{2}\sigma_x, S_y = \frac{\hbar}{2}\sigma_y$ and $S_z = \frac{\hbar}{2}\sigma_z$. According to inequality (3.4),

$$\Delta(S_x)\Delta(S_y) \geq \frac{\hbar}{2}|\langle\psi|[S_x,S_y]|\psi\rangle| = \frac{\hbar}{2}|\langle\psi|S_z|\psi\rangle|. \tag{3.17}$$

Now, if we choose the state $|\psi\rangle = \frac{1}{\sqrt{2}}(1,1)$, which is an eigenvector of S_x, but not of S_y, the left side of the inequality (3.17) becomes zero, which implies that it would be possible to find some pure quantum state by which we can represent our knowledge of S_x and S_y simultaneously, that is, there would be some quantum state that is the eigenvector of both of S_x and S_y. This is however not true.

Example 2. We now consider the deviations of S_x and S_z:

$$\Delta(S_x)\Delta(S_z) \geq \frac{\hbar}{2}|\langle\psi|[S_x,S_z]|\psi\rangle| = \frac{\hbar}{2}|\langle\psi|S_y|\psi\rangle|. \tag{3.18}$$

Now, let us assume that the quantum state would be

$$|\psi\rangle = \frac{1}{2}(1+e^{\mathrm{i}\theta})|0\rangle - \frac{\mathrm{i}}{2}(1-e^{\mathrm{i}\theta})|1\rangle \tag{3.19}$$

in the computational basis $\{|0\rangle, |1\rangle\}$. $|\psi\rangle$ is neither the eigenvector of S_x nor of S_z in general. A straightforward calculation yields that $\Delta(S_x) = \frac{\hbar}{2}\cos(\theta)$ and $\Delta(S_z) = \frac{\hbar}{2}\sin(\theta)$, whose product gives $\Delta(S_x)\Delta(S_z) = \left(\frac{\hbar}{2}\right)^2 \frac{1}{2}\sin(2\theta)$. The product is zero for $2\theta = n\pi, \quad n \in \mathbb{N}$. Moreover and surprisingly, $\langle\psi|S_y|\psi\rangle = 0$. Therefore, it would be possible to find some quantum state for which the deviations could vanish. This shows us that in a finite dimensional Hilbert space one can find some quantum state $|\psi\rangle$ for which we can ascertain some complementary, or mutually unbiased, observables simultaneously with certainty. This is however not true.

Example 3. In this third example, we wish to give a classical example, which has been explored in the reference (Bialynicki-Birula & Rudnicki, 2011). We consider two scenarios that are depicted in Figure 3.2. In the first scenario in Figure 3.2a, the classical particle is allowed to move in the regions I and IV. In the second scenario in Figure 3.2b, the classical particle is allowed to move in the whole box freely; there are no barriers inside the box. Intuitively and reasonably we expect that the inaccuracy in the determination of the particle's position in the first scenario should be less than that of the second scenario. Let us see if this is true. The probability for finding the particle at a position in the first scenario is

(a) A classical particle's motion is confined in the regions I and IV. Each region has the length of $L/4$.

(b) A classical particle's motion is confined in a box of length L, having four regions.

Figure 3.2 The motions of a classical particle in two different scenarios. In (3.2a), regions II and III have been placed on the regions I and IV respectively and they are forbidden. The particle can be placed, and so thus, move in the regions I and IV. In (3.2b), all of the regions together constitute a box of length L and there are no barriers or partitions between the regions. The particle can move freely in the box.

$$p_a(x) = \begin{cases} \frac{2}{L} & \text{if } x \in [0, \frac{L}{4}] \\ \frac{2}{L} & \text{if } x \in [\frac{3L}{4}, L] \\ 0 & \text{elsewhere.} \end{cases} \tag{3.20}$$

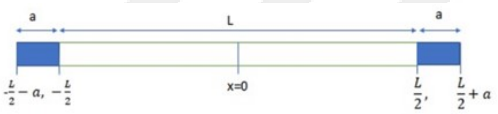
A straightforward calculation of the position deviation yields the result $\Delta_a(X) =$

$\sqrt{\frac{7}{4}}\frac{L}{\sqrt{12}}$. For the second scenario, the corresponding probability is

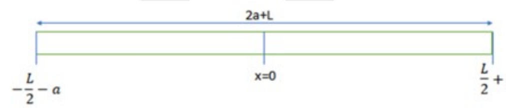
$$p_b(x) = \begin{cases} \frac{1}{L} & \text{if } x \in [0, L] \\ 0 & \text{elsewhere.} \end{cases} \tag{3.21}$$

We estimate the deviation for position of the particle in this case as $\Delta_b(X) = \frac{L}{\sqrt{12}}$. As is seen, $\Delta_a(X) > \Delta_b(X)$, which contradicts our intuition. This shows us that deviation-based uncertainty depends on the variance, which is very sensitive to the tail of probability distribution. In the other words, the uncertainty is very sensitive to labeling the outcomes. Surprisingly, this counter-intuitive fact also arises if we consider a quantum particle.

Example 4. We now consider an electron in two scenarios that are very similar to that of Example 3. In the first scenario, the electron is confined with equal probability in the small boxes with length a attached to the edges of the long box of length L, which has been centered at $x = 0$. In the second scenario, the electron moves inside the box of length $L + 2a$, centered at $x = 0$. We have schematized them in Figure 3.3. We assume that in the first scenario the two small boxes are far



(a) An electron's motion is confined in the boxes of length a, attached to the edges of the box of length L.



(b) An electron's motion is confined in a box of length $L + 2a$. The electron moves freely inside the box.

Figure 3.3 The motions of an electron in two different scenarios. In (3.3a), the electron is allowed to be move in the boxes attached to the edges of a relatively long box of length L. In (3.3b), the electron moves freely in the long box of length $L + 2a$. We do not take other physical quantities such as spin into account that is not affect our conclusion

away from each other enough so that the overlapping of Schrödinger functions in the middle can be ignored. In that case, solving Schrödinger equation for the electron in the first scenario gives

$$\psi_n(x) = \frac{1}{\sqrt{2}}\left(\sqrt{\frac{2}{a}}\sin_I\left(\frac{n\pi}{a}(x + \frac{L}{2})\right) + \sqrt{\frac{2}{a}}\sin_{II}\left(\frac{n\pi}{a}(x - \frac{L}{2})\right)\right) \tag{3.22}$$

for the n^{th} level energy from which we find the variance of the position for electron as

$$(\Delta_a(X))^2 = Var_a(X) = (L+2a)^2 \left(\frac{1}{12} + \frac{L}{6(L+2a)} - \frac{a^2}{2n^2\pi^2(L+2a)^2} \right) \tag{3.23}$$
$$\approx \frac{L^2}{4}, \quad \text{if } L >> a \quad \text{and} \quad n >> 1.$$

In the second scenario, Schrödinger function of n^{th} level energy that are found by solving Schrödinger equation is

$$\psi_n(x) = \sqrt{\frac{2}{L+2a}} \cos\left(\frac{(2n+1)\pi}{L+2a} x \right), \tag{3.24}$$

from which we find the variance of the position for the electron as

$$(\Delta_b(X))^2 = Var_b(X) = (L+2a)^2 \left(\frac{1}{12} - \frac{1}{2\pi^2(2n+1)^2} \right) \tag{3.25}$$
$$\approx \frac{L^2}{12}, \quad \text{if} \quad L >> a \quad n >> 1.$$

Again, $\Delta_a(X) > \Delta_b(X)$, which is unexpected.

Examples (3) and (4) suggest us that deviation-based uncertainties depend on the labeling of the outcomes. This is very absurd since one can reduce uncertainty by simply relabeling the outcomes, or eigenvalues say, of the observables. This is valid for every type of uncertainty relations that are based on the deviations of the observables. We note that the dependence of the variance on labeling also arises if the labeled quantities are finite (Coles et al., 2017). There are also some examples that are to sign the problematic nature of the deviation-based uncertainty relations (Urbanowski, 2020).

3.2 Entropic uncertainty relations

Quantum theory formalism is based on measurements and the corresponding outcomes of the measurements. To be more explicit, let us assume that we have a system with a quantum state ρ. If one wishes to know the value of a physical observables A of the system, they must first perform a measurement of A on the system. Secondly, They need to estimate the probabilities of the possible outcomes of the measurement, which are corresponding to the eigenvalues of A. According to Born's rule, we express the probabilities as $p_k = \mathrm{tr}(\Pi_k \rho)$, where Π_k is the projection operator, or measurement element, that is supported by the eigenspace of the corresponding eigenvalue a_k. However, as is well known, all physical observables are not compatible with each other, that is, we are not able to find a common eigenvector for two observables in general. Therefore, the probability distributions $\{p_{nk} = \mathrm{tr}(\Pi_{nk}\rho)\}_{k=1}^{d}$ and $\{p_{mk} = \mathrm{tr}(\Pi_{mk}\rho)\}_{k=1}^{d}$ that correspond to two different observables A_n and A_m, but for the same quantum state ρ are correlated in general.

The measurement correlations result in restrictions on any entropy function $H(p)$ of the probabilities. When the restrictions are of the form $H(p \mid A_n) + H(p \mid A_m) \geq C > 0$, it is worth considering them as EURs because they do not only avoid the simultaneous vanishing of the individual entropies but also lead to a trade-off relation between the individual entropies, that is, when one increases, the other has to decrease in order to satisfy the inequality. This is the core reasoning of the EURs. We say EURs because there are more than one EUR by choosing different entropies. In the following subsequent sections, we first recover some EURs for positions and momentum, and then, introduce EURs for finite dimensions.

3.2.1 Entropic uncertainty relations for position and momentum

EUR for position and momentum was given firstly by (Bialynicki-Birula & Mycielski, 1975) in terms of Shannon entropy. In their approach they have used a very important theorem, the Sobolev inequality. However, Bialynicki-Birula later improved their results by taking the physical consideration into account and extending the result to Rényi entropy (Bialynicki-Birula, 2006). Firstly, we would like to present Sobolev inequality, and then, EURs of continuous variables in terms of Rényi, Shannon and Tsallis entropies.

Let us consider the function $\psi(\mathbf{r}) \in L^q$ and its Fourier transformation $\phi(\mathbf{k}) \in L^s$,

$$\phi(\mathbf{k}) = \frac{1}{(2\pi\alpha)^{\frac{n}{2}}} \int d^n r \psi(\mathbf{r}) e^{-\frac{\mathbf{i}\mathbf{k}\cdot\mathbf{r}}{\alpha}}, \tag{3.26}$$

in n-dimensional normed space, where $\frac{1}{q} + \frac{1}{s} = 1$ and $\alpha \in \mathbb{R}^+$ such that $\alpha = \hbar$ in quantum theory. The norms in these spaces are defined as

$$\|\psi(\mathbf{r})\|_q = \left(\int d^n r |\psi(\mathbf{r})|^q\right)^{\frac{1}{q}}$$
$$\|\phi(\mathbf{k})\|_s = \left(\int d^n k |\phi(\mathbf{k})|^s\right)^{\frac{1}{s}}. \tag{3.27}$$

Sobolev inequality then is

$$\|\phi(\mathbf{k})\|_s \leq v(q,s)\|\psi(\mathbf{r})\|_q, \quad v(q,s) = \left(\frac{q}{2\pi\hbar}\right)^{\frac{n}{2q}} \left(\frac{s}{2\pi\hbar}\right)^{-\frac{n}{2s}} \quad \text{and} \quad s \geq q. \tag{3.28}$$

We now consider Rényi entropy in order to derive EUR for continuous observables such as position and momentum. We consider 1-dimensional normed space, i.e., $n = 1$. Then, probability densities for position and momentum are defined as the square of Schrödinger function: $P(x) = |\psi(x)|^2$ and $P(p) = |\phi(p)|^2$, for which $q = s = 2$. Ignoring physical considerations and requirements we can take simply $\alpha = 1$ so that x and p are to be regarded as two dimensionless variables. Using Sobolev inequality for $\psi(x)$ and $\phi(p)$ in terms of the corresponding probabilities we have

$$\left(\int_{-\infty}^{\infty} dp (P(p))^a\right)^{\frac{1}{a}} \leq v_1(a,b) \left(\int_{-\infty}^{\infty} dx (P(x))^b\right)^{\frac{1}{b}}, \tag{3.29}$$

where $a = \frac{s}{2}$, $b = \frac{q}{2}$ and

$$v_1(a,b) = \left(\frac{b}{\pi}\right)^{\frac{1}{2b}} \left(\frac{a}{\pi}\right)^{-\frac{1}{2a}}. \tag{3.30}$$

From $\frac{1}{q} + \frac{1}{s}$ we write $\frac{1}{a} + \frac{1}{b} = 2$, which yields $\frac{a}{b} = \frac{a-1}{1-b}$. Since $\psi(x)$ can be also treated as Fourier transformation of $\phi(p)$, Sobolev inequality is also valid by simply exchanging $\psi(x)$ and $\phi(p)$, that is,

$$\|\psi(\mathbf{r})\|_s \leq v(q,s)\|\phi(\mathbf{k})\|_q \tag{3.31}$$

Using this inequality, we obtain

$$\left(\int_{-\infty}^{\infty} dp (P(x))^a\right)^{\frac{1}{a}} \leq v_1(a,b) \left(\int_{-\infty}^{\infty} dp (P(p))^b\right)^{\frac{1}{b}}. \tag{3.32}$$

We will use the equations (3.29) and (3.32) in the following. At this stage, it is easy to obtain an EUR based on Rényi entropy

$$I_a(P(x)) = \frac{1}{1-a} \ln \left(\int_{-\infty}^{\infty} dx P(x)^a\right) \tag{3.33}$$

for two continuous observables. To this aim, if one takes the natural logarithm of both sides of equation (3.29), they achieve to

$$\frac{1}{a}\ln\left(\int_{-\infty}^{\infty} dp(P(p))^a\right) \leq \ln(v_1(a,b)) + \frac{1}{b}\ln\left(\int_{-\infty}^{\infty} dx(P(x))^b\right), \tag{3.34}$$

and multiplying both sides with a and using $\frac{a}{b} = \frac{a-1}{1-b}$, they obtain

$$\begin{aligned} \ln\left(\int_{-\infty}^{\infty} dp(P(p))^a\right) &\leq a\ln(v_1(a,b)) + \frac{a-1}{1-b}\ln\left(\int_{-\infty}^{\infty} dx(P(x))^b\right) \\ \frac{1}{a-1}\ln\left(\int_{-\infty}^{\infty} dp(P(p))^a\right) &\leq \frac{a}{a-1}\ln(v_1(a,b)) + \frac{1}{1-b}\ln\left(\int_{-\infty}^{\infty} dx(P(x))^b\right) \\ \frac{1}{1-a}\ln\left(\int_{-\infty}^{\infty} dp(P(p))^a\right) + \frac{1}{1-b}\ln\left(\int_{-\infty}^{\infty} dx(P(x))^b\right) &\geq -\frac{a}{a-1}\ln(v_1(a,b)) \end{aligned} \tag{3.35}$$

and finally, using the explicit form of $v_1(a,b)$ in equation (3.30), they arrive at the EUR for position and momentum based on Rényi entropy as (Bialynicki-Birula, 2006)

$$I_a(P(p)) + I_b(P(x)) \geq -\frac{1}{2(1-a)}\ln(a) - \frac{1}{2(1-b)}\ln(b) + \ln(\pi). \tag{3.36}$$

By means of equation (2.46), one can write the corresponding EUR of Tsallis entropy (Rajagopal, 1995). If one takes the limits $a \to 1$ and $b \to 1$, they obtain EUR of Shannon entropy,

$$H_S(P) + H_S(X) \geq \ln(\pi e), \tag{3.37}$$

where

$$H_S(P) = -\int_{-\infty}^{\infty} P(p)\ln(P(p))dp \quad \text{and} \quad H_S(X) = -\int_{-\infty}^{\infty} P(x)\ln(P(x))dx. \tag{3.38}$$

The most important fact of equation (3.37) is that it implies Heisenberg uncertainty relation. To see this, we first note that maximization of Shannon entropy of a continuous variable X under the constraints of having a fixed mean value $\overline{x}$, a fixed deviation $\sigma(X)$ of the variable and the normalization condition for the distribution gives rise to the Gaussian distribution,

$$P(x) = \frac{1}{\sqrt{2\pi\sigma(X)^2}}\exp\left(-\frac{(x-\overline{x})^2}{2\sigma(X)^2}\right), \tag{3.39}$$

for which Shannon entropy of the variable takes its maximum value that is equal to $\frac{1}{2}\ln\left(2\pi e\sigma(X)^2\right)$. Accordingly, taking $\sigma(X)$ and $\sigma(P)$ as the deviations of position and momentum respectively, we can write $H_S(X) \leq \frac{1}{2}\ln\left(2\pi e\sigma(X)^2\right)$ and $H_S(P) \leq$

$\frac{1}{2}\ln\Big(2\pi e\sigma(P)^2\Big)$. Using these inequalities and the uncertainty relation of equation (3.37) we obtain Heisenberg uncertainty relation

$$
\begin{aligned}
\ln(2\pi e\sigma(X)\sigma(P)) = & \ln\Big(\sqrt{2\pi e\sigma(X)}\Big) + \ln\Big(\sqrt{2\pi e\sigma(P)}\Big) \\
\geq & H_S(X) + H_S(P) \geq \ln(\pi e) \\
\Rightarrow \sigma(X)\sigma(P) \geq & \frac{1}{2}.
\end{aligned}
\tag{3.40}
$$

In the previous example, we just treated the situation from a very mathematical perspective; we implicitly assumed that the measurement devices have a perfect resolution such that they could detect the ideal values without any inefficiency. However, this is not true because every measurement are performed by a finite resolution such that we always register an interval of the outcomes, not precise real values. To be more realistic and consider the situation in a physical perspective we should handle the individual probabilities with a finite interval. Based on this reasoning, we now consider a realistic situation, following the work of Bialynicki-Birula (Bialynicki-Birula, 2006).

The probabilities of k^{th} and l^{th} outcomes associated with position and momentum of a quantum particle having the corresponding quantum states $|\psi\rangle$ and $|\phi\rangle$ can be expressed as

$$
\begin{aligned}
p_k(\delta x) &= \int_{k\delta x}^{(k+1)\delta x} dx|\langle x|\psi\rangle|^2 = \int_{k\delta x}^{(k+1)\delta x} dx|\psi(x)|^2 \\
p_l(\delta p) &= \int_{l\delta p}^{(l+1)\delta p} dp|\langle p|\phi\rangle|^2 = \int_{l\delta p}^{(l+1)\delta p} dp|\phi(p)|^2
\end{aligned}
\tag{3.41}
$$

respectively, and the integer indices k and l run from $-\infty$ to ∞. We have assumed reasonably that the intervals have equal size. We note that the resolution intervals δx and δp are not about the accuracy of the measuring instruments, but rather about the standard error of the measurement itself. As was pointed out before, if we refer these intervals to the measuring instruments, we come about a vicious circle. $\delta x\delta p$ represents the area of the phase space in which the particle is to be detected. Therefore, we can construe the area as follows: the more accurately one wishes to localize the particle in phase space, the more is the momentum blurring in the phase space.

