UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
CIENCIAS FÍSICAS Y ASTRONÓMICAS

# Nuevas propuestas en Tomografía de estados Cuánticos
# New Proposals in Quantum State Tomography

Profesor Guía: Dr. Aldo Delgado Hidalgo
Departamento de Física
Facultad de Ciencias Físicas y Matemáticas
Universidad de Concepción

**Tesis para optar al grado de Doctor en Ciencias Físicas de la Universidad de Concepción**

**ROBERTO BENJAMÍN SALAZAR VARGAS**
**CONCEPCIÓN - CHILE**
**2013**

Comisión Examinadora : Dr. Aldo Delgado Hidalgo

Dr. Carlos Saavedra

Dr. Guilherme Xavier

*Con cariño, dedico esto a mis padres*

*Sin matemáticas no se penetra hasta el fondo de la filosofía;*
*Sin filosofía no se llega al fondo de las matemáticas;*
*Sin las dos no se ve el fondo de nada.*

*Jean-Baptiste Bordan-Desmoulin*

# Contents

# Resumen

En esta tesis se propone nuevas herramientas teóricas y prácticas para el campo de investigación de tomografía de estados cuánticos.

Como introducción a esta tesis los capítulos 1 y 2 fijan nuestro enfoque de la mecánica cuántica y la teoría de probabilidad. Estos capítulos son esenciales, ya que definen también los conceptos básicos necesarios para entender las propuestas, así como los objetivos filosóficos de nuestra investigación.

En los capítulos 3 y 4, se introduce brevemente los conceptos de discriminación y tomografía cuántica de estados respectivamente. Al final de estos capítulos explicamos nuestros primeros resultados: una nueva herramienta teórica (los estados equidistantes), que ha sido útil en otras investigaciones de nuestros colleages y también un novel esquema de tomografía que combina el enfoque Inverción lineal con las técnicas de discriminación de estados. Este esquema tomográfico también tiene la característica novedosa de que es posible una reconstrucción probabilística del estado original sobre el sistema medido.

En el capítulo 5 se desarrolla la teoría de los marcos de reconstrucción y definimos el conjunto especial de los operadores conocidos como SIC-POVM. Las mediciones definidas por la SIC-POVMs permiten una fórmula de reconstrucción tomográfica que hace de la Tomografía de estados cuánticos más confiable y que también se puede utilizar para sustituir la regla de Born y cambiar el formalismo de la mecánica cuántica. Lamentablemente aún no hay prueba analítica de la existencia del SIC-POVMs de rango uno. Explicamos dos de nuestros intentos para dar una prueba analítica de la existencia desde SIC-POVM de dimensión arbitraria. Estos dos intentos no tuvieron éxito, pero muestran nuevas características de la estructura de la SIC-POVMs. Fueron publicados en revistas ISI.

En el capítulo 6 se propone una noción de optimización de un proceso de reconstrucción y encontramos las condiciones que un marco de reconstrucción debe satisfacer

para que sea óptimo en este sentido. Se demuestra que la SIC-POVMs y las bases mutuamente excluyentes (MUBS) son óptimas en este sentido. Se introduce una generalización de la SIC-POVMs conocida como SIC-POVM condicional (CSI-POVM) y muestran que también son óptimas para la reconstrucción de un subespacio del espacio de estados cuánticos. A continuación, se propone un nuevo conjunto de bases para la reconstrucción que generaliza al de MUBS, utilizando los estados equidistantes. La optimización de esta generalización es sólo una conjectura hasta ahora.

En el capítulo 7 se presenta el nuevo enfoque de la mecánica cuántica conocido como Bayesianismo Cuántico. También se exploran las consecuencias de este enfoque para la tomografía de estados cuánticos, es decir: el teorema cuántico de Finetti y la fórmula de reconstrucción de los SIC-POVM. A continuación, se prueba que los mismos objetivos alcanzados con la fórmula de los SIC-POVM se puede lograr por medio de los CSI-POVM, con la ventaja de que tenemos una prueba analítica de la existencia de los CSI-POVM para dimensiones que son el sucesor de la potencia de un primo. Exploramos las consecuencias de la correspondiente fórmula de reconstrucción de los CSI-POVM para el Bayesianismo Cuántico.

# Abstract

In this thesis we propose new theoretical and practical tools for the research field of Quantum State Tomography.

As an introduction to this thesis, Chapters 1 and 2 set out our approach to quantum mechanics and the theory of probability. These chapters are essential, and which also define the basic concepts necessary to understand proposals and philosophical goals of our research.

In Chapters 3 and 4, we briefly introduce the concepts of quantum state discrimination and quantum state tomography respectively. At the end of this chapters we explain our first results with a new theoretical tool (the equidistant states) that has been useful in many other research of our colleages and also a novel tomography scheme that combines the linear invertion approach with the techniques of quantum state discrimination. This tomographic scheme has also the novel feature that a probabilistic reconstruction of the measured state is possible.