Based on slicing the probabilities as above, we can write the left and right hand sides of equation (3.29) as

$$
\begin{aligned}
\int_{-\infty}^{\infty} dp(P(p))^a &= \sum_{l=-\infty}^{\infty} \int_{l\delta p}^{(l+1)\delta p} dp(P(p))^a, \\
\int_{-\infty}^{\infty} dx(P(x))^b &= \sum_{k=-\infty}^{\infty} \int_{k\delta x}^{(k+1)\delta x} dx(P(x))^b
\end{aligned} \tag{3.42}
$$

and similarly for both sides of the equation (3.32). We now apply Jensen inequality of integral form to each term in the sums. Jensen inequality states that the value of a convex function at the average point is less than or equal to the average value of the function, and vice versa for concave functions. Any function $f(t) = t^{\alpha}$ with $\alpha > 1$ is a convex function while for $\alpha < 1$ is a concave function. Assuming that $a > 1$ and $b < 1$ in equation (3.42), and applying Jensen inequality to each integral in the sums we obtain

$$
\begin{aligned}
\left(\frac{1}{\delta p}\int_{l\delta p}^{(l+1)\delta p} dpP(p)\right)^a &\leq \frac{1}{\delta p}\int_{l\delta p}^{(l+1)\delta p} dp(P(p))^a \\
\frac{1}{\delta x}\int_{k\delta x}^{(k+1)\delta x} dx(P(x))^b &\leq \left(\frac{1}{\delta x}\int_{k\delta x}^{(k+1)\delta x} dxP(x)\right)^b.
\end{aligned} \tag{3.43}
$$

With the use of the definition of probabilities in equation (3.41) we can put the equalities of (3.42) in inequality forms by means of equation (3.43) as

$$
\begin{aligned}
(\delta p)^{1-a} \sum_{l=-\infty}^{\infty} p_l^a(p) &\leq \int_{-\infty}^{\infty} dp\,(P(p))^a \\
\int_{-\infty}^{\infty} dx\,(P(x))^b &\leq (\delta x)^{1-b} \sum_{k=-\infty}^{\infty} p_k^a(x),
\end{aligned} \tag{3.44}
$$

and combining these inequalities with equation (3.29) lead to

$$
\begin{aligned}
\left((\delta p)^{1-a} \sum_{l=-\infty}^{\infty} p_l^a(p)\right)^{\frac{1}{a}} &\leq v_1(a,b)\left((\delta x)^{1-b} \sum_{k=-\infty}^{\infty} p_k^b(x)\right)^{\frac{1}{b}} \\
\left(\sum_{l=-\infty}^{\infty} p_l^a(p)\right)^{\frac{1}{a}} &\leq \left(\frac{\delta x\delta p}{\pi\hbar}\right)^{\frac{1-b}{b}} (a)^{-\frac{1}{2a}} (b)^{\frac{1}{2b}} \left(\sum_{k=-\infty}^{\infty} p_k^b(x)\right)^{\frac{1}{b}} \\
\left(\sum_{l=-\infty}^{\infty} p_l^a(p)\right)^{\frac{1}{a-1}} &\leq \frac{\delta x\delta p}{\pi\hbar}\frac{b^{\frac{1}{2(1-b)}}}{a^{\frac{1}{2(a-1)}}} \left(\sum_{k=-\infty}^{\infty} p_k^b(x)\right)^{\frac{1}{1-b}} \\
\frac{1}{1-a}\ln\left(\sum_{l=-\infty}^{\infty} p_l^a(p)\right) + \frac{1}{1-b}\ln\left(\sum_{k=-\infty}^{\infty} p_k^b(x)\right) &\geq -\frac{1}{2}\left(\frac{\ln(a)}{1-a} + \frac{\ln(b)}{1-b}\right) - \ln\left(\frac{\delta x\delta p}{\pi\hbar}\right)
\end{aligned} \tag{3.45}
$$

from which we finally state EUR for position and momentum based on Rényi entropy as (Bialynicki-Birula, 2006)

$$I_a(P)+I_b(X) \geq -\frac{1}{2}\left(\frac{\ln(a)}{1-a}+\frac{\ln(b)}{1-b}\right)-\ln\left(\frac{\delta x \delta p}{\pi\hbar}\right). \tag{3.46}$$

This result is true for $a>1$, $b<1$ since we came about this conclusion by using Jensen inequality under these restrictions. However, we obtain the same result for $a<1$ and $b>1$ if we start from the equation (3.32). Therefore, the result is general under the condition $\frac{1}{a}+\frac{1}{b}=2$. Taking the limits $a\to 1$ and $b\to 1$ in equation (3.46) we obtain EUR of Shannon entropy

$$H_S(P)+H_S(X) \geq -\ln\left(\frac{\delta x \delta p}{\pi\hbar}\right), \tag{3.47}$$

and using the equation (2.46) one can obtain the corresponding EUR of Tsallis entropy. There are also many other examples of EURs of different entropies for continuous variables, such as EUR of Tsallis entropy for signal processing associated with fractional Fourier transformation (Guanlei et al., 2021); providing link between deviation-based uncertainty relation and EURs (Hertz & Cerf, 2019); application of EUR of Shannon entropy to generalized Hulthen-Yukawa potential (Ikot et al., 2020) and to a two-dimensional nanoring placing in the combination of a transverse uniform magnetic field and the Aharonov-Bohm flux (Olendski, 2019), and also the extensions of EURs to relative entropies (Floerchinger et al., 2021).

3.2.2 Entropic uncertainty relations in finite dimensions

As was pointed out in section 3.1, the problematic nature of deviation-based uncertainty relations inspired Deutsch (Deutsch, 1983) to introduce a new formulation of the uncertainty principle. Let us consider that we have two observables A_1 and A_2 associated with a system in d-dimensional Hilbert space. We also assume that the observables are non-degenerate, and so thus, each of them has d distinct eigenvalues. In addition, let $\{|1e_k\rangle\}_{k=0}^{d-1}$ and $\{|2e_k\rangle\}_{k=0}^{d-1}$ be their respective eigenvectors corresponding to the sets of their eigenvalues $\{a_{1k}\}_{k=0}^{d-1}$ and $\{a_{2k}\}_{k=0}^{d-1}$. Deutsch's aim was to obtain an irreducible lower bound for the non-commutative observables A_1 and A_2 in the finite dimension d, such as for spin-half operators in $d=2$. In accordance with their idea, Deutsch firstly expressed the EUR verbally as

$$\begin{pmatrix}\text{Uncertainty in the result of} \\ \text{a measurement of } A_1 \text{ and } A_2\end{pmatrix} \geq (\text{An irreducible lower bound}).$$

Assuming that the system has the quantum sate $|\psi\rangle$ and considering the probabilities $p_{1k} = |\langle e_{1k}|\psi\rangle|^2$ and $p_{2k} = |\langle e_{2k}|\psi\rangle|^2$ of obtaining the outcomes a_{1k} and a_{2k} through the measurement of the observables A_1 and A_2 respectively, Deutsch expressed the right side of his idea given above in terms of the summation of Shannon entropy of the observables

$$H_S(p \mid A_1) + H_S(p \mid A_1) = -\sum_{k=0,j=0}^{d-1,d-1} p_{1k}p_{2j} \left(\ln(p_{1k}) + \ln(p_{2j})\right), \tag{3.48}$$

and aimed to give a lower bound of this expression. For a fixed k, the term in the parenthesis of equation (3.48) is in general non-positive and it takes its maximum value when the quantum state $|\psi\rangle$ is in the midway between $|e_{1k}\rangle$ and $|e_{2j}\rangle$ as follows:

$$|\psi\rangle = \frac{1}{\sqrt{2(1+|\langle e_{1k}|e_{2j}\rangle|)}} \left(|e_{1k}\rangle + e^{-\mathrm{i}\arg(\langle e_{1k}|e_{2j}\rangle)}\,|e_{2j}\rangle\right). \tag{3.49}$$

Using this fact we can write in general

$$\begin{aligned} H_S(p \mid A_1) + H_S(p \mid A_2) &\geq -2\sum_{k=0}^{d-1} p_{1k}p_{2k} \ln\left(\frac{1}{2}(1+|\langle e_{1k}|e_{2k}\rangle|)\right) \\ &\geq -2\ln\left(\frac{1}{2}(1+\max_{k,j}\{|\langle e_{1k}|e_{2k}\rangle|\})\right) \\ &= 2\ln\left(\frac{2}{1+\max_{k,j}\{|\langle e_{1k}|e_{2k}\rangle|\}}\right). \end{aligned} \tag{3.50}$$

For notational convention we make the definition $c := \max_{k,j} |\langle e_{1k}|e_{2j}\rangle|$. Kraus conjectured that Deutsch's achievement of EUR could be improved further as (Kraus, 1987)

$$H_S(p \mid A_1) + H_S(p \mid A_2) \geq -2\ln(c), \tag{3.51}$$

which was proved by (Maassen & Uffink, 1988). The advantage of equation (3.51) over the deviation-based uncertainty relation in equation (3.4) is that the right hand side is independent of the system state. Therefore, it yields a nontrivial information about the correlation of the probabilities whenever $c < 1$, which in turn becomes an evidence for the non-commutativity of the observables. Since Rényi entropy is mathematically the generalization of Shannon entropy and satisfies SK axioms, it would be useful to introduce the extension of equation (3.51) to Rényi entropy form.

To this aim, we first introduce Riesz theorem:

Theorem 3.1 (Riesz theorem) *Let* $|v\rangle = (v_1, v_2, \dots, v_N)^T \in \mathbb{C}^N$ *and* U *be a unitary transformation such that* $(T|v\rangle)_j = \sum_k T_{jk} v_k$*, and let* $\kappa = \max_{j,k} \left|U_{jk}\right|$*. Then*

$$\kappa^{\frac{1}{a}} \left[\sum_{j=1}^{N} |(U|v\rangle)_j|^a\right]^{\frac{1}{a}} \leq \kappa^{\frac{1}{b}} \left[\sum_{k=1}^{N} |x_k|^b\right]^{\frac{1}{b}}, \quad 2 \geq b \geq 1 \quad \text{and} \quad \frac{1}{a} + \frac{1}{b} = 1.$$

Recalling our previous considerations for two probabilities, $p_{1k} = |\langle e_{1k}|\psi\rangle|^2$ and $p_{2k} = |\langle e_{2k}|\psi\rangle|^2$, Riesz theorem can be readily used if we choose $v_k = \langle e_{1k}|\psi\rangle$, $U_{jk} = \langle e_{2j}|e_{1k}\rangle$ and $(U|v\rangle)_j = \langle e_{2k}|\psi\rangle$, so that it becomes $p_{1k} = |v_k|^2$ and $p_{2j} = |(U|v\rangle)_j|^2$. we can then restate Riesz theorem in terms of probabilities as

$$c^{\frac{1}{\beta}} \left[\sum_{j=1}^{N} p_{2j}^{\beta}\right]^{\frac{1}{\beta}} \leq c^{\frac{1}{\alpha}} \left[\sum_{k=1}^{N} p_{1k}^{\alpha}\right]^{\frac{1}{\alpha}}, \tag{3.52}$$

where $\alpha = \frac{b}{2}$ and $\beta = \frac{a}{2}$ so that we have $\frac{1}{\alpha} + \frac{1}{\beta} = 2$, from which we read $\frac{\alpha}{\beta} = \frac{\alpha - 1}{1 - \beta}$. By imitating the steps in equation (3.45) we reach to EUR of Rényi entropy

$$I_\alpha(p \mid A_1) + I_\beta(p \mid A_2) \geq -2\ln(c). \tag{3.53}$$

Taking the limits $\alpha \to 1$ and $\beta \to 1$ we obtain EUR of Shannon entropy, and again, using the equation (2.46) the corresponding relation can be obtained for Tsallis entropy. Lower bound of equation (3.53) is in general referred as $q_{MU} := -2\ln(c)$ after Maassen and Uffink. Our first application of equation (3.53) is to MUBs of Pauli operators. For a pair of MUBs in 2-dimensional Hilbert space, $q_{MU} = \ln(2)$. In general, for any pair of MUBs in d-dimensional Hilbert space, $q_{MU} = \ln(d)$. This is quite reasonable because it is what we wish to come about. This tells us that the uncertainty of two mutually unbiased bases cannot be removed completely when considering a particular quantum state $|\psi\rangle$ of the system under consideration. In other words, one cannot code a complete knowledge of two MUBs into one quantum state $|\psi\rangle$. Complete knowledge here means that one knows what exact eigenvalues the corresponding observables have. For example, if one knows that spin-half observable S_z has up-spin with certainty, one can confidently represent this complete knowledge with the quantum state $|0\rangle$. Now, we ask to ourselves: Does one also have an opportunity, even in principle, to know a complete knowledge of the spin-half observable S_x? The inequality (3.53) tells us that it is impossible without the use of further resource, such as using another system as a memory, a phenomenon that are to be mentioned below.

The result explored above becomes general only for Shannon entropy when we consider a general quantum state ρ, rather than a pure one, $|\psi\rangle$. It is not tight even for a general quantum state ρ; for example, if the quantum state is maximally mixed sate, $\rho = \frac{1}{d}I$, the left side of equation (3.53) becomes $2\ln(d)$, which is greater than the lower bound. Addressing this problem, Frank and Lieb (Frank & Lieb, 2012) proved the following result conjectured by Rumin (Rumin, 2011)

$$H_S(p \mid A_1) + H_S(p \mid A_2) \geq q_{MU} + S(\rho), \tag{3.54}$$

where $S(\rho) = -\operatorname{tr}(\rho \ln(\rho))$ is the von Neumann entropy. Now, the bound is tighter than the previous result, and this is for any two observables A_1 and A_2.

The above approaches are based on the notion of two observables with the sets of their corresponding eigenvectors. From the perspective of quantum operations formalism, we need to consider two general measurements $\mathcal{M} = \{M_k\}_{k=1}^{n}$ and $\mathcal{N} = \{N_k\}_{k=1}^{m}$ in a d-dimensional Hilbert space. The probabilities then are $\{p_k = \operatorname{tr}(M_k\rho)\}_{k=1}^{n}$ and $\{q_k = \operatorname{tr}(N_k\rho)\}$ and their entropies, $H(\mathcal{M} \mid \rho)$ and $H(\mathcal{N} \mid \rho)$ for any entropy function. Let us define the norm $\|A\|_s := \max_{|\psi\rangle} \{|A|\psi\rangle| : \langle\psi|\psi\rangle = 1\}$. Then, using Rényi entropy in this case for a general quantum state ρ, Rastegin found the following EUR (Rastegin, 2010)

$$I_\alpha(\mathcal{M} \mid \rho) + I_\alpha(\mathcal{N} \mid \rho) \geq -2\ln\left(g(\mathcal{M},\mathcal{N})\right), \tag{3.55}$$

where $g(\mathcal{M},\mathcal{N}) = \max_{j,k} \|M_j^{\frac{1}{2}} N_k^{\frac{1}{2}}\|_s$ and $\frac{1}{\alpha} + \frac{1}{\beta} = 2$ as usual. This is the generalization of the relation in equation (3.53) to the quantum operation formalism. Indeed, if we take $\mathcal{M}$ and $\mathcal{N}$ as the measurements of two MUBs, we recover the result of equation (3.53). Further efforts have been done to improve these uncertainty relations under special cases (Zozor et al., 2014).

The above results are just for two observables and quantum operations. But, what if we consider more than two observables? This question was addressed by Sanchez-Ruiz (Sánchez-Ruiz, 1993, 1995) and they gave an EUR for N MUBs. In order to explore this result we need to introduce the concept of purity of a probability distribution. Let $\{\Pi_{nk}\}_{k=0}^{d-1}$ be measurement elements of a MUB in d-dimensional Hilbert space, ρ be a general quantum state and $\{p_{nk} = \operatorname{tr}(\Pi_{nk}\rho)\}_{k=0}^{d-1}$ be the corresponding probabilities. Then, we define the purity of the probabilities as $C_n(P_n) := \sum_{k=0}^{d-1} p_{nk}^2$. It has been shown that the summation of the purities of N MUBs obeys the following inequality (Ivanovic, 1992; Wu et al., 2009)

$$\sum_{n=1}^{N} C_n(P_n) \leq \operatorname{tr}\left(\rho^2\right) + \frac{N-1}{d}. \tag{3.56}$$

For $N = d+1$, the right side becomes $\mathrm{tr}\left(\rho^2\right)+1$. Based on this inequality and the equation (3.53) for N MUBs, Sanchez-Ruiz obtained (Sánchez-Ruiz, 1993)

$$\sum_{n=1}^{N} H_S(A_n \mid \rho) \geq (N+1)\ln\left(\frac{N+1}{\mathrm{tr}(\rho^2)+1}\right). \tag{3.57}$$

However, this result is not optimal. There are some other works devoted to improve this lower bound (Sánchez-Ruiz, 1995; Wu et al., 2009; Puchała et al., 2015). Using variational calculus, Sanchez-Ruiz succeeded to find optimal lower bound for an informationally complete set of MUBs in dimension two (Sánchez-Ruiz, 1993):

$$H_S(\sigma_x \mid \rho) + H_S(\sigma_y \mid \rho) + H_S(\sigma_z \mid \rho) \geq \frac{3}{2}\ln(2). \tag{3.58}$$

EUR of Rényi entropy for N MUBs and two SIC-POVMs was obtained by (Rastegin, 2013). They found the EUR of Rényi entropy for N MUBs in d-dimensional Hilbert space as

$$\sum_{n=1}^{N} I_\alpha(A_n) \geq \frac{N\alpha}{2(\alpha-1)}\ln\left(\frac{Nd}{d\,\mathrm{tr}(\rho^2)+N-1}\right), \tag{3.59}$$

where the set $\{A_n\}_{n=1}^{N}$ represents the operators corresponding to N MUBs. In the same work, Rastegin showed that the purity of a SIC-POVM $\mathcal{N}$ with measurement elements $\mathcal{N} = \{N_j\}_{j=1}^{d^2}$ and the corresponding probabilities $\{p_j = \mathrm{tr}(N_j\rho)\}$ obeys the following equality

$$C(\mathcal{N}) = \sum_{j=1}^{d^2} p_j^2 = \frac{\mathrm{tr}\left(\rho^2\right)+1}{d(d+1)}. \tag{3.60}$$

Using this equality and keeping the notation for the SIC-POVM above, Rastegin obtained the EUR of Rényi entropy for a SIC-POVM as

$$I_\alpha(\mathcal{N} \mid \rho) \geq \frac{\alpha}{2(\alpha-1)}\ln\left(\frac{d(d+1)}{\mathrm{tr}(\rho^2)+1}\right) \tag{3.61}$$

for $\alpha \in [2,\infty)$ and the lower bound, that is, the right side is $\ln\left(\frac{d(d+1)}{\mathrm{tr}(\rho^2)+1}\right)$ for $\alpha \in (0,2]$. In contrast to MUBs, we cannot talk about EUR of more than one SIC-POVM because a SIC-POVM is informationally complete; once one knows a complete knowledge of a SIC-POVM, knowing the other is trivial. Therefore, it is meaningless to talk about the uncertainty of more than one SIC-POVM, which has been argued in (Rastegin, 2013). In that case, can we spell out entropic uncertainty relation for a single SIC-POVM? We have a chance to interpret it physically: if we regard each element of a SIC-POVM as an operator, then we could argue that knowing the quantum state of the system is one of the SIC-POVM's element with certainty leaves us uncertain about the knowledge of the other elements. In other words, the information content of a SIC-POVM cannot be represented by a single pure quantum state.

So far, we have considered EURs of MUBs and SIC-POVM such that the principle system is not correlated with another system. Let us assume a correlation between the principal system and another secondary system, and an agent whose is going to perform the measurement of two MUBs on the principle system. Can the agent determine both MUBs with certainty if they have access to the knowledge of the secondary system? This is an interesting question that was firstly studied by Berta *et. al.* (Berta et al., 2010). In that case, EUR is illustrated in the frame of a game played by two fictitious characters called Alice and Bob. Bob initially prepares a correlated bipartite state ρ_{BA} of two particles A and B, and sends particle A to Alice while he holds the other as a quantum memory for himself. They agree on the observables, say A_1 and A_2, that will be measured on particle A by Alice. Alice performs her measurement and tells Bob which observables she has measured but not the outcome of her measurement. The goal of the game is for Bob to guess correctly the outcome of Alice's measurement by means of his quantum memory particle. Entropic uncertainty for the observables in the presence of quantum memory was obtained by Berta *et al.* (Berta et al., 2010) as

$$S(A_2 \mid B) + S(A_2 \mid B) \geq -\log_2(c) + S(A \mid B), \tag{3.62}$$

where $S(A_s \mid B) = S(\rho_{BA_s}) - S(\rho_A) = -tr(\rho_{BA_s} log_2(\rho_{BA_s})) + tr(\rho_B log_2(\rho_B))$ is the conditional von Neumann entropy of the post-measurement state $\rho_{BA_s} = \sum_k (|sk\rangle \langle sk| \otimes I) \rho_{BA} (|sk\rangle \langle sk| \otimes I)$ for $s \in \{1, 2\}$, and also $S(A \mid B)$ is conditional von Neumann for the initial state ρ_{BA}. This reformulation is generally called *Quantum-Memory-Assisted Entropic Uncertainty Relation* (QMA-EUR) which is stronger than Maassen and Uffink's uncertainty relation (Maassen & Uffink, 1988), and its most remarkable result is that Bob could guess with certainty if ρ_{BA} is maximally entangled state.

For instance, let us assume that Bob has prepared the *Bell state* $\left|\phi^+\right\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B)$, and Bob and Alice have agreed on the measurements of S_Z and S_X that are going to be performed by Alice. Accordingly, $S(A \mid B) = -log_2(2)$. Then, the right side of equation (3.62) becomes $\log_2(2) - \log_2(2) = 0$, which means that Bob will be able to guess the outcomes of the measurement performed by Alice with certainty. QMA-EUR and its revised versions were experimentally verified (Prevedel et al., 2011; Li et al., 2011; Bergh & Gärttner, 2021; Sponar et al., 2021) and many other experiments based on neutron optics were proposed for testing them (Demirel et al., 2020). QMA-EUR has found room in diverse application areas such as quantum noisy channel (Pourkarimi et al., 2020), entanglement witness (Li et al., 2011), quantum steering (Uola et al., 2020) and EUR of correlated position and momentum (Furrer et al., 2014). As is seen from equation (3.62), the lower bound of QMA-EUR is dependent on the quantum memory system in terms of the conditional entropy $S(A \mid B)$, which in turn can be expressed in terms of mutual information $I(A:B) = S(A) - S(A \mid B)$. Based on this equality, further efforts has been made to tighten the lower bound in equation (3.62) by having recourse to mutual information and Holevo quantity (Adabi et al., 2016; Huang et al., 2018; Ming et al., 2020). The reformulation of QMA-EUR given by (Huang et al., 2018) reveals that the effect of quantum memory on the entropic uncertainty completely reflects into the lower bound when the lower bound is presented in terms of Holevo quantity. Recasting the lower bound into mutual information and Holevo quantity provides one to investigate the dynamics of QMA-EUR, such as the effect of Hawking radiation (Huang et al., 2018; Ming et al., 2019) and of the neutrino oscillation (Wang et al., 2020) on the quantum correlation between the system of interest and the memory system. In addition, extension of QMA-EUR to tripartite system was also worked out (Dolatkhah et al., 2020). A review of EURs and their applications have been collected in (Coles et al., 2017).