In Chapter 5 we develop the theory of reconstruction frames and define the especial set of operators known as SIC-POVM. The measurements defined by SIC-POVMs allow a reconstruction tomographic formula that makes quantum state tomography more reliable and as well, can be used to replace Born's rule and change quantum mechanics formalism. Sadly there is still no analytic proof of the existence of SIC-POVMs of rank one. We explain two of our attemps to give an analytic proof to the existence of SIC-POVM in arbitrary dimension. This two attemps were not successful, but they show novel features of the structure of SIC-POVMs. They were published in ISI reports.

In Chapter 6 we give a notion of optimality for a reconstruction process and found the conditions that a reconstruction frame should satisfy to be optimal in this sense. We show that SIC-POVMs and mutually unbiased bases (MUBs) are optimal in this sense. We introduce a generalization of SIC-POVMs known as conditional SIC-POVM (CSI-POVM) and show that they are also optimal for reconstruction of a subspace of

the quantum state space. Then we propose a new set of bases for reconstruction that generalize the MUBs, using the equidistant states. The optimality of this generalization is a conjecture so far.

In Chapter 7 we introduce the novel approach to quantum mechanics known as Quantum Bayesianism. Also we explore the consequences of this approach for quantum state tomography, that is: the quantum de Finetti theorem and the SIC-POVM reconstruction formula. Then we prove that the same goals achieved with the SIC-POVM formula can be achieved by means of the CSI-POVM, with the advantage that we have an analytical proof for the existence of CSI-POVM for dimensions that are de successor of the power of a prime. We explore the consequences of the corresponding CSI-POVM reconstruction formula for Quantum Bayesianism.

# Publicaciones

## Publicaciones ISI

Esta tesis esta basada en las siguiente publicaciones:

- L. Roa and R. Salazar and C. Hermann-Avigliano and A. B. Klimov, *Conclusive Discrimination among N equidistant pure states*, *Physical Review A* **84** 014302, (2011).
- R. Salazar and A. Delgado, *Quantum tomography via unambiguous state discrimination*, *Physical Review A* **86** 012118, (2012).
- R. Salazar and D. Goyeneche and A. Delgadoa and C. Saavedra, *Constructing symmetric informationally complete positive-operator-valued measures in Bloch space*, *Physics Letters A* **376** 325 to 329, (2012).
- D. Goyeneche and R. Salazar and A. Delgado, *Characterization of fiducial states in prime dimensions via mutually unbiased bases*, *Phys. Scr.* 014031, (2013).

## Draft

Así como las actuales investigaciones:

- R. Salazar and L. Ruppert, *Optimal non-linear estimation and mutually unbiased bases generalization*, *Draft* (2013)
- R. Salazar and A. Delgado, *Quantum Bayesianism via conditional symmetrically informationally complete positive-operator valued measure*, *Draft* (2013)

## Otras publicaciones y trabajos pralelos

Durante el desarrollo de la tesis el autor también desarrollo una investigación paralela en el area de Mecánica de Nambu, con aplicaciones en Mecánica de Fluidos, Magnetohydrodinámica y electromagnetismo:

- R. Salazar and M. V. Kurgansky, *Nambu brackets in fluid mechanics and magnetohydrodynamics* , *Journal of Physics A: Mathematical and Theoretical* **43** 305501, (2010).

- R. Salazar and M. V. Kurgansky, *Nambu brackets for the electromagnetic field*, *http://arxiv.org/pdf/1011.5282* (2011)

# 1 Basics in Quantum Mechanics Formalism

Here we give the basic definitions and statements of quantum mechanics formalism within the context of quantum information to be the framework of the discussions and researches developed in the following chapters. The scheme has been taken from [1] with the necessary changes and rearrangements for a succinct presentation that deals with the requirements of this Thesis.

## 1.1 General approach to physics

Quantum Mechanics is a breakthrough in science, not only because of its formidable applications, but because it changes our way to understand Nature. The ontological problem to define nature is so important, because delimits the range of applicability of Physics (and with this the whole science). At the beginning Aristotle thought nature as an etiology, i.e. the study of the objects of the world and its causalities[2]. Then Galileo changed this conception, and Nature becomes the mathematical determination of events i.e. we substitute objects with mathematical models whose laws will determine the observed events [2, 3].

Now Quantum Mechanics, through the indeterminacy principle forces us to take account of events, not as independent elements of reality, but also connected with us in an unseparable way. This makes phenomena the elementary block of Nature. Phenomena is the aspect that objects offer to our senses [4], i.e. its measurements and Nature is the course of this measures; the law of phenomena [2].

Because of this, every time we face reality within physics three basic concept emerge to grasp the constitution of nature: *physical systems*, *physical process* and *observables*. A *physical system* is any phenomena, describable in finite space-time and its dynamical evolution (free, manipulated or both) is a *physical process* [6]while *observables* are any property we can measure in a physical system [5].