In addition to EUR, the upper bound of EUR is another important concept which puts an upper bound on the summation of the entropies of two or more observables which we henceforth abbreviate as *entropic certainty relation* (ECR). While EUR quantifies the lack of information, ECR is related to the correlation between the observables, i.e., it measures our certainty about the observables. Therefore, to be more precise, EUR and ECR cannot be treated on the same grounds. Because, if one would investigate EUR, for example, for the spin-1/2 observables S_x and S_y, they would always find that EUR is satisfied if there is no accessible memory. However, in case of ECR, it is impossible to say that the spin observables are incompatible if, for instance, the state of the system is one of the eigenstate of the spin-1/2 observable S_z. Because, in that case, S_x and S_y are fully uncertain, i.e., ECR does not reveal any correlation. However, we show that this fact has an advantage for searching the existence of mutually unbiased bases. ECR for the observables set $\{A_n\}_{n=1}^{N}$ is defined as $\sum_n H_S(A_n \mid \rho) \leq f$ for which the upper bound function f is most likely dependent on the dimension (d) of the system, the number (N) of the observables, the state (ρ) of the system and the measurement elements of the observables ($\{|nk\rangle\}$). If such an upper bound is found, mutual information of the observables, which measures the correlation between the observables,

$$I(A_n : Y) := S(\rho_{A_n}) - S(\rho_{A_n|Y}) \tag{3.63}$$

can also be bounded, where Y is a classical (or quantum) memory given its access to the observer and it becomes $S(\rho_{A_n}) = H_S(A_n \mid \rho)$. If the memory Y is classical memory, EUR implies an inequality for conditional entropy, $\sum_n H(A_n) \geq q \Rightarrow \sum_n H(A_n \mid Y) \geq q$, which, by means of equation (3.63), yields directly an inequality for mutual information: $\sum_n I(A_n : Y) \leq \sum_n H(A_n) - q$ [Henceforth, we discard the quantum state ρ in the expression of entropy for notational convention, that is, we shall use $H_S(A_n)$ instead of $H_S(A_n \mid \rho)$, and the same for von Neumann entropy]. This was first presented by Hall (Hall, 1995) for bipartite case and is called *information exclusion relation.* The relation between mutual information and conditional entropy, $I(A_n : B) = S(A_n) - S(A_n \mid B)$, shows that the decrease of our uncertainty about an observable A_n due to the access to a quantum memory B equals to the increase of our certainty about the observable due to our knowledge of the memory. This fact immediately implies that dynamical behaviour of QMA-EUR is completely reflect into mutual information. Therefore, the dynamics of the entanglement between the memory and measurement systems can be also investigated in the context of mutual information (Fuentes-Schuller & Mann, 2005). Indeed, QMA-EUR in equation (3.62) can be directly put in the form of information exclusion relation:

$$I(A_1 : B) + I(A_2 : B) \leq S(A_1) + S(A_2) + \log_2(c) - S(A \mid B), \tag{3.64}$$

where $S(A_s)$ is to be Shannon entropy $H_S(A_s)$ for $s \in \{1,2\}$. For instance, one can conclude that Hawking radiation results in an decrease of the upper bound of information exclusion relation since it gives rise to an increase of the lower bound of QMA-EUR (Huang et al., 2018; Feng et al., 2015). Therefore, it is not ECR that captures the correlation of the observables exhaustively, but information exclusion relation as the counterpart of EUR. As is seen from equation (3.64), the upper bound of the summation of marginal entropies $S(A_s)$, that is ECR, plays a crucial role in information exclusion relation. As we show, in contrast to EUR, ECR can also be used as a criterion in searching the existence of more than three MUBs especially when the dimension of the system is not a power of a prime number. To give an example, the existence of more than three MUBs in six-dimensional Hilbert space can be numerically studied based on the criterion. The extendibility of MUBs is one of the most important question in quantum information theory. We will return to this point in Section.3.3.

Some Optimal upper bound of entropic uncertainty relations (ECRs) have been presented but they are not optimal (Sánchez-Ruiz, 1993, 1995; Puchała et al., 2015) and some of them are valid only for pure states (Puchała et al., 2015). In addition, they have not been applied yet to any physical problem. We obtain optimal ECR of the measurements performed by N MUBs for some state and it becomes valid for general state when $N = d+1$. We also give some applications of our result. Our method is based on the variational calculus with some conditions satisfied by the probability distributions. Since a valid entropy function is a concave function, one of the extremum values of its argument found by means of variational calculus gives indeed its optimal upper bound (Lanczos, 1986).

In the literature, there is not a physically meaningful interpretation of the EUR for SIC-POVMs although some effort has been devoted to find them (Rastegin, 2013). However, as we pointed out above, EUR for a SIC-POVM is meaningful. Likewise, it would be also meaningful to argue about the upper bound of the entropy of a SIC-POVM. Since a SIC-POVM $\mathcal{N}$ has d^2 measurement elements in d-dimensional Hilbert space, the maximum value of the entropy could be at most $H_S(\mathcal{N}) = 2\ln(d)$. If we find a maximum less than this value, then we could argue that all the elements of a SIC-POVM cannot be fully uncertain simultaneously. We indeed found such a bound in our studies.

In this section, we shall present optimal upper bound of EUR for MUBs and SIC-POVMs [2]. We start with N MUBs $\{|nk\rangle\}_{n=1,k=1}^{N,d}$ in Hilbert space H^d, which may be considered as eigenvectors of the observables $\{A_n\}_{n=1}^{N}$. These observables A_n are known as complementary, or mutually exclusive, observables. If there are $(d+1)$ MUBs, we reconstruct the state ρ of a system with the aid of measurement outcomes of the observables as $\rho = \sum_{n=1,k=1}^{d+1,d} p_{nk}\Pi_{nk} - I$, where Π_{nk} is the projection operator onto the eigenspace of the eigenvector $|nk\rangle$ of the observable A_n, and p_{nk} ($=\mathrm{tr}(\Pi_{nk}\rho)$) is the probability of obtaining the corresponding eigenvalue through measurement (Ivanovic, 1981). The relation between the elements of two MUBs can be then rewritten as $\mathrm{tr}(\Pi_{nk}\Pi_{ml}) = \frac{1+(d\delta_{kl}-1)\delta_{nm}}{d}$. The set of probability distributions

[2] We have not studied here the EUR of Rényi and Tsallis entropies since applying our results to them is a trivial problem.

$\{p_{nk}, n = 1,2,\ldots,N; k = 1,2,\ldots,d\}$ of N MUBs obeys the algebraic relation,

$$\sum_{n=1,k=1}^{N,d} p_{nk}^2 \leq \mathrm{tr}\left(\rho^2\right) + \xi, \tag{3.65}$$

which was obtained in (Ivanovic, 1992) with $\xi = 1, N = d+1$, and in (Wu et al., 2009) with $\xi = \frac{N-1}{d}$ independently. Hence, the inequality in equation (3.65) is a restriction on the summation of the purities of N mutually exclusive observables, and the equality is achieved when N is $d+1$. When the summation of entropies of N observables is maximized, this condition on purities has to be taken into account. The optimization of the summation of the purities in equation (3.65) was used in Ref.(Puchała et al., 2015) in order to obtain lower and upper bounds of entropic uncertainty relation of N observables for pure states. Optimal ECR for N MUBs can be obtained if the inequality (3.65) and the summation of probability to unity are considered in the maximization of the entropy-summation of the observables. In (Sánchez-Ruiz, 1993, 1995), Sánchez-Ruiz found ECR for $(d+1)$ MUBs, with the aid of the assumption that the purities of the observables corresponding to MUBs are constant independently. We first extend the equality in equation (3.65) to N MUBs for some state, and then, take it as a necessary condition on the probability distributions; thus, in contrast to Refs.(Sánchez-Ruiz, 1993, 1995), the purities of the observables are considered dependent on each other. The intuitive reason behind our consideration can be seen from the following scheme. If one assumes the probability distribution of an observable as $\{p_{n1} = 1, p_{n2} = p_{n3} = \cdots = p_{nd} = 0\}$, then the probability distributions of the rest observables become equally likely as $\{p_{s1} = p_{s2} = \cdots = p_{sd} = 1/d; s = 1,2,\ldots,n-1,n+1,\ldots,N\}$, which implies that the purities of the observables corresponding to N MUBs are dependent on each other.

Proposition 3.1 *Let $\{|nk\rangle, k = 1,2,\ldots,d\}$ be the orthonormal basis of the observable A_n in Hilbert space H^d. Then, for the states $\rho = \sum_{n=1,k=1}^{N,d} \lambda_{nk} |nk\rangle\langle nk|$, the summation of the purities of N observables is $\sum_{n=1}^{N} C_n := \sum_{n=1,k=1}^{N,d} p_{nk}^2 = \mathrm{tr}\left(\rho^2\right) + \frac{N-1}{d}$.*

We now prove this proposition. When the dimension of the relevant system is a power of a prime number, the expression $\rho = \sum_{n=1,k=1}^{N,d} \lambda_{nk} |nk\rangle\langle nk|$ is valid for any states that can be expanded in terms of N mutual unbiased bases such that $1 \leq N \leq d+1$, because in this case, there are $(d+1)$ MUBs (Wootters & Fields, 1989). If the dimension is not a power of a prime number, the expression given above for states is still valid at least when $1 \leq N \leq 3$ since we know that there exist at least three MUBs in any finite dimensional Hilbert space (Klappenecker & Rötteler, 2003).

Let us assume that $\rho = \sum_{n=1,k=1}^{N,d} \lambda_{nk} |nk\rangle \langle nk|$. Since $\mathrm{tr}(\rho) = 1$ then $\sum_{n=1,k=1}^{N,d} \lambda_{nk} = 1$.

Furthermore, the trace of the square of the sate leads to

$$\begin{aligned}\mathrm{tr}\left(\rho^2\right) &= \sum_{m,n,k,s} \lambda_{nk}\lambda_{ms}\frac{1+(d\delta_{ks}-1)\delta_{nm}}{d} \\ &= \frac{1}{d}+\sum_{nk}\lambda_{nk}^2-\frac{1}{d}\sum_{n,k,s}\lambda_{nk}\lambda_{ns},\end{aligned} \tag{3.66}$$

and the probabilities are

$$\begin{aligned}p_{nk} &:= \mathrm{tr}(\Pi_{nk}\rho) \\ &= \lambda_{nk}+\frac{1}{d}\sum_{m,l}\lambda_{ml}-\frac{1}{d}\sum_{l}\lambda_{nl}.\end{aligned} \tag{3.67}$$

If we consider the probabilities $\{p_{nk}\}$ and the coefficients $\{\lambda_{nk}\}$ as column vectors $\mathbf{p}=(p_{11},p_{12},\ldots,p_{(N)d})^T$ and $\boldsymbol{\lambda}=(\lambda_{11},\lambda_{12},...,\lambda_{(N)d})^T$ respectively, the relation between them can be written by means of an $Nd\times Nd$ symmetric matrix $\mathbf{T}$ as $\mathbf{p}=\mathbf{T}\boldsymbol{\lambda}$. More explicitly,

$$\begin{pmatrix}\vdots\\ p_{nk}\\ \vdots\end{pmatrix}=\begin{pmatrix}I_d & D_1 & D_2 & \ldots & D_{N-1}\\ D_1 & I_d & D_2 & \ldots & D_{N-1}\\ \vdots & & & \ddots & \\ D_1 & D_2 & \ldots & D_{N-1} & I_d\end{pmatrix}\begin{pmatrix}\vdots\\ \lambda_{nk}\\ \vdots\end{pmatrix}, \tag{3.68}$$

where I_d is $d\times d$ identity matrix and the matrices $\{D_i\}_{i=1}^{N-1}$ are also $d\times d$ matrices such that their entries are $\frac{1}{d}$, that is

$$D_1=\cdots=D_{N-1}=D_d=\frac{1}{d}\begin{pmatrix}1&1&\ldots&1\\ 1&1&\ldots&1\\ & &\vdots& \\ 1&1&\ldots&1\end{pmatrix}. \tag{3.69}$$

It is easily seen that $D_d^2=D_d$. The matrix $\mathbf{T}$ is not invertible which implies that a particular distribution $\mathbf{p}=(p_{11},p_{12},...,p_{(N)d})^T$ is not uniquely determined by the state ρ. The summation of the purities of N complementary observables is equal to the square of the norm of $\mathbf{p}$, $\sum_{n=1}^{N}C_n=\sum_{n=1,k=1}^{N,d}p_{nk}^2=\mathbf{p}^T\mathbf{p}$, where it reads

$$\begin{aligned}\mathbf{p}^T\mathbf{p} &= \boldsymbol{\lambda}^T\mathbf{T}^2\boldsymbol{\lambda} \\ &= \frac{N}{d}+\sum_{n,k}\lambda_{nk}^2-\frac{1}{d}\sum_{n,k,s}\lambda_{nk}\lambda_{ns} \\ &= \frac{N}{d}+\mathrm{tr}\left(\rho^2\right)-\frac{1}{d}=\mathrm{tr}\left(\rho^2\right)+\frac{N-1}{d},\end{aligned} \tag{3.70}$$

that was to be shown. Consequently, this equality of purities proves the aforementioned intuitive reasoning of the fact that purities of the observables are dependent on each other. Therefore, the equality has to be taken into account when maximized the summation of the entropies. We obtain the optimal ECR for N MUBs under the following conditions satisfied by the probability distributions of the associated observables

$$\sum_{k=1}^{d} p_{nk} = 1 \tag{3.71}$$

$$\sum_{n=1,k=1}^{N,d} p_{nk}^2 = \mathrm{tr}\left(\rho^2\right) + \frac{N-1}{d}, \tag{3.72}$$

and under the assumption that *the state ρ can be expressed in terms of N MUBs under consideration. For $N = d+1$, the expression of the state in terms of N MUBs is general, and in turn, the following results become true for any state.* We will henceforth abbreviate the trace of the square of the state as $\Pi := \mathrm{tr}\left(\rho^2\right)$. Our method is based on the variation of the function

$$K[\{A_n\}] := \sum_{n=1}^{N} H_S(A_n) = -\sum_{n=1,k=1}^{N,d} p_{nk} \ln p_{nk},$$

where $H_S(A_n)$ is Shannon entropy of the observable A_n. Maximization of the function $K[\{A_n\}]$ under the conditions given above is equivalent to the maximization of the following function

$$\begin{aligned}\Omega(\{p_{nk}\}) :=& -\sum_{n=1,k=1}^{N,d} p_{nk} \ln p_{nk} \\ & -\lambda\Big(\sum_{n=1,k=1}^{N,d} p_{nk}^2 - \Pi - \frac{N-1}{d}\Big) - \beta\Big(\sum_{k=1}^{d} p_{nk} - 1\Big),\end{aligned} \tag{3.73}$$

where λ and β are Lagrange multipliers. Variation of Ω-function reads

$$\delta\Omega = \sum_{k=1}^{d}\left(-\sum_{n=1}^{N} \ln p_{nk} - 2\lambda \sum_{n=1}^{N} p_{nk} - (\beta + N)\right)\delta p_{nk} = 0$$

so that the following equality must be satisfied for all p_{nk}'s, where none of them can be zero,

$$\sum_{n=1}^{N} \ln p_{nk} + 2\lambda \sum_{n=1}^{N} p_{nk} + (\beta + N) = 0, \quad k = 1, 2, \ldots, d. \tag{3.74}$$

Without losing generality, we choose the probabilities set $\{p_{nd} = b_n, p_{nk} = t_{nk} b_n, k =$

$1,2,\ldots,d-1; n=1,2,\ldots,N\}$. Substituting these probabilities into equation (3.74), we obtain two equations

$$\sum_{n=1}^{N} \ln b_n + 2\lambda \sum_{n=1}^{N} b_n = -(\beta+N) \quad for \quad k=d, \tag{3.75}$$

$$\sum_{n=1}^{N} \ln t_{nk} + \sum_{n=1}^{N} \ln b_n + 2\lambda \sum_{n=1}^{N} t_{nk} b_n = -(\beta+N) \quad for \quad k=1,2,\ldots,d-1. \tag{3.76}$$

Substituting $-(\beta+N)$ of equation (3.75) into equation (3.76), we obtain the following equality

$$\frac{\sum_{n=1}^{N} \ln t_{nk}}{\sum_{n=1}^{N} (t_{nk}-1)\, b_n} = -2\lambda; \quad k=1,2,\ldots,d-1. \tag{3.77}$$

The right hand side of equation (3.77) is a constant number for every $k=1,2,\ldots,d-1$, so that the parameter t_{nk} must be independent of index-k, that is, $t_{n1}=t_{n2}=\cdots=t_{n(d-1)}=\frac{1-b_n}{(d-1)b_n}$. Consequently, we obtain the probability distributions as $\{p_{nd}=b_n, p_{nk}=\frac{1-b_n}{d-1}, k=1,2,\ldots,d-1; n=1,2,\ldots,N\}$. According to these distributions, the summation of the entropies is

$$H_T(\{b_n\}) := S[\{A_n\}] = -\sum_{n=1}^{N} b_n \ln b_n - \sum_{n=1}^{N} (1-b_n) \ln\left(\frac{1-b_n}{d-1}\right) \tag{3.78}$$

with the condition

$$\sum_{n=1}^{N} \left(d b_n^2 - 2b_n\right) = \frac{(d-1)\,[d(\Pi+1)-(d+1)]-N}{d}, \tag{3.79}$$

which is the revision of the condition in equation (3.72), since we could not eliminate this condition at the end of the maximization of the function $K[\{A_n\}]$. To find the extremum values of the function $H_T(\{b_n\})$, we define similarly another function as

$$\begin{aligned} \Psi(\{b_n\}) :=& -\sum_{n=1}^{N} b_n \ln b_n - \sum_{n=1}^{N} (1-b_n) \ln\left(\frac{1-b_n}{d-1}\right) \\ &-\mu \left(\sum_{n=1}^{N} \left(d b_n^2 - 2b_n\right) - \frac{(d-1)\,[d(\Pi+1)-(d+1)]-N}{d}\right). \end{aligned} \tag{3.80}$$

The variation of Ψ function reads

$$\sum_{n=1}^{N} \left(\ln\left(\frac{1-b_n}{(d-1)b_n}\right) - 2\mu(d b_n - 1)\right) \delta b_n = 0 \tag{3.81}$$

Since the infinitesimals $\{\delta b_n\}$ are arbitrary, the coefficients must be zero

$$\ln\left(\frac{1-b_n}{(d-1)b_n}\right) - 2\mu(db_n - 1) = 0 \Rightarrow \frac{\ln\left(\frac{1-b_n}{(d-1)b_n}\right)}{db_n - 1} = 2\mu; n = 1,2,\ldots,N. \tag{3.82}$$

The left hand side of equation (3.82) is constant, so that the parameters b_n must be independent of index-n, that is, $b_1 = b_2 = \cdots = b_N$. Bearing in mind this fact, we obtain b_n from equation (3.79) as

$$b_n^{\pm} = \frac{\sqrt{N} \pm \sqrt{(d-1)\left[d(\Pi+1)-(d+1)\right]}}{d\sqrt{N}}; d \geq \left\lceil \frac{d+1}{\Pi+1} \right\rceil, \tag{3.83}$$

where $\lceil . \rceil$ is the ceiling function. The condition on the dimension d in equation (3.83) comes from the fact that the term $\sqrt{(d-1)\left[d(\Pi+1)-(d+1)\right]}$ must be non-negative real number. The value b_n^+ gives the optimal upper bound of the total entropy H_T. Making the abbreviation $\alpha := \sqrt{(d-1)\left[d(\Pi+1)-(d+1)\right]}$, we obtain ECR for N MUBs as

$$\begin{aligned} H_T \leq H_T^+(N,d,\alpha) &= N\ln\left(\frac{d(d-1)\sqrt{N}}{(d-1)\sqrt{N}-\alpha}\right) \\ &- \frac{N+\sqrt{N}\alpha}{d}\ln\left(\frac{(d-1)(\sqrt{N}+\alpha)}{(d-1)\sqrt{N}-\alpha}\right). \end{aligned} \tag{3.84}$$

In order for b_n^- to be a positive real number, it requires that

$$\begin{aligned} b_n^- &= \frac{\sqrt{N} - \sqrt{(d-1)\left[d(\Pi+1)-(d+1)\right]}}{d\sqrt{N}} > 0 \\ &\Rightarrow d < \frac{d+1}{\Pi+1} + \frac{N}{(d-1)(\Pi+1)} \leq \frac{d+1}{\Pi+1} + \frac{d}{d-1} \\ &\Rightarrow d \leq \left\lceil \frac{d+1}{\Pi+1} \right\rceil + 1. \end{aligned} \tag{3.85}$$

Assuming that the state would be a pure state, the restriction in equation (3.85) on the dimension d leads to $d \leq \left\lceil \frac{d+1}{2} \right\rceil + 1$, which is true only if $d \in \{2,3,4\}$. This means that b_n^- cannot be a stationary value for the function $S[\{A_n\}]$ but an extremum (Lanczos, 1986). In passing, we emphasize that the upper bound in equation (3.84) is independent of whatever the set of MUBs has been chosen; it is the same for all type of MUBs.

For a general state, our ECR for $N = d+1$ in equation (3.84) is optimal in contrast to the one

$$H_T \leq (d+1)\ln(d) - \frac{(d-1)(d\Pi-1)\ln(d-1)}{d(d-1)} \tag{3.86}$$

given as equation (10) in Ref.(Sánchez-Ruiz, 1995) since Shannon entropy H_T is a concave function so that its extremum value $\{b_n^+\}$ in equation (3.83) gives its maximum value (Lanczos, 1986). However, equation (3.86) can be used to tighten the upper bound of information exclusion relation in equation (3.64) for MUBs in general. In order to obtain true optimal value of the function $S[\{A_n\}]$ in maximization procedure, the equality satisfied by the purities in equation (3.73) has to be taken in account as has been done above. The upper bound of Shannon entropy $H(\{p_{nk}\})$ for N MUBs given in equation (28) of Ref.(Puchała et al., 2015) is

$$-\sum_{n=1,k=1}^{N,d} p_{nk}\ln(p_{nk}) \leq NH(P) - N\ln(d) \tag{3.87}$$

with the probability distribution $P = \{p_1 = \frac{1+(Nd-1)\sqrt{r}}{Nd}, p_i = \frac{1-\sqrt{r}}{Nd}, i = 2,\ldots,Nd-1\}$, where $r = \frac{Nd\mathcal{P}_{min}-1}{Nd-1}$ and $\mathcal{P}_{min}$ is minimum value of the summation of N purities. Without being optimal, ECR in equation (3.87) is valid only for pure states (Puchała et al., 2015) while our result is valid whenever the state can be expressed in terms of N MUBs. Another difference from the inequality in equation (3.87) is that $H_T^+(N,d,\alpha)$ is state-independent for pure states. To conclude, as its main difference, our ECR for MUBs is optimal whenever the state can be expressed in terms of N MUBs and for any state when $N = d+1$.

We have confirmed our result by some numerical estimations. For a pure state ρ in dimension $d = 2$,

$$\rho = \begin{bmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{bmatrix}, \tag{3.88}$$

we can estimate ECR for spin observables (operators) $\{\sigma_X, \sigma_Y, \sigma_Z\}$. In addition, taking the eigenstates of spin operators as columns for constructing the unitary matrices

$$U_z = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, U_x = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, U_y = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}, \tag{3.89}$$

we can calculate the probabilities as $p_{nk} = \langle 1k| U_n^\dagger \rho U_n |1k\rangle$, where $\{|11\rangle = |0\rangle, |12\rangle = |1\rangle\}$ is computational basis. Without losing generality, if we choose $\alpha = \sqrt{r}$ and $\beta = \sqrt{1-r}\exp(i\phi)$, then we obtain the probability distributions of spin observables S_z, S_x, S_y as in Table 3.1.

Table of MUBs and their probabilities, $d = 2$		
S_z	$p_{11} = r$	$p_{12} = 1 - r$
S_x	$p_{21} = \frac{1}{2}(1 + 2\sqrt{r(1-r)}cos(\phi))$	$p_{22} = \frac{1}{2}(1 - 2\sqrt{r(1-r)}cos(\phi))$
S_y	$p_{31} = \frac{1}{2}(1 - 2\sqrt{r(1-r)}sin(\phi))$	$p_{32} = \frac{1}{2}(1 + 2\sqrt{r(1-r)}sin(\phi))$

Table 3.1 The probability distributions table of MUBs in $d = 2$ when the state is pure. The leftmost column stands for MUBs ($S_n, n = z, x, y.$), and the other columns for probabilities of obtaining their first and second eigenvalues, respectively.

Writing total Shannon entropy of the observables ($H_S(\sigma_n), n = z, x, y.$)

$$H_T(r,\phi) := \sum_{n=1}^{3} H_S(\sigma_n) = -\sum_{n=1,k=1}^{3,2} p_{nk} \ln(p_{nk}), \tag{3.90}$$

we can estimate numerically the maximum value of H_T by running over the parameters r and ϕ. The maximum values, as ca be seen in Figure 3.4, is 1.547120, achieving when $r = 0.2113$ and $\phi = \frac{\pi}{4}$, which coincides with the value of ECR ($H_T^+(N,d,\alpha)$) given in equation (3.84).

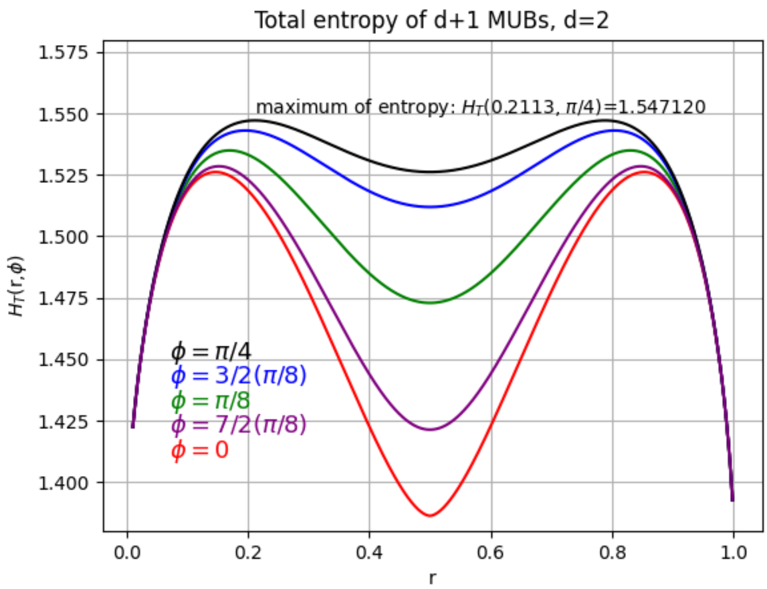


Figure 3.4 Total Shannon entropy of 3 MUBs in equation (3.89) for different values of the parameter ϕ.

We also confirmed our result for $d = 3$, $N = d + 1$. Like in dimension $d = 2$, the general pure state in dimension $d=3$ can be written as follows

$$\rho=\begin{bmatrix}|\alpha|^2 & \alpha\beta^* & \alpha\gamma^*\\ \alpha^*\beta & |\beta|^2 & \beta\gamma^*\\ \alpha^*\gamma & \beta^*\gamma & |\gamma|^2\end{bmatrix}, \tag{3.91}$$

and we choose the unitary matrices

$$\begin{aligned} U_1&=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}, U_2=\begin{bmatrix}\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{3}}&\frac{\omega}{\sqrt{3}}&\frac{\omega^2}{\sqrt{3}}\\ \frac{1}{\sqrt{3}}&\frac{\omega^2}{\sqrt{3}}&\frac{\omega}{\sqrt{3}}\end{bmatrix},\\ U_3&=\begin{bmatrix}\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{3}}\\ \frac{\omega}{\sqrt{3}}&\frac{\omega^2}{\sqrt{3}}&\frac{1}{\sqrt{3}}\\ \frac{\omega}{\sqrt{3}}&\frac{1}{\sqrt{3}}&\frac{\omega^2}{\sqrt{3}}\end{bmatrix}, U_4=\begin{bmatrix}\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{3}}\\ \frac{\omega^2}{\sqrt{3}}&\frac{\omega}{\sqrt{3}}&\frac{1}{\sqrt{3}}\\ \frac{\omega^2}{\sqrt{3}}&\frac{1}{\sqrt{3}}&\frac{\omega}{\sqrt{3}}\end{bmatrix}, \end{aligned} \tag{3.92}$$

of an informationally complete set of MUBs, where $\omega=\exp\left(\frac{2\pi i}{3}\right)$. Then, the probability of obtaining the eigenvalue λ_k of the observable A_n is $p_{nk}=\langle 1k|\,[U_n^\dagger\rho U_n\,|1k\rangle$. Without losing generality, we choose $\alpha=\sqrt{r},\beta=\sqrt{q}\exp(i\phi_1)$ and $\gamma=\sqrt{1-(r+q)}\exp(i\phi_2)$, leading to the probability distributions in Table 3.2:

Table of MUBs and their probabilities, d=3			
A_1	$p_{11}=r$	$p_{12}=q$	$p_{13}=1-(r+q)$
A_2	$p_{21}=\frac{1}{3}(1+2f_{21})$	$p_{22}=\frac{1}{3}(1+2f_{22})$	$p_{23}=\frac{1}{3}(1+2f_{23})$
A_3	$p_{31}=\frac{1}{3}(1+2f_{31})$	$p_{32}=\frac{1}{3}(1+2f_{32})$	$p_{33}=\frac{1}{3}(1+2f_{33})$
A_4	$p_{41}=\frac{1}{3}(1+2f_{41})$	$p_{42}=\frac{1}{3}(1+2f_{42})$	$p_{43}=\frac{1}{3}(1+2f_{43})$

Table 3.2 The probability distributions table of MUBs in d=3 when the state is pure. The leftmost column stands for MUBs ($A_n, n=1,2,3,4.$), and the other columns for probabilities of obtaining their first, second and third eigenvalues, respectively.

The functions f_{nk}'s are as follows:

$$\begin{aligned} &f_{21}=\sqrt{rq}cos(\phi_1)+\sqrt{r(1-(r+q))}cos(\phi_2)+\sqrt{q(1-(r+q))}cos(\phi_1-\phi_2),\\ &f_{22}=f_{21}(r,q,\phi_1-\frac{2\pi}{3},\phi_2+\frac{2\pi}{3}), f_{23}=f_{21}(r,q,\phi_1+\frac{2\pi}{3},\phi_2-\frac{2\pi}{3}),\\ &f_{31}=f_{21}(r,q,\phi_1-\frac{2\pi}{3},\phi_2-\frac{2\pi}{3}), f_{32}=f_{21}(r,q,\phi_1+\frac{2\pi}{3},\phi_2),\\ &f_{33}=f_{21}(r,q,\phi_1,\phi_2+\frac{2\pi}{3}), f_{41}=f_{21}(r,q,\phi_1+\frac{2\pi}{3},\phi_2+\frac{2\pi}{3}),\\ &f_{42}=f_{21}(r,q,\phi_1-\frac{2\pi}{3},\phi_2), f_{43}=f_{21}(r,q,\phi_1,\phi_2-\frac{2\pi}{3}). \end{aligned} \tag{3.93}$$

Like in dimension $d = 2$, the maximum value of total Shannon entropy $H_T(r, q, \phi_1, \phi_2)$ can be estimated, searching over its parameters r, q, ϕ_1 and ϕ_2. We obtained numerically the (maximum) value as ≈ 3.44912, achieving when $r = 0.210, q = 0.395, \phi_1 = \phi_2 = 5.236 \approx \frac{5\pi}{3}$, which almost coincides with the value 3.47025 of $H_T^+(N, d, \alpha)$. These numerical results suggests that ECR in equation (3.84) is optimal.

The most physical significance of ECR is that it arises in searching mutually coherent states, which are related to the existence of MUBs. By definition, $|\psi_{coh}\rangle$ is a mutually coherent state with respect to N MUBs associated with the set of observables $\{A_n\}_{n=1}^N$, iff $\{\mathrm{tr}(\Pi_{nk} |\psi_{coh}\rangle \langle\psi_{coh}|) = \frac{1}{d}, \forall n, k; n = 1, 2, \ldots, N; k = 1, 2, \ldots, d\}$. Even if the existence of 3 MUBs is known (Klappenecker & Rötteler, 2003), whether there are more than three MUBs in non-prime power dimension is still an open question. If $\{|\psi_k\rangle\}_{k=1}^d$ are mutually coherent states with respect to N MUBs, the set of N MUBs can be extended to $(N+1)$ MUBs (Mandayam et al., 2014). Stating in a reverse manner, (i) *if there is no a mutually coherent state* $|\psi_{coh}\rangle$ *with respect to* N *MUBs, this set of* N *MUBs cannot be extended to* $(N+1)$ *MUBs.* It is straightforward to see that in case of the state of the system being either a maximally mixed state or a mutually coherent state (with respect to N MUBs in question), total entropy of N MUBs must achieve its maximum value, that is, $N \ln(d)$. We now wish to show how ECR ($H_T(N, d, \alpha)$) covers this fact. First of all, when the state is a maximally mixed state (i.e., $\rho = \frac{1}{d} I$), the parameter α becomes zero and $H_T^+(N, d, 0) = N \ln(d)$, an extreme case in which we are not interested. Secondly, we assume that the state of the system of inquiry could be written in terms of N MUBs and a mutually coherent state $|\psi_{coh}\rangle$, that is,

$$\rho = \sum_{n=1,k=1}^{N,d} \lambda_{nk} |nk\rangle \langle nk| + r |\psi_{coh}\rangle \langle\psi_{coh}|. \tag{3.94}$$

For such state, the only change in our maximization procedure for total entropy happens to the parameter α and the condition on the validity of b_n^- in equation (3.85) such that $\alpha \mapsto \bar{\alpha} = \sqrt{(d-1)\left[d(\Pi+1) - (d+1) - r^2(d-1)\right]}$ and $d \leq \left\lceil \frac{d+1}{\Pi+1} \right\rceil + 1 \mapsto d \leq \left\lceil \frac{d+1}{\Pi+1} + \frac{r^2(d-1)}{\Pi+1} \right\rceil + 1$. In that case, b_n^- is not valid when $d \geq 6$ for $\sqrt{\frac{d-5}{d-1}} \geq r > 0$ since the condition $d \leq \left\lceil \frac{d+1}{\Pi+1} + \frac{r^2(d-1)}{\Pi+1} \right\rceil + 1$ is not satisfied anymore. In addition, we need to make the revision $H_T^+(N, d, \alpha) \mapsto H_T^+(N, d, \bar{\alpha})$ in equation (3.84). Now, if ρ is any mutually coherent state with respect to N MUBs, it must be $\forall \lambda_{nk} = 0, r = 1$, which makes the parameter $\bar{\alpha} = 0$, and thereby, $H_T^+(N, d, 0)$ reduces to $N \ln(d)$ that was to be shown. This is another justification that ECR in equation (3.84) is indeed optimal. Since $H_T^+(N, d, \alpha) \leq H_T^+(N, d, \bar{\alpha})$, and since the ECR in equation (3.84) is optimal upper bound, we can, in consequence, assert that (ii) *there is no a mutually* *coherent state with respect to N MUBs if the ECR in equation (3.84) cannot be exceeded.* As a result, from the two premises (i) and (ii) above, we make the following inference: *(iii) If ECR for N MUBs in equation (3.84) cannot be exceeded, this set of N MUBs cannot be extended to* $(N+1)$ *MUBs.* This inference sets forth a quantitative criterion for the existence of mutually coherent states, and thus, for the extendibility of MUBs. We emphasis that this inference is clearly valid not only for the states that can be written in terms of N MUBs but for any state. Therefore, ECR in equation (3.84) is not just an incremental but also a substantial improvement of the previous ECRs. To give an example, since the existence of 4 MUBs in 6-dimensional Hilbert space is still a conundrum, this criterion can be used as a numerical ground in order to show the non-existence of fourth MUB. If ECR in equation (3.84) cannot be exceeded for 3 MUBs in six dimensional Hilbert space, there is no fourth MUB. For a simple application, one can first prepare a general pure state of the system, and then, perform the measurement of three MUBs. If the total entropy of the outcomes of three MUBs does not exceed the bound in equation (3.84), one can infer that there are no further MUBs. Since $H_T^+(N,d,\alpha) \leq H_T^+(N,d,\bar{\alpha})$, the criterion also implies that total entropy of N MUBs cannot be between $H_T^+(N,d,\alpha)$ and $N\ln(d)$ if there is no a mutually coherent state, and so thus, further MUBs. Otherwise, a mistake must have happened when performed the measurements. In other words, there is no state which results in a total entropy between $H_T^+(N,d,\alpha)$ and $N\ln(d)$ for N MUBs unless it consists of at least a mutually coherent state as its part like in equation (3.94) . Therefore, if we are sure about performing the measurement correctly, any total entropy of N MUBs that exceeds $H_T^+(N,d,\alpha)$ must be interpreted as an evidence for the existence of a mutually coherent state.

Another direct consequence of ECR in equation (3.84) is to decrease the upper bound of information exclusion relation for MUBs. As was stated before, if the total entropy of the observables set $\{A_n\}_{n=1}^{d+1}$ has a lower bound such as $\sum_n H_S(A_n) \geq q$, the total entropy of the observables, where each of them is conditioned with a classical memory Y, satisfies the inequality $\sum_n H_S(A_n \mid Y) \geq q$, which yields information exclusion relation, $\sum_n I(A_n : Y) \leq \sum_n H_S(A_n) - q$. This means that, if the observables are not compatible (i.e., $q > 0$), it is impossible to construct a deterministic correlation between a classical memory and each of the observables of the set $\{A_n\}_{n=1}^{N}$. For the complementary observables $\{A_n\}_{n=1}^{d+1}$ in d-dimensional Hilbert space, $q = (d+1)\ln\left(\frac{d+1}{\Pi+1}\right)$ (Rastegin, 2013) and using the inequality in equation (3.84) for $(d+1)$ MUBs, we obtain an upper bound on the summation of the mutual information as

$$\sum_{n=1}^{d+1} I(A_n:Y) \leq I_T^{+}(N,d,\alpha) = (d+1)\ln\left(\frac{d(d-1)(\Pi+1)\sqrt{d+1}}{(d+1)\left[(d+1)\sqrt{d+1}-\alpha\right]}\right)$$
$$-\frac{d+1+\sqrt{d+1}\alpha}{d}\ln\left(\frac{(d-1)\sqrt{d+1}+\alpha}{(d-1)\sqrt{d+1}-\alpha}\right). \tag{3.95}$$

This result has a significant implication. Not only the lower bound but also the upper bound of the total entropy for observables makes it impossible to construct a deterministic correlation with classical memory which implies the incompatibility of the observables. Because, if the $(d+1)$ observables were compatible, the upper bound of total mutual information would be $(d+1)\ln(d)-q$, which is greater than $I_T^{+}(N,d,\alpha)$. We have shown that the upper bound of information exclusion relation of the observables is dependent on the ECR for the observables. ECR becomes a part of the upper bound of information exclusion relation together with q irrelevant to the case of having a memory. Therefore, in expression of information exclusion relation in equation (3.64), the term $S(A_1)+S(A_2)+log_2(c)$ in information exclusion relation plays the role of the lower bound q of EUR.

We can also apply the variational method explored above to Shannon entropy of SIC-POVMs. We shall try to find an optimal upper bound of EUR for SIC-POVMs. We are not going to use Rényi and Tsallis entropy, but only Shannon entropy.

Let $\mathcal{N}=\left\{N_i=\frac{1}{d}\left|\phi_i\right\rangle\left\langle\phi_i\right|; i=1,2,...,d^2\right\}$ be a SIC-POVM and ρ be the density matrix of the system under consideration in d-dimensional Hilbert space. We then read the probability distribution as $\{p_i=Tr(N_i\rho), i=1,2,...,d^2\}$. Since the SIC-POVM is considered as an informationally complete measurements, the quantum state can be reconstructed. We have two conditions on the probability distribution of SIC-POVM just as those on that of MUBs (Rastegin, 2013),

$$\sum_{i=1}^{d^2} p_i = 1 \tag{3.96}$$

$$\sum_{i=1}^{d^2} p_i^2 = \frac{\Pi+1}{d(d+1)}; \quad \Pi := \mathrm{tr}\left(\rho^2\right) \tag{3.97}$$

Under these two conditions, we define the following potential function:

$$\zeta := -\sum_{i=1}^{d^2} p_i\ln(p_i) - \gamma\left(\sum_{i=1}^{d^2} p_i - 1\right) - \kappa\left(\sum_{i=1}^{d^2} p_i^2 - \frac{\Pi+1}{d(d+1)}\right) \tag{3.98}$$

The variation of the potential ζ leads to

$$\delta\zeta = \sum_{i=1}^{d^2} \left(\ln(p_i) + 2\kappa p_i + (1+\gamma)\right)\delta p_i = 0$$

Since the variations δp_i's are arbitrary, the coefficient of each δp_i must be zero,

$$\ln(p_i) + 2\kappa p_i + (1+\gamma) = 0; i = 1,2,...,d^2 \tag{3.99}$$

According to equation (3.99), none of p_i can be zero, otherwise the equation cannot be satisfied. We now choose the distribution $\{p_1 = q, p_i = \beta_i q; i = 2,3,...,d^2\}$, from which we obtain two equation by means of the condition in equation (3.99):

$$\ln(q) + 2\kappa q = -(1+\gamma); \quad \text{and} \quad \ln(\beta_i) + \ln(q) + 2\kappa\beta_i q = -(1+\gamma), \quad i = 2,3,...,d^2.$$

Substituting the first equation in the second one, we obtain

$$2\kappa q = -\frac{\ln(\beta_i)}{\beta_i - 1}, i = 2,3,...,d^2, \tag{3.100}$$

The left hand side of equation (3.100) is a constant and the same for each index-$\{i\}$, so that the parameters β_i has to be independent of the index-$\{i\}$, that is, $\beta_2 = \beta_3 = ... = \beta_{d^2}$. Then, we obtain β_i's from the normalization condition in equation 3.96 as $\{\beta_i = (1-q)/(d^2-1); i = 2,3,...,d^2\}$. In addition, using the second condition on the probability distribution in equation (3.97), we obtain two values for the positive real parameter q,

$$q^{\pm} = \frac{1 \pm \sqrt{(d-1)\left[d(\Pi+1) - d - 1\right]}}{d^2}. \tag{3.101}$$

The solution q^- is not a stationary point, but an extremum which is the particular value $1/d^2$ of q^+, corresponding to the pure mixed state, $\rho = \frac{1}{d}I$. Therefore, it cannot be considered as an optimal value in general. Making the conventional abbreviation $\alpha := \sqrt{(d-1)\left[d(\Pi+1) - d - 1\right]}$, the optimal upper bound of the entropy is achieved by choosing $q = q^+$:

$$H \le H^+(\alpha, d) = \frac{1+\alpha}{d^2}\ln\left(\frac{d^2}{1+\alpha}\right) + \frac{d^2-1-\alpha}{d^2}\ln\left(\frac{d^2(d^2-1)}{d^2-1-\alpha}\right). \tag{3.102}$$

We note that ECR of equation (3.102) is independent of whatever SIC-POVM has been chosen. We can check if the equation (3.102) is indeed optimal. To this aim, we take the SIC-POVM $\mathcal{N} = \left\{ N_i = \frac{1}{2} |\phi_i\rangle \langle\phi_i| ; i = 1, 2, ..., d^2 \right\}$ as follows:

$$
\begin{aligned}
&N_1 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, N_2 = \frac{1}{6}\begin{pmatrix} 1 & \frac{2}{\sqrt{2}} \\ \frac{2}{\sqrt{2}} & 2 \end{pmatrix}, \\
&N_3 = \frac{1}{6}\begin{pmatrix} 1 & \frac{2e^{-\frac{2\pi i}{3}}}{\sqrt{2}} \\ \frac{2e^{\frac{2\pi i}{3}}}{\sqrt{2}} & 2 \end{pmatrix}, N_4 = \frac{1}{6}\begin{pmatrix} 1 & \frac{2e^{\frac{2\pi i}{3}}}{\sqrt{2}} \\ \frac{2e^{-\frac{2\pi i}{3}}}{\sqrt{2}} & 2 \end{pmatrix},
\end{aligned} \tag{3.103}
$$

and considering the pure quantum state in equation (3.91), the corresponding probabilities are

$$
\begin{aligned}
p_1 &= \frac{1}{2}r, \quad p_2 = \frac{1}{6}\left(2 - r + 2\sqrt{2r(1-r)}\cos(\phi)\right) \\
p_3 &= \frac{1}{6}\left(2 - r - \sqrt{2r(1-r)}\left(\cos(\phi) + \sqrt{3}\sin(\phi)\right)\right) \\
p_4 &= \frac{1}{6}\left(2 - r - \sqrt{2r(1-r)}\left(\cos(\phi) - \sqrt{3}\sin(\phi)\right)\right).
\end{aligned} \tag{3.104}
$$

The maximum value of Shannon entropy of the SIC-POVM $\mathcal{N}$ can be numerically estimated as ≈ 1.242 approximately that can be seen in Figure 3.5.