When measuring an observable various outcomes are simultaneously possible, but their plausibility shows to be different in general. To give a quantitative account of this plausibility is developed the concept of probability [7, 8]. Is the belief of the author that *probabilities* are just a way to extend our logical reasoning for cases when deductive reasoning is not possible due to a lack of enough information. This will be discussed in more detail in Chapter two where this conception of probabilities and its relation with *statistical frequency* will be clarified.

For what follows we just need to explain that *statistical frequency* is a particular kind of data which is the number of actual outcomes which its plausibility we are interested in, divided by the number of all the actual outcomes in the observable. A probability of a particular outcome is in general a function of the statistical frequency and our prior information.

It must be remarked that the mathematical framework here developed includes this general definition as well the more simplistic identification of statistical frequency limit (i.e. extrapolation) with the probability, which is the core belief of *frequentist interpretation* of probabilities. Having say that, then we call a *state* of a physical system the characterization of the probabilities of the various outcomes to every conceivable measuring of a particular observable of the physical system[5].

Any theory whose predictions are just probabilities about the results of different observables is called a *statistical theory* and such is the case of quantum mechanics. *Statistical experiments* are experiments in which the prediction power of a statistical theory is tested. They require to be repeated according to the same procedure as it can be set out in a detailed laboratory manual. Two kinds of procedures are required:

- *Preparation procedures:* Which prepare a certain kind of physical system in a distinguished state

- *Registration procedures:* Also known as *tests* are the procedures for measuring a particular observable.

A mathematical description of such a setup basically consist of two sets $\mathcal{S}$ , $\mathcal{E}$ and a map $\mathcal{S} \times \mathcal{E} \ni (\rho\, , \, A) \rightarrow \rho(A) \in [0, 1]$. The elements of $\mathcal{S}$ describe the states i.e. preparations, while the $A \in \mathcal{E}$ known as *effects*, represent all results of a measurement of any observable of the system. The probability to get the result represented by the effect $A$ on a system prepared in the state $\rho$ , is given by $\rho\,(A)$. This is a very general scheme applicable not only to quantum mechanics but also to a very broad class of statistical models, containing in particular classical probability theory. In order to make use of it we have to specify the precise structure of the sets $\mathcal{S}$ , $\mathcal{E}$ and $\rho\,(A)$ for the types of systems we want to discuss. Most of the following technical results that complete the mathematical structure required for this approach are taken from different chapters of the book-review *Fundamentals of quantum information theory* by Michael Keyl [1].

### 1.1.1 Operator Algebras

The scheme we are going to discuss is based on an algebra $\mathcal{A}$ [1] of bounded operators acting on a Hilbert space $\mathcal{H}$. More precisely $\mathcal{A}$ is a (closed) linear subspace of $\mathcal{B}\,(\mathcal{H})$, the algebra of bounded operators on $\mathcal{H}$, which contains the identity ($\mathbb{I} \in \mathcal{A}$) and is closed under products ($A,\, B \in \mathcal{A} \Rightarrow AB \in \mathcal{A}$) and adjoints ($A \in \mathcal{A} \Rightarrow A^* \in \mathcal{A}$). For simplicity

[1] An algebra $\mathcal{A}$ (or algebra over a field) is a vector space equipped with a bilinear product. If $V$ is a vector space a bilinear product is a function $\phi : V \times V \rightarrow V$ such that for any $v \in V$, the maps $w \rightarrow \phi(w, v)$ and $w \rightarrow \phi(v, w)$ are linear maps $V \rightarrow V$.

we will refer to each such $\mathcal{A}$ as an *observable algebra.* The key observation is now that each type of system we will study in the following can be completely characterized by its observable algebra $\mathcal{A}$, i.e. once $\mathcal{A}$ is known there is a systematic way to derive the sets $\mathcal{S}$ , $\mathcal{E}$ and the map $(\rho\,,\,A) \to \rho(A)$ from it. We frequently make use of this fact by referring to systems in terms of their observable algebra $\mathcal{A}$, or even by identifying them with their algebra and saying that $\mathcal{A}$ is the system.