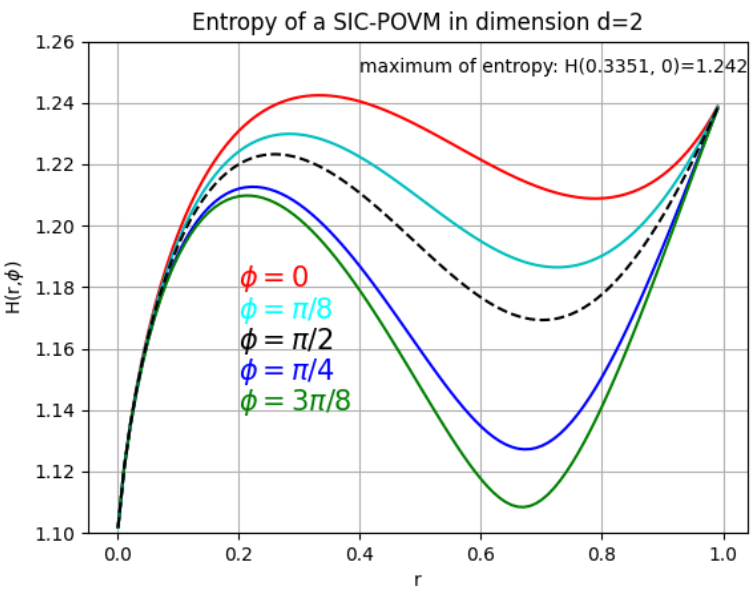


Figure 3.5 Shannon entropy of the SIC-POVM $\mathcal{N}$ in equation (3.103) for different values of the parameter ϕ.

The theoretical value exactly coincides with the numerical value. Like the question we raised as to whether EUR for SIC-POVM is meaningful, we also concern with the question of whether ECR for SIC-POVMs is significant. The duality between knowledge and ignorance, or certainty and uncertainty, reflects into the relation between EUR and ECR, for that we are able to consider ECR for SIC-POVMs. In accordance with this reflection, we argue that since the upper bound in equation (3.102) is less than $\ln\left(d^2\right)$ with the exception $\rho = \frac{1}{d}I$, the measurement elements of SIC-POVMs are correlated to each other. This correlation comes from the constant value of the inner product of the measurement elements of SIC-POVMs.

As we noted before, if there is a SIC-POVM in a d-dimensional Hilbert space, we have more than one. Let us see how this is possible. We consider that we have the SIC-POVM $\mathcal{N} = \{N_i = \frac{1}{d}|\psi_i\rangle\langle\psi_i|\}_{i=1}^{d^2}$. Now, we take the complex conjugation of the vectors $|\psi_i\rangle$, and express it as $|\psi_i^\star\rangle$. For example, the complex conjugation of the vector $|v\rangle = \frac{1}{\sqrt{2}}(1,\mathrm{i})^T$ is $|v^\star\rangle = \frac{1}{\sqrt{2}}(1,-\mathrm{i})^T$. The characteristic condition on the SIC-POVM's elements, as was presented before in Section 2.1.2, is $|\langle\psi_i|\psi_j\rangle|^2 = \langle\psi_i|\psi_j\rangle\langle\psi_j|\psi_i\rangle = \frac{d\delta_{ij}+1}{d+1}$ for all $i,j \in \{1,2,\ldots,d^2\}$. Since $\left\langle\psi_i^\star\middle|\psi_j^\star\right\rangle = \langle\psi_j|\psi_i\rangle$, it is clear that the relation $\left|\left\langle\psi_i^\star\middle|\psi_j^\star\right\rangle\right|^2 = \frac{d\delta_{ij}+1}{d+1}$ is also valid for all $i,j \in \{1,2,\ldots,d^2\}$. Therefore, we have always more than one SIC-POVM provided that we have at least one. Henceforth, we shall call this second SIC-POVM as *star-SIC-POVM*. In that case, classification of the SIC-POVMs is an important issue in searching for their existence, their importance in application and grouping them. We studied two questions related to two more than one SIC-POVM:

1) Can we consider the star-SIC-POVM as the time reversal of its counterpart SIC-POVM?

2) Is there a distinctive feature of SIC-POVMs that cannot be encapsulated in the frame of classical information theory?

When studying the first question we recognized that the concept of the celebrated *Quantum Time Reversal* is not compatible with the time reversal operation in quantum mechanics. We also notice that the information coded in SIC-POVMs cannot be represented in the classical information theory. In the following section, we shall explore our results triggered by these two questions.

3.4 Time reversal and non-classical feature of SIC-POVMs

In quantum mechanics time reversal operation is expressed in general by a unitary operation U and complex conjugation operation $\mathcal{K}$. Let us define $\theta := U\mathcal{K}$. Having a quantum state ρ in a d-dimensional Hilbert space $\mathcal{H}^d$, time reversal operation on

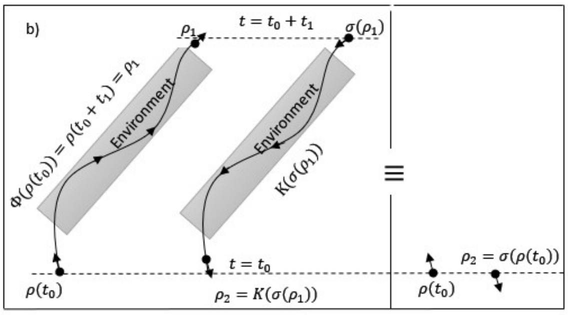


Figure 3.6 Picture of the definition of classical time-reversal symmetry: Left side schematizes the left side of Eg.(3.105), while the right side schematizes the right side of equation (3.105). The (time-reversal) operation K represents Φ of the SIC-POVM.

the space, $\sigma : \mathcal{L}(\mathcal{H}^d) \to \mathcal{L}(\mathcal{H}^d)$, is defined as $\sigma(\rho) := \theta\rho\theta^\dagger$.

Time reversal operation in quantum mechanics is characterized as follows:

> "Time-reversal invariance requires that two different sequences of operations applied to an arbitrary state $|\psi(t_0)\rangle$ [i.e., $\rho(t_0)$] lead to the same state. In the first sequence we allow $|\psi(t_0)\rangle$ [i.e.,$\rho(t_0)$] to evolve for a time t, whereupon we reverse all momenta, and then permit a further evolution for a time t. In the second we merely reverse all momenta in $|\psi(t_0)\rangle$ [i.e., $\rho(t_0)$]. Thus, [by θ to refer to time-reversal operator], we demand that" (Gottfried, 1989, p. 316)

$$e^{-iHt/\hbar}\theta e^{-iHt/\hbar}\rho(t_0)e^{iHt/\hbar}\theta^\dagger e^{iHt/\hbar} = \theta\rho(t_0)\theta^\dagger, \tag{3.105}$$

which must be true for all states of the system of inquiry. We have depicted the two aforementioned sequences in Figure 3.6. This perspective of time reversal operation is just the expression of the classical notion of time reversal operation in quantum mechanics. Therefore, we call it classical time reversal.

We wonder if the star-SIC-POVM can be considered as the time reversal of its corresponding SIC-POVM. To be more explicit, we wish to see if a SIC-POVM is time reversal invariant. With regard to time reversal operation, we first perform a quantum operation of the SIC-POVM $\mathcal{M}_1 = \{M_i = \frac{1}{d}|\psi_i\rangle\langle\psi_i|, i = 1,2,\ldots,d^2\}$, on the state ρ of the system under consideration:

$$\Phi(\rho) := \sum_i dp_i M_i = \rho_1, \tag{3.106}$$

where $p_i = \mathrm{tr}(M_i\rho)$. Then, we take complex conjugation of the resultant state: $\overline{\rho}_1 = \sum_i dp_i\overline{M}_i$, where $\overline{M}_i = \frac{1}{d}\left|\mathcal{K}\psi_i\right\rangle\left\langle\mathcal{K}\psi_i\right| = \frac{1}{d}\left|\psi_i^\star\right\rangle\left\langle\psi_i^\star\right|$. On the other hand, we first apply time reversal operation on the state: $\rho \to \sigma(\rho) = \theta\rho\theta^\dagger = \rho'$. After that, we perform the star-SIC-POVM measurement, $\mathcal{M}_1^\star = \{\overline{M}_i, i = 1, 2, \ldots, d^2\}$, on the new state: $\overline{\Phi}(\rho') := \sum_{j=1} dq_j\overline{M}_i$, where $q_j = Tr(\overline{M}_j\rho')$. Our question is: Are two procedures above equivalent; or to be more formal, is $\overline{\Phi}(\sigma(\rho))$ equal to $\overline{\Phi(\rho)}$? Let us rake together what we did above: In the first place, we simply took the complex conjugation of the resultant state $\Phi(\rho)$, while in the second case, we applied the star-SIC-POVM operation $\overline{\Phi}$ on the time-reversed state $\rho' = \sigma(\rho)$. These two operations exactly comply with the sequences that have been mentioned in the above quotation with the exception that we did just apply the operation $\mathcal{K}$ on the quantum operation Φ to obtain $\overline{\Phi}$ instead of applying time reversal operation $\theta = U\mathcal{K}$. In 2-dimensional Hilbert space, time reversal operator θ is $-\mathrm{i}\sigma_y\mathcal{K}$, where σ_y is Pauli matrix of y-component. We showed for spin-1/2 that $\overline{\Phi}(\rho')$ is transpose of $\overline{\rho}_1$, where we took the SIC-POVM in equation (3.103). However, transpose operation cannot be represented by any accessible quantum operation (Nielsen & Chuang, 2010). Therefore, we cannot link them with the aid of any accessible quantum operation.

We can also apply classical time reversal operation on the quantum operation Φ of the SIC-POVM and see if it is time reversal invariance; that is, we can check if $\Phi \circ \sigma \circ \Phi(\rho) = \sigma(\rho)$. Let us consider the SIC-POVM in equation (3.103) and the general quantum state

$$\rho = \begin{pmatrix} \frac{1+z}{2} & \frac{x-iy}{2} \\ \frac{x+iy}{2} & \frac{1-z}{2} \end{pmatrix}. \tag{3.107}$$

Then, we can define the quantum operation $\Phi(\rho) := \sum_{i=1}^4 A_i\rho A_i^\dagger$ with Kraus operators $A_i = \frac{2}{\sqrt{2}}N_i$. They satisfy the completeness condition: $\sum_{i=1}^4 A_i^\dagger A_i = I$. Recalling time reversal operator $\sigma = -\mathrm{i}\sigma_y\mathcal{K}$, we obtain

$$\Phi \circ \sigma \circ \Phi(\rho) = \begin{pmatrix} \frac{9-z}{18} & \frac{-x+\mathrm{i}y}{18} \\ -\frac{x+\mathrm{i}y}{18} & \frac{9+z}{18} \end{pmatrix}. \tag{3.108}$$

On the other hand, if we apply time reversal operation σ on the state ρ, we obtain

$$\sigma(\rho) = \begin{pmatrix} \frac{1-z}{2} & \frac{-x+\mathrm{i}y}{2} \\ -\frac{x+\mathrm{i}y}{2} & \frac{1+z}{2} \end{pmatrix}. \tag{3.109}$$

As is seen, $\Phi \circ \sigma \circ \Phi(\rho)$ is not equal to $\sigma(\rho)$, which means that quantum operations based on SIC-POVMs are not classical time reversal invariant.

In the classical perspective, we represent the physically accessible processes by Markov transition matrices $R = (r_{ij})_{m\times n}$ such that r_{ij} is the transition probability from state i to state j for a forward evolution chain. Similarly, $\tilde{R} = (\tilde{r}_{ij})_{n\times m}$ is the transition matrix for the reversed chain. We also consider a probability distribution $p^{(e)}$ for the equilibrium of both chains such that $Rp^{(e)} = \tilde{R}p^{(e)}$. In the equilibrium, the probability of the transition $i \mapsto j$ in forward chain is equal to the probability of the transition $j \mapsto i$ in the reversed chain. To express it formally, $\tilde{r}_{ij}p^{(e)}(j) = r_{ji}p^{(e)}(i)$ for all i, j. Putting it in the matrix multiplication, we can rewrite it as

$$\tilde{R} = D(p^{(e)})R^T D^{-1}(p^{(e)}). \tag{3.110}$$

A transition matrix R is *balanced* if $Rp^{(e)} = p^{(e)}$, and is *detailed balance* if it is time reversal invariant, that is, if $\tilde{R} = R$. In line with this classical frame of the processes, and considering a quantum operation $\Phi(\rho) := \sum_{i=}^{m} A_i\rho A_i^\dagger$, Crooks defined time reversal for quantum operations as (Crooks, 2008)

$$\tilde{\Phi}(\rho) := \sum_{i}^{m} \tilde{A}_i\rho\tilde{A}_i^\dagger, \tag{3.111}$$

where $\tilde{A}_i = \rho_0^{1/2}A_i^\dagger\rho_0^{-1/2}$ and ρ_0 is a quantum state such that $\Phi(\rho_0) = \rho_0$. The reason behind this definition was given by Crooks as follows:

> "Since we cannot observe a sequence of states for the quantum dynamics (at least not without measuring and therefore disturbing the system), we instead focus on the sequence of transitions. Each operator of a Kraus operator sum represents a particular interaction with the environment that an external observer could, in principle, measure and record. We can therefore define the dynamical history by the observed sequence of Kraus operators. For each Kraus operator of forward dynamics, A_α, there should be a corresponding operator, $\tilde{A}_\alpha$, of the reversed dynamics such that starting from equilibrium, the probability of observing any sequence of Kraus operators in the forward dynamics is the same as the probability of observing the reversed sequence of reversed operators in the reversed dynamics" (Crooks, 2008, p. 2).

According to Crooks, quantum time reversal must be expressed as the time reversal of the environment in contrast to classical time reversal operation in equation (3.105), in which we seek for time reversal of the considered system, not the environment.

If $\tilde{\Phi} = \Phi$, we say that Φ is in *detailed balance* (Crooks, 2008). This means that if a quantum operation is detailed balance, then it is also time reversal invariant. Indeed, this is the case in stochastic processes. Therefore, we reasonably demand that if a quantum operation is detailed balance it has to be also time reversal invariance. We state our demand as follows:

Corollary 3.1 *If a quantum operation Φ has detailed balance, then it is time reversal invariant.*

In accordance with classical time reversal invariance expressed in equation (3.105), time reversal invariance for quantum operation Φ in literature is defined formally as $\sigma \circ \Phi \circ \sigma = \Phi$ (Fagnola & Umanitá, 2008)[3]. Then, if we combine the expression of time reversal invariance of Crooks and this last one, we state that a quantum operation is time reversal invariant if

$$\sigma \circ \Phi \circ \sigma = \tilde{\Phi}. \tag{3.112}$$

We emphasis that we have brought together two different aspects of time reversal of quantum operations: one aspect, that is $\sigma \circ \Phi \circ \sigma$, has been stated based on classical time reversal in equation (3.105), while other aspect was inspired by the time reversal invariance of stochastic processes. However, even if we can state such condition for time reversal invariance, the conceptual content of the aspect are not compatible with each other. As we pointed out above, Crooks focused on time reversal of environment, while classical time reversal is defined over the quantum state of the system under consideration, which is thought of as separated from its environment. According to the definition of time reversal of Crooks, SIC-POVMs are time reversal invariant; however, it is not true if we take the time reversal condition $\sigma \circ \Phi \circ \sigma = \Phi$. The problem is that while the former approach ignores the system itself, the latter ignores the environment. We can remedy the problem by requiring that there should not be made any difference between the system and its environment. Indeed, if we reverse time, then it must be reversed for every part of the universe. Realistically, it is not meaningful to speak of a partial time reversal. Based on this critics, we first revise the definition of the time reversal of Crooks as follows: for a quantum operation $\Phi(\rho) := \sum_{i=}^{m} A_i \rho A_i^\dagger$, we define its quantum-classical time-reversal operation as $\tilde{\Phi}_{qc}(\rho) := \sum_i^m \tilde{A}_i \rho \tilde{A}_i^\dagger$, where $\tilde{A}_i = \rho_0^{1/2} \sigma(A_i^\dagger) \rho_0^{-1/2}$ and ρ_0 is a

[3] From equation (3.105), it must be $\sigma \circ \Phi \circ \sigma = \tilde{\Phi}$.

quantum state such that $\sigma(\rho_0) = \rho_0$ and $\Phi(\rho_0) = \rho_0$. Now, regarding this remedy, we define quantum-classical time-reversal invariance as follow: *If a quantum operation* Φ *is quantum-classical time-reversal invariant, the following statement hold: If we first perform a quantum operation* Φ *on (the state of) a system and then reverse the time, this whole operation is equivalent to firstly reversing time and then performing quantum-classical time-reversal operation* $\tilde{\Phi}_{qc}$ *of the quantum operation* Φ*; to put it in a formal language,* $\sigma \circ \Phi \equiv \tilde{\Phi}_{qc} \circ \sigma$. We have pictured this definition in Figure 3.7.

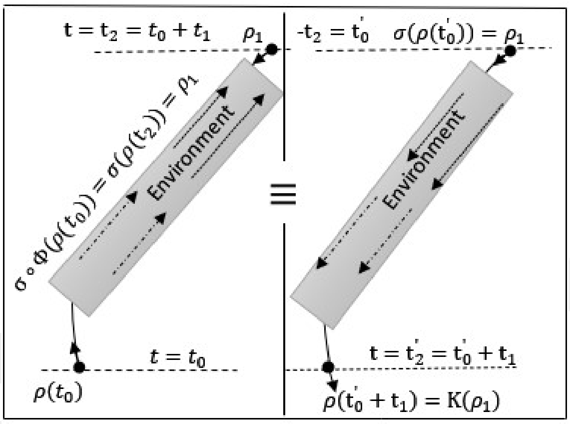


Figure 3.7 The quantum-classical time-reversal operation. The left side represents the operation $\sigma \circ \Phi$, and the right side represents the operation $\tilde{\Phi}_{qc} \circ \sigma$.

We point out that this definition includes classical time-reversal invariance that is given in equation (3.105). If we take the quantum operation as the unitary evolution of the system, $\Phi(\rho) = e^{-itH/\hbar}\rho e^{itH/\hbar}$, its quantum-classical time-reversal operation becomes, $\tilde{\Phi}_{qc}(\rho) = (\theta e^{itH/\hbar}\theta^\dagger)\rho(\theta e^{-itH/\hbar}\theta^\dagger)$. Then, $\theta e^{-itH/\hbar}\rho e^{itH/\hbar}\theta^\dagger = (\theta e^{itH/\hbar}\theta^\dagger)\theta\rho\theta^\dagger(\theta e^{-itH/\hbar}\theta^\dagger)$ if and only if, $\theta e^{itH/\hbar}\theta^\dagger = e^{-itH/\hbar}$, or $\theta H \theta^\dagger = H$, which is exactly what is required.

We show that SIC-POVMs are quantum-classical time-reversal invariant. To this aim, we first note that the equilibrium state for a quantum operation that is defined by a SIC-POVM is maximally mixed state: $\rho_0 = \frac{1}{d}I$. For example, let us consider time dependent quantum state in 2-dimensional Hilbert space,

$$\rho(t) = \begin{pmatrix} \frac{1+z}{2} & \frac{x-iy}{2}e^{-iw_L t} \\ \frac{x+iy}{2}e^{iw_t t} & \frac{1-z}{2} \end{pmatrix} \tag{3.113}$$

and the quantum operation

$$\Phi(\rho) = \sum_i^m A_i \rho A_i^\dagger := \sqrt{2}\sum_i p_i L_i \tag{3.114}$$

whose Kraus operators are defined as $A_i = L_i$ via the measurement elements of the

SIC-POVM

$$L_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} \frac{3+\sqrt{3}}{6} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & \frac{3-\sqrt{3}}{6} \end{pmatrix}, L_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} \frac{3+\sqrt{3}}{6} & -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} & \frac{3-\sqrt{3}}{6} \end{pmatrix}$$
$$L_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} \frac{3-\sqrt{3}}{6} & -\frac{i}{\sqrt{6}} \\ \frac{i}{\sqrt{6}} & \frac{3+\sqrt{3}}{6} \end{pmatrix}, L_4 = \frac{1}{\sqrt{2}}\begin{pmatrix} \frac{3-\sqrt{3}}{6} & \frac{i}{\sqrt{6}} \\ -\frac{i}{\sqrt{6}} & \frac{3+\sqrt{3}}{6} \end{pmatrix} \tag{3.115}$$

such that $p_i = \mathrm{tr}\left(L_i \rho L_i^\dagger\right)$. We note that $L_i = L_i^\dagger$ for all i. The completeness condition is satisfied: $\sum_{i=1}^{4} L_i^\dagger L_i = I$. Considering time reversal operator $\theta = -i\sigma_y K_0$, we showed that

$$\tilde{\Phi}_{qc}(\rho(t)) = \begin{pmatrix} \frac{3+z}{6} & \frac{x-iy}{6} e^{-iw_L t} \\ \frac{x+iy}{6} e^{iw_L t} & \frac{3-z}{6} \end{pmatrix} = \sigma \circ \Phi \circ \sigma(\rho(t)) \tag{3.116}$$

Therefore, the quantum operation of the SIC-POVM given above is quantum-classical time-reversal invariant. In addition, we note that star-SIC-POVM is quantum-classical time reversal of its SIC-POVM up to a unitary operation.