Although $\mathcal{A}$ and $\mathcal{H}$ can be infinite dimensional in general, we will consider only finite dimensional Hilbert spaces, as long as nothing else is explicitly stated. Hence we can choose $\mathcal{H} = \mathbb{C}^d$ and $\mathcal{B}(\mathcal{H}) = \mathcal{M}_d(\mathbb{C})$ the algebra of $d \times d$ matrices. Since $\mathcal{A}$ is a subalgebra of $\mathcal{B}(\mathcal{H})$ it operates naturally on $\mathcal{H}$ and it inherits from $\mathcal{B}(\mathcal{H})$ the *operator norm* $\|A\| = sup_{\|\psi\|=1}\|A\psi\|$ and the *operator ordering* $A \geq B \Leftrightarrow \langle\psi,\,A\psi\rangle \geq \langle\psi,\,B\psi\rangle\ \forall\psi \in \mathcal{H}$. Now we can define:

$$\mathcal{S}(\mathcal{A}) = \{\rho \in \mathcal{A}^* |\, \rho \geq 0,\ \rho(\mathbb{I}) = 1\} \tag{1.1}$$

where $\mathcal{A}^*$denotes the *dual space* of $\mathcal{A}$, i.e. the set of all linear functionals on $\mathcal{A}$, and $\rho \geq 0$ means $\rho(A) \geq 0 \quad \forall A \geq 0$. Elements of $\mathcal{S}(\mathcal{A})$ describe the states of the system in question while effects are given by

$$\mathcal{E}(\mathcal{A}) = \{A \in \mathcal{A} \,|\, A \geq 0,\ A \leq \mathbb{I}\} \tag{1.2}$$

More generally we can look at $\rho(A)$ for an arbitrary $A$ as the *expectation value* of $A$ in a system of state $\rho$. Hence the idea behind equation (1.1) is to define states in terms of their expectation value functional.

Both spaces are *convex* i.e. $\rho,\,\sigma \in \mathcal{S}(\mathcal{A})$ and $0 \leq \lambda \leq 1$ implies that $\lambda\rho + (1-\lambda)\,\sigma \in \mathcal{S}(\mathcal{A})$ and similarly for $\mathcal{E}(\mathcal{A})$. The *extremal points* of $\mathcal{S}(\mathcal{A})$ respectively $\mathcal{E}(\mathcal{A})$, i.e. those elements which do not admit a proper convex decomposition,

$$x = \lambda y + (1-\lambda)z \Rightarrow \lambda = 1 \textit{ or } \lambda = 0 \textit{ or } y = z = x$$

play a distinguished role: the extremal points of $\mathcal{S}(\mathcal{A})$ are *pure states* and those of $\mathcal{E}(\mathcal{A})$ are the *propositions* of the system in question. The latter represents the effects which register a property with certainty.

### 1.1.2 Quantum systems

For quantum systems we have:

$$\mathcal{A} = \mathcal{B}(\mathcal{H}) \tag{1.3}$$

where we have chosen again $\mathcal{H} = \mathbb{C}^d$. The corresponding systems are called $d$-level systems, qudits or qubits when $d = 2$ holds. To avoid clumsy notations we frequently write $\mathcal{S}(\mathcal{H})$ and $\mathcal{E}(\mathcal{H})$ instead of $\mathcal{S}(\mathcal{B}(\mathcal{H}))$ and $\mathcal{E}(\mathcal{B}(\mathcal{H}))$. From equation (1.2) we immediately see that an operator $A \in \mathcal{B}(\mathcal{H})$ is an effect iff it is positive and bounded above by $\mathbb{I}$. An element $P \in \mathcal{E}(\mathcal{H})$ is a proposition iff $P$ is a projection operator ($P^2 = P$).

States are described in quantum mechanics usually by density matrices, i.e. positive and normalized trace class operators. To make contact to the general definition in equation (1.1) note first that $\mathcal{B}(\mathcal{H})$ is a Hilbert space with the Hilbert-Schmidt scalar product $\langle A,\, B\rangle = Tr\left(A^{*}B\right)$. Hence each linear functional $\rho \in \mathcal{B}(\mathcal{H})^{*}$ can be expressed in terms of a (trace class) operator $\tilde{\rho}$ by $A \to \rho(A) = Tr\left(\tilde{\rho}A\right)$. In this way each $\tilde{\rho}$ defines a unique functional $\rho$. If we start on the other hand with $\rho$ we can recover the matrix elements of $\tilde{\rho}$ from $\rho$ by $\tilde{\rho}_{kj} = Tr\left(\tilde{\rho}\,|j\rangle\langle k|\right) = \rho\left(|j\rangle\langle k|\right)$ , where $|j\rangle\langle k|$denotes the canonical basis of $\mathcal{B}(\mathcal{H})$. More generally we get for $\psi,\, \phi \in \mathcal{H}$ the relation $\langle \phi, \tilde{\rho}\psi\rangle = \rho\left(|\psi\rangle\langle\phi|\right)$, where $|\psi\rangle\langle\phi|$ now denotes the rank one operator which maps $\eta \in \mathcal{H}$ to $\langle\phi,\eta\rangle\,\psi$. In the following we drop the $\sim$ and use the same symbol for the operator and the functional whenever confusion can be avoided. Due to the same abuse of language we will interpret elements of $\mathcal{B}(\mathcal{H})^{*}$frequently as (trace class) operators instead of linear functionals (and write $Tr(\rho A)$ instead of $\rho(A)$ ). However we do not identify $\mathcal{B}(\mathcal{H})^{*}$ with $\mathcal{B}(\mathcal{H})$ in general, because the two different notations help to keep track of the distinction between spaces of states and spaces of observables.

Positivity of the functional $\rho$ implies positivity of the operator $\rho$ due to $0 \leq \rho\left(|\psi\rangle\langle\psi|\right) = \langle\psi,\rho\psi\rangle$ and the same holds for normalization: $1 = \rho(\mathbb{I}) = Tr\left(\rho\right)$. Hence we can identify the state space from equation (1.1) with the set of density matrices, as expected for quantum mechanics. Pure states of a quantum system are the one dimensional projectors. As usual we will frequently identify the density matrix $|\psi\rangle\langle\psi|$ with the wave function $\psi$ and call the latter in abuse of language a state.

### 1.1.3 Classical systems

The observable algebra $\mathcal{A}$ of such a system is the space:

$$\mathcal{A} = \mathcal{C}(X) = \{f \,:\, X \to \mathbb{C}\} \tag{1.4}$$

of complex valued functions on the finite set $X$ of *elementary events.* To interpret this as an operator algebra acting on a Hilbert space $\mathcal{H}$ choose an arbitrary but fixed orthonormal basis $|x\rangle$ , $x \in X$ in $\mathcal{H}$ and identify the function $f \in \mathcal{C}(X)$ with the operator $f = \sum_x f_x\,|x\rangle\langle x| \in \mathcal{B}(\mathcal{H})$ (we use the same symbol for the function and the operator, provided confusion can be avoided). Most frequently we have $X = \{1,\ldots,d\}$ and we can choose $\mathcal{H} = \mathbb{C}^d$ and the canonical basis for $|x\rangle$. Hence $\mathcal{C}(X)$ becomes the algebra of diagonal $d \times d$ matrices. Using equation (1.2) we see that $f \in \mathcal{C}(X)$ is an effect iff $0 \leq f_x \leq 1$, $\forall x \in X$. Physically we can interpret $f_x$ as the probability that the effect $f$ registers the elementary event $x$ This makes the distinction between propositions and fuzzy effects very transparent: $P \in \mathcal{E}(X)$ is a proposition iff we have either $P_x = 1$ or $P_x = 0$ for all $x \in X$. Hence the propositions $P \in \mathcal{C}(X)$ are in one to one correspondence with the subset $\omega_P = \{x \in X \,|\, P_x = 1\} \subset X$ which in turn describe the *events* of the system. Hence $P$ register the event $\omega_P$ with certainty, while a fuzzy effect $f < P$ does this with a probability less than one.

Since $\mathcal{C}(X)$ is finite dimensional and admits the distinguished basis $|x\rangle\langle x|$ , $x \in X$ it is naturally isomorphic to its dual $\mathcal{C}^{*}(X)$. More precisely: each linear functional

$\rho \in \mathcal{C}^*(X)$ defines and is uniquely defined by the function $x \to \rho_x = \rho(|x\rangle\langle x|)$ and we have $\rho(f) = \sum_x f_x \rho_x$. As in the quantum case we will identify the function $\rho$ with the linear functional and use the same symbol for both, although we keep the notation $\mathcal{C}^*(X)$ to indicate we are talking about states rather than observables.

Positivity of $\rho \in \mathcal{C}^*(X)$ is given by $\rho_x \geq 0$ for all $x$ and normalization leads to $1 = \rho(\mathbb{I}) = \rho\left(\sum_x |x\rangle\langle x|\right) = \sum_x \rho_x$. Hence $\rho$ to be a state, $\mathcal{C}^*(X)$ must be a *probability distribution* on $X$ and $\rho_x$ the probability that the elementary event $x$ occurs during statistical experiments with systems in the state $\rho$. More generally $\rho(f) = \sum_x f_x \rho_x$ is the probability to measure the effect $f$ on systems in the state $\rho$. If $P$ is a particular proposition, $\rho(P)$ gives the probability for the event $\omega_P$. The pure states of the system are the Dirac measures $\delta_x$, $x \in X$ ; with $\delta_x(|y\rangle\langle y|) = \delta_{xy}$. Hence each $\rho \in \mathcal{S}(X)$ can be decomposed in a unique way into a convex linear combination of pure states.