Now, we wish to know if there is any distinctive feature of SIC-POVM that cannot be handled in the frame of classical information theory. We recall that any quantum state can be expressed in terms of a SIC-POVM since the latter is informationally complete. Let us assume a quantum state ρ, and $\mathcal{M} = \{M_i = d^{-1} |\psi_i\rangle \langle\psi_i|, i = 1,2,\ldots,d^2\}$ and $\mathcal{N} = \{N_i = d^{-1} |\varphi_i\rangle \langle\varphi_i|, i = 1,2,\ldots,d^2\}$ be two SIC-POVMs with the corresponding probability distributions $\{p_i = \frac{1}{d}\langle\psi_i|\rho|\psi_i\rangle, i = 1,2,\ldots,d^2\}$ and $\{q_i = \frac{1}{d}\langle\varphi_i|\rho|\varphi_i\rangle, i = 1,2,\ldots,d^2\}$ in a d-dimensional Hilbert space.

According to equation (2.34), we can write

$$\rho = d(d+1)\sum_{i=1}^{d^2} p_i M_i - I = d(d+1)\sum_{i=1}^{d^2} q_i N_i - I, \tag{3.117}$$

from which we read the matrix that provides transition from $\{p_i = \frac{1}{d}\langle\psi_i|\rho|\psi_i\rangle, i = 1,2,\ldots,d^2\}$ to $\{q_i = \frac{1}{d}\langle\varphi_i|\rho|\varphi_i\rangle, i = 1,2,\ldots,d^2\}$ as

$$\Lambda = \begin{pmatrix} \frac{(d+1)|\langle\varphi_1|\psi_1\rangle|^2-1}{d} & \frac{(d+1)|\langle\varphi_1|\psi_2\rangle|^2-1}{d} & \cdots & \frac{(d+1)\left|\langle\varphi_1|\psi_{d^2}\rangle\right|^2-1}{d} \\ \frac{(d+1)|\langle\varphi_2|\psi_1\rangle|^2-1}{d} & \frac{(d+1)|\langle\varphi_2|\psi_2\rangle|^2-1}{d} & \cdots & \frac{(d+1)\left|\langle\varphi_2|\psi_{d^2}\rangle\right|^2-1}{d} \\ \vdots & \vdots & \ddots & \\ \frac{(d+1)\left|\langle\varphi_{d^2}|\psi_1\rangle\right|^2-1}{d} & \frac{(d+1)\left|\langle\varphi_{d^2}|\psi_2\rangle\right|^2-1}{d} & \cdots & \frac{(d+1)\left|\langle\varphi_{d^2}|\psi_{d^2}\rangle\right|^2-1}{d} \end{pmatrix}. \tag{3.118}$$

We simply write the vector form $\mathbf{q} = \Lambda\mathbf{p}$, where $\mathbf{q} = (q_1, q_2, \ldots, q_{d^2})^T$ and $\mathbf{p} = (p_1, p_2, \ldots, p_{d^2})^T$. We note that $\Lambda = (\lambda_{ij})_{d^2 \times d^2}$ is an orthogonal matrix, each of whose rows and columns sum to one, that is, $\sum_i^{d^2} \lambda_{ij} = \sum_{j=1}^{d^2} \lambda_{ij} = 1$ for all i, j, which are the conditions that define a doubly stochastic matrix together with the non-negativity of the entries. The entries of Λ could be negative in general. Therefore, due to its properties, we can call Λ pseudo doubly stochastic matrix. Our first quest is to seek for a possible physical description of Λ. If we would like matrix Λ of equation (3.118) to be a stochastic map, we stress on all of its entries the condition that they must be non-negative. In that case, Λ becomes a permutation matrix, which is a trivial transition, that is, it takes a probability distribution to itself up to a permutation. We know that Shannon entropies of two probability distributions are equal if one of them is a permutation of the other. Inversely, if Shannon entropies of two probability distributions are equal for all possible quantum states, then the probability distributions are permutations of one another. However, this is not what we demand for an entropy function as the measure of the information content. SIC-POVMs encode the information content of a quantum state in the same way. Therefore, we can reasonably expect that any function that quantifies the information content of a system should give the same amount of information for all the measurements that are the same type; for example it should give the same amount of information for all SIC-POVMs. Indeed, this is the main reason that pushed Brukner and Zeilinger to define a new measure of information content of systems (Brukner & Zeilinger, 1999). We will turn back to this point in Section 3.6.

That the matrix Λ could consists of negative entries implies that stochastic maps do not provide a frame for the description of quantum operation, especially the quantum operation of SIC-POVMs. If we remove the non-negativity condition on the entries, it becomes then possible to nail down a physical description of Λ. Accordingly, an alternative formalism of quantum operations has been constructed such that, in the formalism, probability vector plays the role of density matrix and a pseudo-stochastic matrix Λ takes place of quantum operation (Yashin et al., 2020; Fuchs & Schack, 2013).

To sum up this section, we first note that the expressions of time reversal operations in quantum mechanics and in the quantum operation formalism are not compatible with each other. We propose another expression which harmonizes them. We note that the distinction between principle system and environment disappears in this revised expression of time reversal operation. Apart from this change, we have preserved the formal structure of time reversal operation. Therefore, we cannot conclude that time reversal operation of total system (i.e., principle system+environment) is unitary.

In addition we showed that quantum operation of SIC-POVMs cannot be encapsulated in the frame of classical information theory, a fact that can be construed as the non-classical feature of SIC-POVMs.

3.5 An algebraic relation between MUBs and SIC-POVM

The existence of a SIC-POVM and of a set of $d+1$ MUBs is still a conundrum. We are more auspicious about the existence of SIC-POVMs than that of MUBs in every dimension. Although we have SIC-POVMs in dimension $d=6$, we do not yet have a set of 7 MUBs. There are some geometrical approaches that aim to reveal the connection between MUBs and SIC-POVMs (Wootters, 2006), however, we still do not know how exactly they are related to each other, if any exists. Since they are informationally complete, they can be connected to each other algebraically. Then, we can seek for the existence of a set of $d+1$ MUBs in dimension $d=6$ by means of some algebraic relations with a SIC-POVM in that dimension. Here, we present an algebraic relation between MUBs and SIC-POVMs.

Let us consider a set of $d+1$ MUBs, $\{\Pi_{nk}\}_{n,k=1,1}^{d+1,d}$, and a particular SIC-POVM, $\{N_j = \frac{1}{d}\Pi_j = \frac{1}{d}|\psi_j\rangle\langle\psi_j|\}_{j=1}^{d^2}$. Since the SIC-POVM is informationally complete and each Π_{nk} is a projector, we can write

$$\Pi_{nk} = \sum_{j=1}^{d^2} c_{nk}^{(j)}\Pi_j. \tag{3.119}$$

Taking the square of equation (3.119)

$$\Pi_{nk}^2 = \Pi_{nk} = \sum_{j,l=1,1}^{d^2,d^2} c_{nk}^{(j)}c_{nk}^{(l)}\Pi_j\Pi_l, \tag{3.120}$$

and using the equality $\Pi_j\Pi_l = \sum_{m=1}^{d^2}\alpha_{jlm}\Pi_m$, where $\alpha_{jlm} = \mathrm{tr}(\Pi_j\Pi_l\Pi_m)$ (Fuchs & Schack, 2013), and adjusting some dummy indices, we obtain an algebraic equation for the coefficients as

$$\sum_{m=1}^{d^2} c_{nk}^{(m)}\Pi_m = \sum_{m=1}^{d^2}\left(\sum_{j,l=1,1}^{d^2,d^2} c_{nk}^{(j)}c_{nk}^{(l)}\alpha_{jlm}\right)\Pi_m \Rightarrow c_{nk}^{(m)} = \sum_{j,l=1,1}^{d^2,d^2} c_{nk}^{(j)}c_{nk}^{(l)}\alpha_{jlm}. \tag{3.121}$$

Defining $|\mathbf{C}_{nk}\rangle := (c_{nk}^{(1)}, c_{nk}^{(2)}, \ldots, c_{nk}^{(d^2)}) \in \mathbb{C}^{d^2}$ and the matrix $\Lambda^{(m)} := (\alpha_{jlm}) \in M_{d^2 \times d^2}(\mathbb{C})$ over the indices j and l, we finally acquire

$$\Lambda^m |\mathbf{C}_{nk}\rangle = c_{nk}^m |\mathbf{C}_{nk}\rangle, \tag{3.122}$$

which is an eigenvalue problem for the coefficients. According to this equation, the eigenvalues of Λ^m are possible choices for the particular coefficient $c_{nk}^{(j)}$ in equation (3.119); that is, for each $c_{nk}^{(j)}$, we need to solve the eigenvalue problem of equation (3.122). Since the entries α_{jlm} of the matrices $\Lambda^{(m)}$ are the triple product of the elements of the SIC-POVM we consider, and since there exist SIC-POVMs in dimension $d = 6$, then we can seek at least numerically for the existence of MUBs in dimension $d = 6$ by means of the algebraic equation (3.122).

3.6 Information energy and quantum measurements

In classical information theory, we generally consider a random variable that has a set of events, or outcomes. One of the events is realized each time when the system is observed or measured. Let us call the random variable X with its corresponding set of events $\{x_i, i = 1, 2, \ldots, W\}$. Information content of the random variable X, is generally measured by entropy. To be more concrete, if a random variable has W possible outcomes, $\{x_i, i = 1, 2, \ldots, W\}$, through which the system manifests itself, and if the probabilities of the outcomes are given by $\{p_i, i = 1, 2, \ldots, W\}$ respectively, then the information content of the random variable is averagely expressed by Shannon entropy

$$H_S(p) := -\sum_{i=1}^{W} p_i \ln(p_i). \tag{3.123}$$

However, in the classical frame of information theory, the understanding of the information content of a system is fundamentally based on the occurrences of the events, and the state that exhibits the geometrical and dynamical structure of the system is not considered. When we turn our attention to the realm of quantum systems, we need to take the quantum state of the systems into account, which gives rise to a substantially different understanding of information.

The quantum state of a system in d-dimensional Hilbert space is characterized by a set of $d^2 - 1$ parameters. If one performs a measurement, like the measurement of a

SIC-POVM, or a set of measurements, such as $d+1$ MUBs, these real parameters can be determined, so that the quantum state of the system is ascertained uniquely. Such measurement or set of measurements are called informationally complete since the quantum state can be determined by means of performing one of them. Informationally complete means that the information content of the system can be encapsulated, or encoded, completely in terms of those measurements. On the other hand, we do not have only one SIC-POVM or one set of $d+1$ MUBs in a d-dimensional Hilbert space: If there is one SIC-POVM or one set of $d+1$ MUBs, then there are more than one of them. Intuitively, we expect that every SIC-POVM should encode the same amount of information, as well as we expect the sets of MUBs do the same. However, if the information encoded by, for instance, two SIC-POVMs is measured by Shannon entropy, we do not obtain the same amount of information. This is also valid for two different sets of $d+1$ MUBs. Eventually, we look for a function for measuring the information encoded by those informationally complete measurements such that it should give the same amount of information for the same type of measurements, for example, for every SIC-POVM.

Addressing to the problem stated above, Brukner and Zeilinger (Brukner & Zeilinger, 1999, 2009) proposed an alternative measure of information in terms of square of the probabilities

$$I(p) = \mathcal{N} \sum_{i}^{W} p_i^2 \tag{3.124}$$

which is generally known as information energy in the literature (Pardo & Taneja, 1991). $\mathcal{N}$ is normalization constant; it is determined according to the requirements that I(p) is equal to k bits of information, that is $W = 2^k$, when one probabilities is 1 and all the others are zero (Brukner & Zeilinger, 1999). Considering a set of $d+1$ MUBs with the set of respective probabilities $\{p_{nk}, n = 1, 2, \ldots, d+1; k = 1, 2, \ldots, d\}$, the measure of information is

$$I(p) = \sum_{n=1}^{d+1} I_n(p_n) = \mathcal{N} \sum_{n,k=1,1}^{d+1,d} p_{nk}^2, \tag{3.125}$$

where $I_n(p_n) = \mathcal{N} \sum_{k=1}^{d} p_{nk}^2$ is the information energy for the n^{th} measurement in the set of $d+1$ MUBs. It was generalized to any measurement scenario. Assume that $\mathcal{M} = \{A_j\}_{j=1}^{n}$ be a POVM, ρ be the quantum state of the relevant system and $p_j = \mathrm{tr}\left(A_j \rho A_j^\dagger\right)$ be the probabilities. In addition, let $\rho_\star$ be the invariant state under the quantum operation $\mathcal{E}$ such that $\mathcal{E}(\rho_\star) := \sum_j A_j \rho_\star A_j^\dagger = \rho_\star$. Defining the index function $\mathcal{C}(\mathcal{M} \mid \rho) := \sum_{j=1}^{n} p_j^2$, The Brukner and Zeilinger approach to the measure of information can be expressed as (Rastegin, 2015)

$$I(\mathcal{M} \mid \rho) = \mathcal{C}(\mathcal{M} \mid \rho) - \mathcal{C}(\mathcal{M} \mid \rho_{\star}). \tag{3.126}$$

According to (Brukner & Zeilinger, 1999), information is the fundamental entity for quantum theory, and there is an irreducible information unit such that every physical system is an information building block of this unit. rukner and Zeilinger to the measure of information has been criticized in different respects, especially for its physical reasoning (Shafittee et al., 2006; Khrennikov, 2016). Here, we provided a physical ground in the perspective of Stokes parameters for the use of the Brukner and Zeilinger approach to the measure of information.

To determine experimentally the state of the polarization of an arbitrary beam of electromagnetic radiation (photon), one must make a set of four measurements. The most convenient set of four measurements are those that yield the following information (Fano, 1949; McMaster, 1961):

(i) The intensity of beam,

(ii) The degree of plane polarization with respect to two arbitrary orthogonal axes,

(iii) The degree of plane polarization with respect to a set of axes oriented at 45° to the right of the orthogonal axes in item (ii).

(iv) The degree of circular polarization.

In optics, the second and third measurements can be performed by Nicol prisms, while the fourth requires additional use of a quarter-wave plate (Fano, 1949). Let us assume the general expression for polarization

$$\mathbf{E} = E_1 \mathrm{e}^{i(wt+\delta_1)} \hat{e}_1 + E_2 \mathrm{e}^{i(wt+\delta_2)} \hat{e}_2. \tag{3.127}$$

This general expression reduces to certain polarization under the following conditions:

(i) we have planed-polarized radiation if $\phi = \delta_1 - \delta_2 = 0$;

(ii) we have circular polarized radiation if $E_1 = E_2$ and $\phi = \delta_1 - \delta_2 = \pm\frac{\pi}{2}$;

(iii) we have elliptical polarization if $E_1 \neq E_2 \neq 0$ and $\phi \neq 0$.

A plane-polarized light can be expressed by the following quantum state $|\psi_L\rangle = a\,|e_1\rangle + b\,|e_2\rangle$. A measurement of linearly polarization then is

$$\Pi = \left(sin(\theta)\left|\psi_L^{\perp}\right\rangle + cos(\theta)\,|\psi_L\rangle\right)\left(sin(\theta)\left\langle\psi_L^{\perp}\right| + cos(\theta)\,\langle\psi_L|\right). \tag{3.128}$$

The probability is $\mathrm{tr}(\langle\psi_L|\Pi|\psi_L\rangle) = \frac{1}{2}(1+cos(2\theta))$. If the quantum state is a partially linearly polarized instead of being fully plane-polarized, it can be expressed as $\rho = \frac{1}{2}(1-P)I + P|\psi_L\rangle\langle\psi_L|$ (McMaster, 1954). Then, $\mathrm{tr}(\Pi\rho) = \frac{1}{2}(1+Pcos(2\theta))$. Here P is the degree of polarization which is determined by observing the maximum of transmitted light (i.e., $(1+P)/2$ for $\theta = 0$) or minimum fraction of transmitted light (i.e., $(1-P)/2$ for $\theta = \pi/2$) (Fano, 1949).

Now, Let us assume that we have a general polarized light (or a spin-1/2 system) with the quantum state

$$|\psi\rangle = \alpha_1|\psi_1\rangle + \alpha_2|\psi_2\rangle, \tag{3.129}$$

where the coefficients in general are complex and describes the amplitude and phase of the polarized light along the orthogonal unit vectors. We assume that intensity is normalized to unity: $I = |\alpha_1|^2 + |\alpha_2|^2 = 1$. Considering the unit vectors in equation (3.129) as the vectors of computational basis, one can rewrite the vector in the form of matrix

$$\rho_\alpha = |\psi\rangle\langle\psi| = \begin{pmatrix} \alpha_1\alpha_1^\star & \alpha_1\alpha_2^\star \\ \alpha_1^\star\alpha_1 & \alpha_2\alpha_2^\star \end{pmatrix} = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix}. \tag{3.130}$$

An orientation coefficient was defined as

$$P(|\psi_1\rangle, |\psi_2\rangle) = \alpha_1\alpha_1^\star - \alpha_2\alpha_2^\star = \rho_{11} - \rho_{22}, \tag{3.131}$$

which gives the difference of the intensity measurements of the pure (basis) states defined by $|\psi_1\rangle$ and $|\psi_2\rangle$ (McMaster, 1954).

We now perform a measurement for the degree of plane polarization with respect to a set of axes oriented at 45° to the right of the computational basis $\{|\psi_1\rangle, |\psi_2\rangle\}$. With these choice of axes, the quantum state in this case is

$$|\psi\rangle = \beta_1|\phi_1\rangle + \beta_2|\phi_2\rangle, \tag{3.132}$$

where

$$|\phi_1\rangle = cos(\pi/4)|\psi_1\rangle + sin(\pi/4)|\psi_2\rangle \tag{3.133}$$

$$|\phi_2\rangle = -sin(\pi/4)|\psi_1\rangle + cos(\pi/4)|\psi_2\rangle. \tag{3.134}$$

Then, we have

(3.135)
$$
\begin{aligned}
|\psi\rangle &= \beta_1(cos(\pi/4)\,|\psi_1\rangle + sin(\pi/4)\,|\psi_2\rangle) + \beta_2(-sin(\pi/4)\,|\psi_1\rangle + cos(\pi/4)\,|\psi_2\rangle) \\
&= (\beta_1 cos(\pi/4) - \beta_2 sin(\pi/4))\,|\psi_1\rangle + (\beta_1 sin(\pi/4) + \beta_2 cos(\pi/4))\,|\psi_2\rangle \\
&= \alpha_1\,|\psi_1\rangle + \alpha_2\,|\psi_2\rangle\,,
\end{aligned}
$$

which yields the equalities

$$
\begin{aligned}
\beta_1 cos(\pi/4) - \beta_2 sin(\pi/4) &= \alpha_1 \\
\beta_1 sin(\pi/4) + \beta_2 cos(\pi/4) &= \alpha_2.
\end{aligned} \tag{3.136}
$$

Putting equation (3.136) in the matrix-vector multiplication form, we obtain

$$
\begin{pmatrix} cos(\pi/4) & -sin(\pi/4) \\ sin(\pi/4) & cos(\pi/4) \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}, \tag{3.137}
$$

from which one leads to

$$
\begin{aligned}
\beta_1 &= cos(\pi/4)\alpha_1 + sin(\pi/4)\alpha_2 \\
\beta_2 &= -sin(\pi/4)\alpha_1 + cos(\pi/4)\alpha_2
\end{aligned}
$$

A second orientation coefficient for this new frame was defined as (McMaster, 1954)

$$
\begin{aligned}
P(|\phi_1\rangle, |\phi_2\rangle) &= \beta_1\beta_1^{\star} - \beta_2\beta_2^{\star} = \rho_\beta(11) - \rho_\beta(22) \\
&= \rho_\alpha(12) + \rho_\alpha(21) = \rho_{12} + \rho_{21}.
\end{aligned} \tag{3.138}
$$

Let us assume that $|\psi\rangle = \alpha_1\,|\psi_1\rangle + \alpha_2\,|\psi_2\rangle = b_1 e^{\mathrm{i}\delta_1}\,|\psi_1\rangle + b_2 e^{\mathrm{i}\delta_2}\,|\psi_2\rangle$ is to be a general polarized state. Then, a general quantum state of a beam is (Fano, 1949; McMaster, 1954)

$$
\rho = \frac{1}{2}(1-P)I + P\,|\psi\rangle\langle\psi| = \begin{pmatrix} \frac{1}{2}(1-P) + Pb_1^2 & Pb_1b_2e^{i(\delta_1-\delta_2)} \\ Pb_1b_2e^{-i(\delta_1-\delta_2)} & \frac{1}{2}(1-P) + Pb_2^2 \end{pmatrix}. \tag{3.139}
$$

In regard to the general quantum state in equation (3.139) we revised the orientation coefficient in equation (3.138) as $P(|\phi_1\rangle, |\phi_2\rangle) = \rho_{12} + \rho_{21} = 2Pb_1b_2cos(\delta_1 - \delta_2)$.