### 1.1.4 Observables

We can think an observable $E$ taking its values in a finite set $X$ as a map which associates to each possible outcome $x \in X$ the effect $E_x \in \mathcal{E}(\mathcal{A})$ (when $\mathcal{A}$ is the observable algebra of the system in question) which is true if $x$ is measured and false otherwise. If the measurement is performed on systems in the state $\rho$ we get for each $x \in X$ the probability $p_x = \rho(E_x)$ to measure $x$. Hence the set of $p_x$ should be a probability distribution on $X$, and implies that $E$ should be a *positive operator valued measure* (POVM) on $X$. Then we have the following definition:

Definition 1.1: *Consider an observable algebra $\mathcal{A} \subset \mathcal{B}(\mathcal{H})$ and a finite set $X$. A family $E = (E_x)_{x\in X}$ of effects in $\mathcal{A}$ (i.e. $0 \leq E_x \leq \mathbb{I}$ ) is called a* positive operator valued measure *(POVM) on $X$ if $\sum_{x\in X} E_x = \mathbb{I}$ (known as* completeness relation *) holds. If all $E_x$ are projections, $E$ is called* projection valued measure *(PVM).*

In quantum mechanics we know that observables are described by self adjoint operators on a Hilbert space $\mathcal{H}$. If $A$ is such observable operator then it has the form: $A = \sum_{\lambda\in\sigma(A)} \lambda P_\lambda$ where $\sigma(A)$ denotes the *spectrum* of $A$, i.e. the set of eigenvalues and $P_\lambda$ denotes the projection onto the corresponding eigenspace. Hence there is a unique PVM, $P = (P_\lambda)_{\lambda\in\sigma(A)}$ associated to $A$ which is called *spectral measure* of $A$. It is uniquely characterized by the property that *the expectation value* $\sum_{\lambda\in\sigma(A)} \lambda\rho(P_\lambda)$ of $P$ in the state $\rho$ is given for any state $\rho$ by $\rho(A) = Tr(\rho A)$. Hence the traditional way to define observables within quantum mechanics perfectly fits into the scheme just outlined. However it only covers the projection valued case and therefore admits no fuzziness. For this reason POVMs are sometimes called *generalized observables.*

## 1.2 Composite systems

Composite systems occur in many places in quantum information theory. In this cases we are allow to construct states and observables from the subsystems. In quantum mechanics this is done in terms of tensor products, and we will review in the following some of the most relevant material

### 1.2.1 Tensor products

Consider two (finite dimensional) Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$. To each pair of vectors $\psi_1 \in \mathcal{H}$, $\psi_2 \in \mathcal{K}$ we can associate a bilinear form $\psi_1 \otimes \psi_2$, called the *tensor product* of $\psi_1$ and $\psi_2$, by $\psi_1 \otimes \psi_2 \left(\phi_1 , \phi_2\right) = \langle \psi_1, \phi_1 \rangle \langle \psi_{2,} \phi_2 \rangle$. For two product vectors $\psi_1 \otimes \psi_2$ and $\eta_1 \otimes \eta_2$ their scalar product is defined by $\langle \psi_1 \otimes \psi_2, \, \eta_1 \otimes \eta_2 \rangle = \langle \psi_1, \eta_1 \rangle \langle \psi_{2,} \eta_2 \rangle$ and it can be shown that this definition extends in a unique way to the span of all $\psi_1 \otimes \psi_2$ which therefore defines the tensor product $\mathcal{H} \otimes \mathcal{K}$. If we have more than two Hilbert spaces $\mathcal{H}_j$, $j = 1, \ldots, N$ their tensor product $\mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_2$ can be defined similarly.

The tensor product $A_1 \otimes A_2$ of two bounded operators $A_1 \in \mathcal{B}\left(\mathcal{H}\right)$, $A_2 \in \mathcal{B}\left(\mathcal{K}\right)$ is defined first for product vectors $\psi_1 \otimes \psi_2 \in \mathcal{H} \otimes \mathcal{K}$ by $A_1 \otimes A_2 \left(\psi_1 \otimes \psi_2\right) = \left(A_1 \psi_1\right) \otimes \left(A_2 \psi_2\right)$and then extended by linearity. The space $\mathcal{B}\left(\mathcal{H} \otimes \mathcal{K}\right)$ coincides with the span of all $A_1 \otimes A_2$. If $\rho \in \mathcal{B}\left(\mathcal{H} \otimes \mathcal{K}\right)$ is not of product form (and of trace class for infinite dimensional $\mathcal{H}$ and $\mathcal{K}$) there is nevertheless a way to define restrictions to $\mathcal{H}$ respectively $\mathcal{K}$ called the *partial trace* of $\rho$. It is defined by the equation

$$Tr_{\mathcal{H}} \left[Tr_{\mathcal{K}} \left(\rho\right) A\right] = Tr_{\mathcal{H} \otimes \mathcal{K}} \left(\rho A \otimes \mathbb{I}\right) \; \forall A \in \mathcal{B}\left(\mathcal{H}\right) \tag{1.5}$$

If two orthonormal bases $\phi_1, \ldots, \phi_n$ and $\psi_1, \ldots, \psi_m$ are given in $\mathcal{H}$ respectively $\mathcal{K}$ we can consider the product basis $\phi_1 \otimes \psi_1, \ldots, \phi_n \otimes \psi_m$ in $\mathcal{H} \otimes \mathcal{K}$, and we can expand each $\Psi \in \mathcal{H} \otimes \mathcal{K}$ as $\Psi = \sum_{jk} \Psi_{jk} \phi_j \otimes \psi_k$ with $\Psi_{jk} = \langle \phi_j \otimes \psi_k, \, \Psi \rangle$. This procedure works for an arbitrary number of tensor factors. However, if we have exactly a twofold tensor product, there is a more economic way to expand $\Psi$, called *Schmidt decomposition* in which only diagonal terms of the form $\phi_j \otimes \psi_j$ appear.

Proposition 1.2: *For each element $\Psi \in \mathcal{H} \otimes \mathcal{K}$ there are orthonormal systems $\phi_1, \ldots, \phi_n$ and $\psi_1, \ldots, \psi_n$ (not necessarily bases because $n$ can be smaller than $\dim\left(\mathcal{H}\right)$ and $\dim\left(\mathcal{K}\right)$) of $\mathcal{H}$ and $\mathcal{K}$ respectively such that $\Psi = \sum_j \sqrt{\lambda_j} \phi_j \otimes \psi_j$ holds. The $\phi_j$ and $\psi_j$ are uniquely determined by $\Psi$. The expansion is called* Schmidt decomposition *and the numbers $\sqrt{\lambda_j}$ are the* Schmidt coefficients.

For a proof see Chapter 2, section 2 page 16 in [1]. As an immediate application of this result we can show that each mixed state $\rho \in \mathcal{B}\left(\mathcal{H}\right)^*$ (of the quantum system $\mathcal{B}\left(\mathcal{H}\right)$) can be regarded as a pure state on a larger Hilbert space $\mathcal{H} \otimes \mathcal{H}'$. We just have to consider the eigenvalue expansion $\rho = \sum_j \lambda_j \left|\phi_j\right\rangle \left\langle\phi_j\right|$ of $\rho$ and choose an arbitrary orthonormal system $\psi_j$, $j = 1, \ldots, n$ in $\mathcal{H}'$. Using Proposition 1.2 we get:

Corollary 1.3: *Each state $\rho \in \mathcal{B}\left(\mathcal{H}\right)^*$ can be extended to a pure state $\Psi$ on a larger system with Hilbert space $\mathcal{H} \otimes \mathcal{H}'$ such that $Tr_{\mathcal{H}'} \left(\left|\Psi\right\rangle \left\langle\Psi\right|\right) = \rho$ holds.*

### 1.2.2 Compound and hybrid systems

To discuss the composition of two arbitrary (i.e. classical or quantum) systems it is very convenient to use the scheme developed in subsection 1.1.1 and to talk about the two subsystems in terms of their observable algebras $\mathcal{A} \subset \mathcal{B}\left(\mathcal{H}\right)$ and $\mathcal{B} \subset \mathcal{B}\left(\mathcal{K}\right)$. The observable algebra of the composite system is then simply given by the tensor product of $\mathcal{A} \otimes \mathcal{B}$, i.e.

$$\mathcal{A} \otimes \mathcal{B} := span\left\{A \otimes B \mid A \in \mathcal{A}, \; B \in \mathcal{B}\right\} \subset \mathcal{B}\left(\mathcal{H} \otimes \mathcal{K}\right) \tag{1.6}$$

The dual of $\mathcal{A} \otimes \mathcal{B}$ is generated by product states, $(\rho \otimes \sigma)(A \otimes B) = \rho(A)\,\sigma(B)$ and we therefore write $\mathcal{A}^* \otimes \mathcal{B}^*$ for $(\mathcal{A} \otimes \mathcal{B})^*$.

We will consider the special cases arising from different choices for $\mathcal{A}$ and $\mathcal{B}$. If both systems are quantum $\mathcal{A} = \mathcal{B}(\mathcal{H})$ and $\mathcal{B} = \mathcal{B}(\mathcal{K})$ we get

$$\mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{K}) = \mathcal{B}(\mathcal{H} \otimes \mathcal{K}) \tag{1.7}$$

as expected. For two classical systems $\mathcal{A} = \mathcal{C}(X)$ and $\mathcal{B} = \mathcal{C}(Y)$ recall that elements of $\mathcal{C}(X)$ (respectively $\mathcal{C}(Y)$) are complex valued functions on $X$ (on $Y$). Hence the tensor product $\mathcal{C}(X) \otimes \mathcal{C}(Y)$ consists of complex valued functions on $X \times Y$, i.e. $\mathcal{C}(X) \otimes \mathcal{C}(Y) = \mathcal{C}(X \times Y)$. In other words states and observables of the composite system $\mathcal{C}(X) \otimes \mathcal{C}(Y)$ are given by probability distributions and random variables on the Cartesian product $X \times Y$.