We now choose a third measurement that is for circularly polarized light. To make this measurement, we insert a quarter-wave plate with its fast axis 45° to the right of $|\psi_1\rangle$ and make intensity measurements with the transmission axis of Nicol prism oriented along $|\psi_1\rangle$ and $|\psi_2\rangle$. That is, we are making the choice

$$|\nu_1\rangle = \frac{i}{\sqrt{2}}|\psi_1\rangle + \frac{1}{\sqrt{2}}|\psi_2\rangle$$
$$|\nu_2\rangle = \frac{1}{\sqrt{2}}|\psi_1\rangle + \frac{i}{\sqrt{2}}|\psi_2\rangle,$$

from which we obtain

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}(|\nu_2\rangle - i|\nu_1\rangle)$$
$$|\psi_2\rangle = \frac{1}{\sqrt{2}}(|\nu_1\rangle - i|\nu_2\rangle).$$

The quantum state in equation (3.129) can be rewritten as

$$|\psi\rangle = \gamma_1|\nu_1\rangle + \gamma_2|\nu_2\rangle = \frac{1}{\sqrt{2}}(\alpha_2 - i\alpha_1)|\nu_1\rangle + \frac{1}{\sqrt{2}}(\alpha_1 - i\alpha_2)|\nu_2\rangle. \tag{3.140}$$

Similar to the previous orientation coefficients, a third orientation coefficient was defined as (McMaster, 1954)

$$P(|\nu_1\rangle, |\nu_2\rangle) = \gamma_1\gamma_1^\star - \gamma_2\gamma_2^\star = \rho_\nu(11) - \rho_v(22) = \mathrm{i}(\rho_{21} - \rho_{12}), \tag{3.141}$$

which is equal to $2Pb_1b_2\sin(\delta_1 - \delta_2)$ in regard to the general statement of the quantum state in equation (3.139). The orientation coefficients defined above become equal to Stokes parameters together with the intensity (I), which is equal to trace of the general quantum state in equation (3.139) in our case, that is, $I = \mathrm{tr}(\rho)$. The first Stokes parameter is equal to the difference of the probabilities of the spin's outcomes whose eigenbasis can be considered computational basis. To be more explicit, if we consider $|\psi_1\rangle$ and $|\psi_2\rangle$ in equation (3.129) as the eigenvectors of spin operator σ_z along z-axis, then $\{\Pi_{31} = |\psi_1\rangle\langle\psi_1|, \Pi_{32} = |\psi_2\rangle\langle\psi_2|\}$ is the measurement elements of σ_z, and the probability of the outcomes are $p_{31} = \mathrm{tr}(\Pi_{31}\rho) = \rho_{11}; p_{32} = \mathrm{tr}(\Pi_{32}\rho) = \rho_{22}$. Similarly, we are to consider $\{\Pi_{11} = |\phi_1\rangle\langle\phi_1|, \Pi_{12} = |\phi_2\rangle\langle\phi_2|\}$ be the measurement elements of spin operator σ_x along x-axis with the corresponding probabilities $\{p_{11} = \mathrm{tr}(\Pi_{11}\rho), p_{12} = \mathrm{tr}(\Pi_{12}\rho)\}$, and $\{\Pi_{21} = |\nu_1\rangle\langle\nu_1|, \Pi_{22} = |\nu_2\rangle\langle\nu_2|\}$ be the measurement elements of spin operator σ_y along y-axis with the corresponding probabilities $\{p_{21} = \mathrm{tr}(\Pi_{21}\rho), p_{22} = \mathrm{tr}(\Pi_{22}\rho)\}$. Now, it is clear that there is relation between Stokes parameters and the probabilities as $P(|\psi_1\rangle, |\psi_2\rangle) = p_{31} - p_{32}$, $P(|\phi_1\rangle, |\phi_2\rangle) = p_{11} - p_{12}$ and $P(|\nu_1\rangle, |\nu_2\rangle) = p_{21} - p_{22}$. We have the celebrated inequality

$$P(|\psi_1\rangle, |\psi_2\rangle)^2 + P(|\phi_1\rangle, |\phi_2\rangle)^2 + P(|\nu_1\rangle, |\nu_2\rangle)^2 \leq I = \mathrm{tr}(\rho), \tag{3.142}$$

where equality is satisfied if light is fully polarized, that is, if $P = 1$ in equation (3.139). We can rewrite equation (3.142) in terms of probabilities as

$$(p_{11}-p_{12})^2+(p_{21}-p_{22})^2+(p_{31}-p_{32})^2 \leq \mathrm{tr}(\rho), \tag{3.143}$$

which can be expressed formally as

$$\begin{aligned} \frac{1}{2}\sum_{n,k,s=1,1,1}^{d+1,d,d}(p_{nk}-p_{ns})^2 &\leq \mathrm{tr}(\rho) \\ d\sum_{n,k=1,1}^{d+1,d}p_{nk}^2-(d+1) &\leq \mathrm{tr}(\rho) \\ \sum_{n,k=1,1}^{d+1,d}p_{nk}^2 &\leq \frac{1}{d}\mathrm{tr}(\rho)+\frac{d+1}{d}, \end{aligned} \tag{3.144}$$

where the left side of the inequality in the last line is nothing but just the Brukner-Zeilinger approach to the measure of information up to the normalization constant $\mathcal{N}$. This fact suggests us that the square of the probabilities difference, so thus the Brukner-Zeilinger approach, could be more physical in order to measure the information content of a system.

If, for example, the quantum state of the system is one of the computational basis elements, say $|0\rangle$, then we have $\{p_{31}=1, p_{32}=0, p_{11}=p_{12}=p_{21}=p_{22}=\frac{1}{2}\}$, and thus, $(p_{11}-p_{12})^2+(p_{21}-p_{22})^2+(p_{31}-p_{32})^2=1$. This fact characterizes our certainty about the quantum state of the system, that is, the summation of the probabilities difference can be considered as the measure of the information that an observer has about the system. Therefore, we can restate Brukner and Zeilinger approach to the measure of information as

$$I(p)=\sum_{n=1}^{d+1}I_n(p_n) := \frac{\mathcal{N}}{2}\sum_{n,k,s=1,1,1}^{d+1,d,d}(p_{nk}-p_{ns})^2. \tag{3.145}$$

However, this expression is for a set of $d+1$ MUBs; but, how can we define it for a SIC-POVM? The most reasonable way is to make use of the informationally completeness of MUBs and SIC-POVMs. Let us consider the quantum state ρ of the system under consideration, a set of $d+1$ MUBs $\{\Pi_{nk}\}_{n,k=1,1}^{d+1,d}$ with probabilities $\{q_{nk}=\mathrm{tr}(\Pi_{nk}\rho)\}_{n,k=1,1}^{d+1,d}$, and a SIC-POVM $\mathcal{N}=\{\frac{1}{d}\Pi_j\}_{j=1}^{d^2}$ with probabilities $\{p_j = \frac{1}{d}\mathrm{tr}(\Pi_j\rho)\}_{j=1}^{d^2}$. By means of equations (2.16) and (2.34) we can write

$$\sum_{n,k=1,1}^{d+1,d}q_{nk}\Pi_{nk}-I=\rho=(d+1)\sum_{j=1}^{d^2}p_j\Pi_j-I, \tag{3.146}$$

from which one can obtain $q_{nk}+1=(d+1)\sum_j p_j t_{nk}^{(j)}$ with $t_{nk}^{(j)}=\mathrm{tr}(\Pi_j\Pi_{nk})$ such that $\sum_j t_{nk}^{(j)}=d$ and $\sum_k t_{nk}^{(j)}=1$. Substituting these equalities into the equation (3.145) and using the fact $\sum_{n=1,k=1}^{d+1,d} t_{nk}^{(i)}t_{nk}^{(j)}=\frac{d(1+\delta_{ij})+2}{d+1}$, we obtain

$$
\begin{aligned}
\frac{1}{2}\sum_{n,k,s=1,1,1}^{d+1,d,d}(q_{nk}-q_{ns})^2 &= d^2(d+1)\sum_{j=1}^{d^2}p_j^2-(d+1) \\
&= \frac{d+1}{2}\sum_{i,j=1,1}^{d^2,d^2}(p_i-p_j)^2 .
\end{aligned}
\tag{3.147}
$$

This equality is a consequence of the formal expressions of MUBs and SIC-POVMs, and is valid if MUBs and SIC-POVMs do exist in every dimension, that is, the validity is ensured by the existence of MUBs and SIC-POVMs. Could we still meaningfully argue this equality if MUBs and SIC-POVMs do not exist? For example, we do not know whether or not there are $d+1$ MUBs in dimension six. We wish to recall the discussion in Section 2.2; the consequences (e.g., SK axioms) from the formal statement of some entity (e.g., entropy) cannot replace the being of the entity (e.g., entropy).

4. CONCLUSION

In the context of the thesis, we studied entropic uncertainty relations in two quantum measurements that are MUBs and SIC-POVMs. These two measurements are important due to the fact that they have found room in diverse application areas in quantum information theory. In addition, entropic uncertainty relations provide an alternative approach to quantify uncertainty in quantum mechanics. The advantage of entropic uncertainty relations over deviation-based uncertainty relations is that they do not suffer from many drawbacks that the latter relations have.

We have obtained the optimal ECR ($H_T^+(N, d, \alpha)$) for N MUBs if the state of the relevant system can be expressed in terms of N MUBs whose importance is two folds. First of all, this bound implies that the entropies of the observables cannot achieve to their maximum values ($\ln(d)$) simultaneously which, together with lower bound of EUR, determines the upper bound of information exclusion relation for MUBs. The crucial point in our derivation is the condition satisfied by the purities of the observables. As pointed out, the purities of the observables corresponding to N MUBs are dependent on each other; therefore, we have considered them in the maximization of the total entropy. If an equality relation for the summation of the purities of N MUBs exists for a general density matrix, our result can be extended directly. An equality of this sort will be most likely related to the dimension of the system (d), the state (ρ) and the number of MUBs (N). Furthermore, we have also obtained an optimal upper bound of Shannon entropy for any SIC-POVM in d-dimensional Hilbert space. We noted that it is possible to speak of ECR for SIC-POVM meaningfully if we consider the measurement elements of the SIC-POVM as observables. However, it is meaningless to speak of EUR or ECR of two SIC-POVMs.

Secondly, we have shown that ECR for MUBs provides a criterion for the existence of mutually coherent states, which are related to the existence of MUBs. In that case, we have inferred that any quantum state resulting in a total entropy of N MUBs more than the upper bound $H_T^+(N, d, \alpha)$ of total entropy of MUBs consists of at least a mutually coherent state as its part. A question can be argued in connection with the criterion: Can we assert that the optimal upper bound, $H_T^+(N, d, \alpha)$, cannot be exceeded if there is no a mutually coherent states? The optimal upper bound cannot be exceeded on the condition that the density matrix is not maximally mixed state. Therefore, the premise that "if there is no a mutually coherent state $|\psi_{coh}\rangle$ with respect to N MUBs, this set of N MUBs cannot be extended to $(N+1)MUBs$" becomes biconditional if density matrix is not maximally mixed state.

We have also applied entropic certainty relation to the summation of the mutual information of $(d+1)$ complementary observables $\{A_n\}_{n=1}^{d+1}$ conditioned with a classical memory. Our result implies that it is not possible to construct a deterministic correlation between a classical memory and each of $(d+1)$ observables simultaneously. One can make use of our result to detect whether the observables are compatible. If the mutually information of the observables set $\{A_n\}_{n=1}^{d+1}$ correlated with a classical memory cannot achieve its maximum value $((d+1)\ln(d))$, one can infer that the observables are not compatible. In addition, if an observables becomes more correlated to the memory, the other becomes less correlated with the memory.

We gave some concrete examples that show the problematic nature of the deviation-based uncertainty relations in the sense that for some realistic examples, they loose their ability of quantifying uncertainty and give rise to some counter-intuitive results. Like the problems that exist for the deviation-based uncertainty relations, it is also claimed that entropic uncertainty relations do not capture exactly the physical content of the uncertainty principle that were examined by Heisenberg. To be more explicit, Hilgevoord and Uffink argued that entropic uncertainty relations do not quantify the complementarity of the observables; instead, the aforementioned relations put forward the following statement:

> "In quantum mechanics, it is impossible to prepare any system in a state $|\psi\rangle$ such that its position and momentum are both precisely predictable, in the sense of having both the expected spread in a measurement of position and the expected spread in a momentum measurement arbitrarily small" (Hilgevoord & Uffink, 2016).

Quantification of uncertainty of the observables on distinct but the same prepared state is called *preparation uncertainty relation*. In the expression of entropic uncertainty relations, entropy of the observables are computed on different copies of the quantum state; that is, they are not computed in terms of successive measurements. In that case, it might be true that entropic uncertainty relations do not grasp the physical content of the notion of uncertainty, but provide a quantification of the preparation uncertainty. However, if we restrict the notion of uncertainty to the commutation relation of the observables, we can safely claim that entropic uncertainty relation indeed reflects in themselves the exact meaning of uncertainty. But, if we take uncertainty as the impossibility of joint measurement of the observables, then it is true that entropic uncertainty relations do not represent the notion of uncertainty comprehensively. The contentious nature of uncertainty avoids us to develop a quantitative measure of it. There are some works devoted to express entropic uncertainties of successive measurements that were to give rise to the same lower bound as EURs (Srinivas, 2003; Baek & Son, 2016; Rastegin, 2018).

Classification of SIC-POVMs is an important issue by which we can economize the efforts that is to be spent for finding SIC-POVMs analytically. When comparing a SIC-POVM with its star-SIC-POVM, we recognized that the notions of time reversal in quantum mechanics and in quantum operation formalism are not compatible with each other. The reason is that while in the former the system itself is considered for reversing in time, in the latter it is the environment that is going to be reversed. We proposed a harmonization of the two aspects form which we showed that SIC-POVMs are time reversal invariant, such that star-SIC-POVM is time reversal of its corresponding SIC-POVM up to a unitary transformation. In addition, we obtain an algebraic relation between MUBs and SIC-POVMs, which, for example, enable us to search the existence of fourth MUB numerically in dimension 6. If we succeed to find the fourth MUB by means of this algebraic relation, it would be a great progress for the existence of further MUBs; otherwise, we can read it as an evidence for the non-existence of further MUBs.

Finally, we provided a physical ground for the use of information energy in quantum information theory. It seems that the alternative expression of information energy in terms of the square of the difference of probabilities serves to quantify our information about the system, while any entropy function itself is to quantify our lack of the information. It is well known that Rényi entropy $I_\alpha(p)$ of parameter $\alpha = 2$ fulfills the requirements that has been argued by Brukner and Zeilinger for the invariance of the information content of a system. Being inspired by kinetic and potential energies, could we speculate that there might be information potential which can be expressed in terms of Réyi entropy? The only thing that can be said at this moment is that it is worth considering this question.

BIBLIOGRAPHY

Abe, S. (2002). Stability of Tsallis entropy and instabilities of Rényi and normalized Tsallis entropies: A basis for q-exponential distributions. *Physical Review E*, *66*(4), 046134.

Abe, S. & Thurner, S. (2005). Anomalous diffusion in view of Einstein's 1905 theory of Brownian motion. *Physica A: Statistical Mechanics and its Applications*, *356*(2), 403–407.

Adabi, F., Salimi, S., & Haseli, S. (2016). Tightening the entropic uncertainty bound in the presence of quantum memory. *Physical Review A*, *93*, 062123.

Albouy, O. & Kibler, M. R. (2007). A unified approach to sic-povms and mubs. *Journal of Russian Laser Research*, *28*, 429–438.

Appleby, D. M. (2005). Symmetric informationally complete–positive operator valued measures and the extended clifford group. *Journal of Mathematical Physics*, *46*(5), 052107.

Appleby, D. M. (2009). SIC-POVMS and MUBS: Geometrical Relationships in Prime Dimension. *AIP Conference Proceedings*, *1101*(1), 223–232.

Appleby, D. M., Bengtsson, I., Grassl, M., Harrison, M., & McConnell, G. (2021). IC-POVMs from Stark Units.

Appleby, M. & Bengtsson, I. (2019). Simplified exact SICS. *Journal of Mathematical Physics*, *60*(6), 062203.

Appleby, M., Bengtsson, I., Dumitru, I., & Flammia, S. (2017). Dimension towers of SICs. I. Aligned SICs and embedded tight frames. *Journal of Mathematical Physics*, *58*(11), 112201.

Appleby, M., Chien, T. Y., Flammia, S., & Waldron, S. (2018). Constructing exact symmetric informationally complete measurements from numerical solutions. *Journal of Physics A: Mathematical and Theoretical*, *51*(16), 165302.

Appleby, M., Flammia, S., McConnell, G., & Yard, J. (2017). SICs and Algebraic Number Theory. *Foundations of Physics*, *47*, 1042–1059.

Baek, K. & Son, W. (2016). Entropic Uncertainty Relations for Successive Generalized Measurements. *Mathematics*, *4*(2).

Bandyopadhyay, S., Boykin, P. O., Roychowdhury, V., & Vatan, F. (2002). A new proof for the existence of mutually unbiased bases. *Algorithmica*, *34*, 512–528.

Barnett, S. M. (2009). *Quantum Information* (1 ed.). New York: Oxford University Press.

Beckner, W. (1975). Inequalities in Fourier analysis. *Ann. Math.*, *102*(1), 159.

Bengtsson, I. (2017). The number behind the simplest sic–povm. *Foundations of Physics*, *47*, 1031–1041.

Bera, A., Das, D., Sen(De), A., & Sen, U. (2019). Universal quantum uncertainty relations: Minimum-uncertainty wave packet depends on measure of spread. *Physics Letters A*, *383*(16), 1850–1855.

Bergh, B. & Gärttner, M. (2021). Entanglement detection in quantum many-body systems using entropic uncertainty relations. *Physical Review A*, *103*(5), 052412.

Berta, M., Christandl, M., Colbeck, R., Renes, J. M., & Renner, R. (2010). The uncertainty principle in the presence of quantum memory. *Nature Physics*, *6*(9), 659.

Bialynicki-Birula, I. (2006). Formulation of the uncertainty relations in terms of the rényi entropies. *Physical Review A*, *74*(5), 052101.

Bialynicki-Birula, I. & Mycielski, J. (1975). Uncertainty Relations for Information Entropy in Wave Mechanics. *Communications in Mathematical Physics*, *44*, 129.

Bialynicki-Birula, I. & Rudnicki, L. (2011). Entropic Uncertainty Relations in Quantum Physics. In K. D. Sen (Ed.), *Statistical Complexity* (1 ed.). (pp. 4–6). Dortrecht: Springer Netherlands.

Born, M., Heisenberg, W., & Jordan, P. (1926). Zur Quantenmechanik. II. *Zeitschrift für Physik*, *35*, 557–615.

Born, M. & Jordan, P. (1925). Zur Quantenmechanik. *Zeitschrift für Physik*, *34*, 858–888.

Bouchard, F., Heshami, K., England, D., Fickler, R., Boyd, R. W., Englert, B., Sánchez-Soto, L. L., & Karimi, E. (2018). Experimental investigation of high-dimensional quantum key distribution protocols with twisted photons. *Quantum*, *2*, 111.

Brierley, S. & Weigert, S. (2008). Maximal sets of mutually unbiased quantum states in dimension 6. *Physical Review A*, *78*(4), 042312.

Brukner, C. & Zeilinger, A. (2009). Information invariance and quantum probabilities. *Foundations of Physics*, *39*, 677–689.

Brukner, i. c. v. & Zeilinger, A. (1999). Operationally invariant information in quantum measurements. *Physical Review Letters*, *83*(17), 3354–3357.

Busch, P., Heinonen, T., & Lahti, P. (2007). Heisenberg's uncertainty principle. *Physics Reports*, *452*(6), 155–176.

Bush, P., Lahti, P., & Werner, R. F. (2013). Proof of Heisneberg's Error-Disturbance Relation. *Physical Review Letters*, *111*(16), 160405.

Calderbank, A. R., Rains, E. M., Shor, P. W., & Sloane, N. J. A. (1997). Quantum Error Correction and Orthogonal Geometry. *Phys. Rev. Lett.*, *78*(3), 405.

Canturk, B. & Gedik, Z. (2021). Optimal upper bound of entropic uncertainty relation for mutually unbiased bases. *Physica A: Statistical Mechanics and its Applications*, *582*, 126275.

Cerf, N. J., Bourennane, M., Karlsson, A., & Gisin, N. (2002). Security of quantum key distribution using d-level systems. *Physical Review Letters*, *88*(12), 127902.

Chen, B., Li, T., & Fei, S. M. (2015). General SIC measurement-based entanglement detection. *Qunatum Information Process*, *14*, 2281–2290.

Chen, L. & Yu, L. (2017). Product states and schmidt rank of mutually unbiased bases in dimension six. *Journal of Physics A: Mathematical and Theoretical*, *50*(47), 475304.

Chen, L. & Yu, L. (2018). Mutually unbiased bases in dimension six containing a product-vector basis. *Quantum Information Processing*, *17*, 198.

Coles, P. J., Berta, M., Tomamichel, M., & Wehner, S. (2017). Entropic uncertainty relations and their applications. *Rev. Mod. Phys.*, *89*(1), 015002.

Combes, J., Ferrie, C., Leifer, M. S., & Pusey, M. F. (2018). Why protective measurement does not establish the reality of the quantum state. *Quantum Studies: Mathematics and Foundations*, *5*(2), 189–211.

Crooks, G. E. (2008). Quantum operation time reversal. *Physical Review A*, *77*, 034101.

D'Ambrosio, V., Cardano, F., Karimi, E., Nagali, E., Santamato, E., Marrucci, L., & Sciarrino, F. (2013). Test of mutually unbiased bases for six-dimensional photonic quantum systems. *Scientific Reports*, *3*, 2726.

Dammeier, L., Schwonnek, R., & Werner, R. F. (2015). Uncertainty relations for angular momentum. *New Journal of Physics*, *17*(9), 093046.

Demirel, B., Sponar, S., & Hasegawa, Y. (2020). Measurements of entropic uncertainty relations in neutron optics. *Applied Sciences*, *10*(3).

Deutsch, D. (1983). Uncertainty in Quantum Measurements. *Phys. Rev. Lett.*, *50*(9), 631.

Dolatkhah, H., Haseli, S., Salimi, S., & Khorashad, A. S. (2020). Tightening the tripartite quantum-memory-assisted entropic uncertainty relation. *Physical Review A*, *102*(5), 052227.

Durt, T., Englert, B., Bengtsson, I., & Życzkowski, K. (2010). On Mutually Unbiased Bases. *International Journal of Quantum Information*, *08*(04), 535.

Erhart, J., Sponar, S., Sulyok, G., Badurek, G., Ozawa, M., & Hasegawa, Y. (2012). Experimental demonstration of a universally valid error–disturbance uncertainty relation in spin measurements. *Nature Physics*, *8*, 185–189.

Everett, H. (1957). "Relative State" Formulation of Quantum Mechanics. *Rev. Mod. Phys.*, *29*(3), 454.

Fagnola, F. & Umanitá (2008). Detailed balance, time reversal, and generators of quantum Markov semigroups. *Mathematical Notes*, *84*, 108–115.

Fano, U. (1949). Remarks on the classical and quantum-mechanical treatment of partial polarization. *J. Opt. Soc. Am.*, *39*(10), 859–863.

Feng, J., Zhang, Y. Z., & Gould, M. D.and Fan, H. (2015). Uncertainty relation in Schwarzschild spacetime. *Physics Letters B*, *743*, 198.

Floerchinger, S., Haas, T., & Hoeber, B. (2021). Relative entropic uncertainty relation. *Physical Review A*, *103*(6), 062209.

Frank, R. L. & Lieb, E. H. (2012). Entropy and the Uncertainty Principle. *Annales Henri Poincaré*, *13*, 1711–1717.

Frauchiger, D. & Renner, R. (2018). Quantum theory cannot consistently describe the use of itself. *Nature Communications*, *9*(1), 3711.

Fuchs, C. A., Hoang, M. C., & Stacey, B. C. (2017). The SIC Question: History and State of Play. *Axioms*, *6*(3).

Fuchs, C. A., Mermin, N. D., & Schack, R. (2014). An introduction to qbism with an application to the locality of quantum mechanics. *American Journal of Physics*, *82*(8), 749.

Fuchs, C. A. & Schack, R. (2013). Quantum-bayesian coherence. *Review of Modern Physics*, *85*(4), 1693–1715.

Fuentes-Schuller, I. & Mann, R. B. (2005). Alice Falls into a Black Hole: Entanglement in Noninertial Frames. *Physical Review Letters*, *95*(12), 120404.

Furrer, F., Berta, M., Tomamichel, M., Scholz, V. B., & Christandl, M. (2014). Position-momentum uncertainty relations in the presence of quantum memory. *Journal of Mathematical Physics*, *55*(12), 122205.

Gottfried, K. (1989). *Quantum Mechanics* (3 ed.). Massachusetts: Perseus Books.

Goyeneche, D. (2013). Mutually unbiased triplets from non-affine families of complex hadamard matrices in dimension 6. *Journal of Physics A: Mathematical and Theoretical*, *46*(10), 105301.

Grassl, M. & Scott, A. J. (2017). Fibonacci-Lucas SIC-POVMs. *Journal of Mathematical Physics*, *58*(12), 122201.

Guanlei, X., Xiaogang, X., & Xiaotong, W. (2021). Generalized uncertainty relations of Tsallis entropy on FrFT. *Signal,Image and Video Processing*, *15*, 9–16.

Hall, M. J. W. (1995). Information Exclusion Principle for Complementary Observables. *Physical Review Letters*, *74*(17), 3307.

Heisenberg, H. (1925). Über quantentheoretische umdeutung kinematischer und mechanischer beziehungen. *Zeitschrift für Physik*, *33*(1), 879–893. English translation is available in (Van der Waerden, 1967, pp. 261-276).

Heisenberg, W. (1927). Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, *43*, 172. English translation is available in (Wheeler & Zurek, 1983, pp. 62-84).

Hertz, A. & Cerf, N. J. (2019). Continuous-variable entropic uncertainty relations. *Journal of Physics A: Mathematical and Theoretical*, *52*(17), 173001.

Hilgevoord, J. & Uffink, J. (2016). The uncertainty principle. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 ed.). Metaphysics Research Lab, Stanford University.

Hirschman, I. (1957). A note on entropy. *Am. J. Math.*, *79*(1), 152.

Hoggar, S. G. (1998). 64 lines from a quaternionic polytope. *Geometriae Dedicata*, *69*, 287–289.

Huang, J. L., Gan, W. C., Xiao, Y., Shu, F. W., & Yung, M. H. (2018). Holevo bound of entropic uncertainty in Schwarzschild spacetime. *The European Physical Journal C*, *78*, 545.

Hughston, L. P. & Salamon, S. M. (2016). Surveying points in the complex projective plane. *Advances in Mathematics*, *286*, 1017–1052.

Ikot, A., Rampho, G., Amadi, P., Okorie, U., Sithole, M., & Lekala, M. (2020). Quantum information-entropic measures for exponential-type potential. *Results in Physics*, *18*, 103150.

Inoue, Y. & Ozawa, M. (2020). Error-disturbance relation in Stern-Gerlach measurements. *Phys. Rev. A*, *102*(4), 042226.

Ivanovic, I. D. (1981). Geometrical description of quantal state determination. *Journal of Physics A: Mathematical and General*, *14*, 3241.

Ivanovic, I. D. (1992). An inequality for the sum of entropies of unbiased quantum measurements. *Journal of Physics A: Mathematical and General*, *25*(7), L363.

Jaming, P., Matolcsi, M., & Móra, P. (2010). The problem of mutually unbiased bases in dimension 6. *Cryptography and Communications*, *2*, 211–220.

Jaynes, E. (1957a). Information theory and statistical mechanics. *Physical Review*, *106*(4), 620–630.

Jaynes, E. (1957b). Information theory and statistical mechanics. II. *Physical Review*, *108*(2), 171–190.

Jedwab, J. & Wiebe, A. (2016). Constructions of complex equiangular lines from mutually unbiased bases. *Designs, Codes and Cryptography*, *80*, 73–89.

Jizba, P. & Korbel, J. (2019). Maximum entropy principle in statistical inference: Case for non-shannonian entropies. *Physical Review Letters*, *122*(12), 120601.

Jizba, P. & Korbel, J. (2020). When shannon and khinchin meet shore and johnson: Equivalence of information theory and statistical inference axiomatics. *Physical Review E*, *101*(4), 042126.

Kennard, E. (1927). Zur quantenmechanik einfacher bewegungstypen. *Zeitschrift für Physik*, *44*, 326–352.

Khinchin, A. I. (1957). *Mathematical foundations of information theory* (1 ed.). New York: Dover Publication.

Khrennikov, A. (2016). Reflections on zeilinger–brukner information interpretation of quantum mechanics. *Foundations of Physics*, *46*, 836–844.

Kimura, G. (2003). The bloch vector for n-level systems. *Physics Letters A*, *314*(5), 339–349.

Klappenecker, A. & Rötteler, M. (2003). Constructions of mutually unbiased bases.

Kondepudi, D. & Prigogine, I. (2015). *Modern thermodynamics: from heat engines to dissipative structures* (2 ed.). New Delhi: John Wiley & Sons, Inc.

Kopp, G. S. (2021). SIC-POVMs and the Stark Conjectures. *International Mathematics Research Notices*, *2021*(18), 13812–13838.

Kraus, K. (1987). Complementary observables and uncertainty relations. *Physical Review D*, *35*(10), 3070–3075.

Lanczos, C. (1986). *The Variational Principles of Mechanics.* New York: Dover Publications.

Lesche, B. (1982). Instabilities of Reńyi entropies. *Journal of Statistical Physics*, *27*, 419–422.

Lewis, P. G., Jennings, D., Barrett, J., & Rudolph, T. (2012). Distinct quantum states can be compatible with a single state of reality. *Physical Review Letters*, *109*(15), 150404.

Li, C. F., Xu, J. S., Xu, X. Y., Li, K., & Guo, G. C. (2011). Experimental investigation of the entanglement-assisted entropic uncertainty principle. *Nature Physics*, *7*, 752.

Liang, M., Hu, M., Sun, Y., Chen, L., & Chen, X. (2021). Real entries of complex Hadamard matrices and mutually unbiased bases in dimension six. *Linear and Multilinear Algebra*, *69*(15), 2908–2925.

Maassen, H. & Uffink, J. (1988). Generalized Entropic Uncertainty Relations. *Phys. Rev. Lett.*, *60*(12), 1103.

Mafu, M., Dudley, A., Goyal, S., Giovannini, D., Mclaren, M., Padgett, M. J., Konrad, T., Petruccione, F., Lütkenhaus, N., & Forbes, A. (2013). Higher-dimensional orbital-angular-momentum-based quantum key distribution with mutually unbiased bases. *Phys. Rev. A*, *88*, 032305.

Mandayam, P., Bandyopadhyay, S., Grassl, M., & Wootters, W. K. (2014). Unextendible Mutually Unbiased Bases from Pauli Classes. *Quantum Information and Computation*, *14*(9-10), 823.

Margenau, H. (1963). Measurements in quantum mechanics. *Annals of Physics*, *23*(3), 469–485.

Maroney, O. (2009). Information Processing and Thermodynamic Entropy. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2009 ed.). Metaphysics Research Lab, Stanford University.

McMaster, W. H. (1954). Polarization and the stokes parameters. *American Journal of Physics*, *22*(6), 351–362.

McMaster, W. H. (1961). Matrix representation of polarization. *Review of Modern Physics*, *33*, 8–28.

Ming, F., Wang, D., Fan, X. G., Shi, W. N., Ye, L., & Chen, J. L. (2020). Improved tripartite uncertainty relation with quantum memory. *Physical Review A*, *102*, 012206.

Ming, F., Wang, D., & Ye, L. (2019). Dynamical Measurement's Uncertainty in the Curved Space-Time. *Annals of Physics (Berlin)*, *531*, 1900014.

Mondal, D., Bagchi, S., & Pati, A. K. (2017). Tighter uncertainty and reverse uncertainty relations. *Physical Review A*, *95*(5), 052117.

Moreira, R., Carvalho, A., Mendes, M., Moreno, M., Ferraz, J., Parisio, F., Acioli, L., & Felinto, D. (2018). Toward an experimental test for the finite-time wave function collapse. *Optics Communications*, *426*, 212–218.

Moy, G. M., Hope, J. J., & Savage, C. M. (1999). Born and markov approximations for atom lasers. *Physical Review A*, *59*(1), 667–675.

Nairz, O., Arndt, M., & Zeilinger, A. (2002). Experimental verification of the heisenberg uncertainty principle for fullerene molecules. *Physical Review A*, *65*(3), 032109.

Nielsen, M. A. & Chuang, I. L. (2010). *Quantum Computation and Quantum Information* (10th Aniversary ed.). Cambridge University Press.

Nikolic, D. & Nesic, L. J. (2011). Verification of the uncertainty principle by using diffraction of light waves. *European Journal of Physics*, *32*(2), 467–477.

Oikonomou, T. & Bagci, G. B. (2019). Rényi entropy yields artificial biases not in the data and incorrect updating due to the finite-size data. *Physical Review E*, *99*(3), 032134.

Olendski, O. (2019). Quantum information measures of the aharonov–bohm ring in uniform magnetic fields. *Physics Letters A*, *383*(11), 1110–1122.

Oreshkov, O., Calsamiglia, J., Muñoz-Tapia, R., & Bagan, E. (2011). Optimal signal states for quantum detectors. *New Journal of Physics*, *13*(7), 073032.

Ozawa, M. (2003). Universally valid reformulation of the heisenberg uncertainty principle on noise and disturbance in measurement. *Physical Review A*, *67*(4), 042105.

Pardo, L. & Taneja, I. (1991). Information energy and its aplications. In P. W. Hawkes (Ed.), *Advances in Electronics and Electron Physics*, volume 80 (pp. 165–241). Academic Press.

Plastino, A. R. & Wedemann, R. S. (2020). Nonlinear Fokker–Planck Equation Approach to Systems of Interacting Particles: Thermostatistical Features Related to the Range of the Interactions. *Entropy*, *22*(2).

Pourkarimi, M. R., Haddadi, S., & Haseli, S. (2020). Exploration of entropic uncertainty bound in a symmetric multi-qubit system under noisy channels. *Physica Scripta*, *96*(1), 015101.

Pressé, S. (2014). Nonadditive entropy maximization is inconsistent with bayesian updating. *Physical Review E*, *90*(5), 052149.

Pressé, S., Ghosh, K., Lee, J., & Dill, K. A. (2013). Nonadditive entropies yield probability distributions with biases not warranted by the data. *Physical Review Letters*, *111*(18), 180604.

Prevedel, R., Hamel, D. R., Colbeck, R., Fisher, K., & Resch, K. J. (2011). Experimental investigation of the uncertainty principle in the presence of quantum memory and its application to witnessing entanglement. *Nat. Phys.*, *7*(10), 757.

Puchała, Z., Rudnicki, L., Chabuda, K., Paraniak, M., & Życzkowski, K. (2015). Certainty relations, mutual entanglement, and nondisplaceable manifolds. *Physical Review A*, *92*(3), 032109.

Pusey, M. F., Barrett, J., & Rudolph, T. (2012). On the reality of the quantum state. *Nature Physics*, *8*, 475.

Qu, D., Wang, W., Xiao, L., Zhan, X., & Xue, P. (2021). Experimental demonstration of strong unitary uncertainty relations. *Optics Express*, *29*(18), 29567–29575.

Rajagopal, A. K. (1995). The sobolev inequality and the tsallis entropic uncertainty relation. *Physics Letters A*, *205*(1), 32–36.

Rastegin, A. E. (2010). Rényi formulation of the entropic uncertainty principle for POVMs. *Journal of Physics A: Mathematical and Theoretical*, *43*(15), 155302.

Rastegin, A. E. (2013). Uncertainty relations for MUBs and SIC-POVMs in terms of generalized entropies. *The European Physical Journal D* , *67*, 269.

Rastegin, A. E. (2015). On the brukner–zeilinger approach to information in quantum measurements. *Proceedings of the Royal Society A*, *471*, 20150435.

Rastegin, A. E. (2018). Entropic Uncertainty Relations for Successive Measurements in the Presence of a Minimal Length. *Entropy*, *20*(5).

Renes, J. M. (2005). Equiangular spherical codes in quantum cryptography. *Quantum Information and Computation*, *5*(1), 81–92.

Renes, J. M., Blume-Kohout, R., Scott, A. J., & Caves, C. M. (2004). Symmetric informationally complete quantum measurements. *Journal of Mathematical Physics*, *45*(6), 2171–2180.

Renes, J. M., Scholz, V. B., & Huber, S. (2017). Uncertainty relations: An operational approach to the error-disturbance tradeoff. *Quantum*, *1*, 20.

Reńyi, A. (1961). On measure of entropy and information. In *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics* , volume 1, (pp. 547–561). University of California Press.

Reńyi, A. (1965). On the foundations of information theory. *Review of the International Statistical Institute*, *33*(1), 1–14.

Robertson, H. P. (1929). The uncertainty principle. *Physcal Review*, *34*(1), 163–164.

Rodríguez, A., Nobre, F. D., & Tsallis, C. (2019). d-Dimensional Classical Heisenberg Model with Arbitrarily-Ranged Interactions: Lyapunov Exponents and Distributions of Momenta and Energies. *Entropy*, *21*(1).

Rozema, L. A., Darabi, A., Mahler, D. H., Hayat, A., Soudagar, Y., & Steinberg, A. M. (2012). Violation of heisenberg's measurement-disturbance relationship by weak measurements. *Physics Review Letters*, *109*(10), 100404.

Rumin, M. (2011). Balanced distribution-energy inequalities and related entropy bounds. *Duke Mathematical Journal*, *160*(3), 567–597.

Sánchez-Ruiz, J. (1993). Entropic uncertainty and certainty relations for complementary observables. *Physics Letters A*, *173*(3), 233.

Sánchez-Ruiz, J. (1995). Improved bounds in the entropic uncertainty and certainty relations for complementary observables. *Physics Letters A*, *201*(2-3), 125.

Schrödinger, E. (1928). *Collected Papers on Wave Mechanics*. London: Blackie and Son.

Schrödinger, E. (1930). Zum heisenbergschen unschärfeprinzip. *Proc. Prussian Acad. Sci.*, *19*. Available at *arXiv:quant-ph/9903100 v2 15*.

Schwinger, J. (1960). Unitary operator bases. *Proceedings of the National Academy of Sciences*, *46*(4), 570–579.

Scott, A. J. (2006). Tight informationally complete quantum measurements. *Journal of Physics A: Mathematical and General*, *39*(43), 13507–13530.

Shafittee, A., Safinejad, F., & Naqsh, F. (2006). Information and the brukner-zeilinger interpretation of quantum mechanics: A critical investigation. *Foundations of Physics Letters*, *19*, 1–20.

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

Shore, J. & Johnson, R. (1980). Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Transactions on Information Theory*, *26*(1), 26–37.

Spengler, C., Huber, M., Brierley, S., Adaktylos, T., & Hiesmayr, B. C. (2012). Entanglement detection via mutually unbiased bases. *Phys. Rev. A*, *86*(2), 022311.

Sponar, S., Danner, A., Pecile, V., Einsidler, N., Demirel, B., & Hasegawa, Y. (2021). Experimental test of entropic noise-disturbance uncertainty relations for three-outcome qubit measurements. *Physical Review Research*, *3*(2), 023175.

Srinivas, M. D. (2003). Optimal entropic uncertainty relation for successive measurements in quantum information theory. *Pramana*, *60*, 1137–1152.

Sulyok, G., Sponar, S., Erhart, J., Badurek, G., Ozawa, M., & Hasegawa, Y. (2013). Violation of Heisenberg's error-disturbance uncertainty relation in neutron-spin measurements. *Physical Review A*, *88*(2), 022110.

Szymusiak, A. & Słomczyński, W. (2016). Informational power of the hoggar symmetric informationally complete positive operator-valued measure. *Physical Review A*, *94*(1), 012122.

Tsallis, C. (1988). Possible generalization of boltzmann-gibbs statistics. *Journal of Statistical Physics*, *52*, 479–487.

Uffink, J. (1995). Can the maximum entropy principle be explained as a consistency requirement? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, *26*(3), 223–261.

Uffink, J. (1996). The constraint rule of the maximum entropy principle. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, *27*(1), 47–79.

Uffink, J. (2001). Bluff your way in the second law of thermodynamics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, *32*(3), 305–394.

Uola, R., Costa, A. C. S., Nguyen, H. C., & Gühne, O. (2020). Quantum steering. *Reviewof Modern Physics*, *92*(1), 015001.

Urbanowski, K. (2020). Uncertainty Relations: Curiosities and Inconsistencies. *Symmetry*, *12*(10).

Van der Waerden, B. L. (1967). *Sources of Quantum Mechanics*. New York: Dover Publications.

Visser, M. (2013). Zipf's law, power laws and maximum entropy. *New Journal of Physics*, *15*(4), 043021.

Wang, D., Ming, F., Song, X. K., Ye, L., & Chen, J. L. (2020). Entropic uncertainty relation in neutrino oscillations. *The European Physical Journal C*, *80*(8), 800.

Wang, K. & Zheng, Z. J. (2021). Constructing Entanglement Witnesses from Two Mutually Unbiased Bases. *International Journal of Theoretical Physics*, *60*, 274–283.

Watrous, J. (2018). *The Theory of Quantum Information*. Cambridge University Press.

Wheeler, J. A. & Zurek, W. H. (1983). *Quantum Theory and Measurement*. New Jersey: Princeton University Press.

Wigner, E. (1963). The problem of measurement. *American Journal of Physics*, *31*(1), 6.

Wootters, W. K. (2006). Quantum Measurements and Finite Geometry. *Foundations of Physics*, *36*, 112–126.

Wootters, W. K. & Fields, B. D. (1989). Optimal state-determination by mutually unbiased measurements. *Annals of Physics*, *191*(2), 363.

Wu, S., Yu, S., & Mølmer, K. (2009). Entropic uncertainty relation for mutually unbiased bases. *Physical Review A*, *79*(2), 022104.

Xiong, T. P., Yan, L. L., Ma, Z. H., Zhou, F., Chen, L., Yang, W. L., Feng, M., & Busch, P. (2017). Optimal joint measurements of complementary observables by a single trapped ion. *New Journal of Physics*, *19*(6), 063032.

Yashin, V. I., Kiktenko, E. O., Mastiukova, A. S., & Fedorov, A. K. (2020). Minimal informationally complete measurements for probability representation of quantum dynamics. *New Journal of Physics*, *22*(10), 103026.

Zauner, G. (2011). Quantum designs: Foundations of a noncommutative design theory. *International Journal of Quantum Information*, *09*(01), 445–507.

Zhu, H. (2010). SIC POVMs and Clifford groups in prime dimensions. *Journal of Physics A: Mathematical and Theoretical*, *43*(30), 305305.

Zozor, S., Bosyk, G. M., & Portesi, M. (2014). General entropy-like uncertainty relations in finite dimensions. *Journal of Physics A: Mathematical and Theoretical*, *47*(49), 495302.