If only one of the subsystems is classical and the other is quantum, we have a hybrid system. The elements of this observable algebra $\mathcal{C}(X) \otimes \mathcal{B}(\mathcal{H})$ can be regarded as operator valued functions on $X$, i.e. $X \ni x \to A_x \in \mathcal{B}(\mathcal{H})$ and $A_x$ is an effect iff $0 \leq A_x \leq \mathbb{I}$ holds for all $x \in X$. The elements of the dual $\mathcal{C}^*(X) \otimes \mathcal{B}^*(\mathcal{H})$ are in a similar way $\mathcal{B}^*(X)$ valued functions $X \ni x \to \rho_x \in \mathcal{B}^*(\mathcal{H})$ and $\rho$ is a state iff each $\rho_x$ is a positive trace class operator on $\mathcal{H}$ and $\sum_x \rho_x = \mathbb{I}$. The probability to measure the effect $A$ in the state $\rho$ is $\sum_x \rho_x(A)$.

### 1.2.3 Correlations

Let us now consider two effects $A \in \mathcal{A}$ and $B \in \mathcal{B}$ then $A \otimes B$ is an effect of the composite system $\mathcal{A} \otimes \mathcal{B}$. It is interpreted as the joint measurement of $A$ on the first and $B$ on the second subsystem, where the "yes" outcome means "both outcomes give yes". In particular $A \otimes \mathbb{I}$ means to measure $A$ on the first subsystem and to ignore the second one completely. If $\rho$ is a state of $\mathcal{A} \otimes \mathcal{B}$ we can define its *restrictions* by $\rho^{\mathcal{A}}(A) = \rho(A \otimes \mathbb{I})$ and $\rho^{\mathcal{B}}(A) = \rho(\mathbb{I} \otimes A)$. If both systems are quantum the restrictions of $\rho$ are the partial traces, while in the classical case we have to sum over the $\mathcal{B}$, respectively $\mathcal{A}$ variables. For two states $\rho_1 \in \mathcal{S}(\mathcal{A})$ and $\rho_2 \in \mathcal{S}(\mathcal{B})$ there is always a state $\rho$ of $\mathcal{A} \otimes \mathcal{B}$ such that $\rho_1 = \rho^{\mathcal{A}}$ and $\rho_2 = \rho^{\mathcal{B}}$ holds: We just have to choose the product state $\rho_1 \otimes \rho_2$. However in general we have $\rho \neq \rho^{\mathcal{A}} \otimes \rho^{\mathcal{B}}$ which means nothing else than $\rho$ also contains *correlations* between the two subsystem systems.

Definition 1.4: *A state $\rho$ of a bipartite system $\mathcal{A} \otimes \mathcal{B}$ is called* correlated *if there are some $A \in \mathcal{A}$, $B \in \mathcal{B}$ such that $\rho(A \otimes B) \neq \rho^{\mathcal{A}}(A)\,\rho^{\mathcal{B}}(B)$ holds.*

From this we see that $\rho = \rho_1 \otimes \rho_2$ implies $\rho(A \otimes B) = \rho_1(A)\,\rho_2(B) = \rho^{\mathcal{A}}(A)\,\rho^{\mathcal{B}}(B)$ hence $\rho$ is not correlated. If on the other hand $\rho(A \otimes B) = \rho^{\mathcal{A}}(A)\,\rho^{\mathcal{B}}(B)$ we get $\rho = \rho^{\mathcal{A}} \otimes \rho^{\mathcal{B}}$. Hence, the definition of correlations just given perfectly fits into our intuitive considerations.

An important issue in quantum information theory is the comparison of correlations between quantum systems on the one hand and classical systems on the other. Hence let us have a closer look on the state space of a system consisting of at least one classical subsystem.

PROPOSITION 1.5: *Each state $\rho$ of a composite system $\mathcal{A} \otimes \mathcal{B}$ consisting of a classical and an arbitrary system has the form:*

$$\rho = \sum_{j \in X} \lambda_j \rho_j^{\mathcal{A}} \otimes \rho_j^{\mathcal{B}} \tag{1.8}$$

*with positive weights $\lambda_j > 0$ and $\rho_j^{\mathcal{A}} \in \mathcal{S}(\mathcal{A})$, $\rho_j^{\mathcal{B}} \in \mathcal{S}(\mathcal{B})$.*

For a proof see Chapter 2, section 2 page 17 in [1]. If $\mathcal{A}$ and $\mathcal{B}$ are two quantum systems it still possible for them to be correlated in the way just described. We can simply prepare them with a classical random generator which triggers two preparations devices to produce systems in the states $\rho_j^{\mathcal{A}}$, $\rho_j^{\mathcal{B}}$ with probability $\lambda_j$. The overall state produced by this setup is obviously the $\rho$ from equation (1.8). However, the crucial point is that *not all* correlations of quantum systems are of this type. This is a consequence of the definition of pure states $\rho = |\Psi\rangle \langle\Psi| \in \mathcal{S}(\mathcal{H})$ : Since there is no proper convex decomposition of $\rho$, it can be written as in Proposition 1.5 iff $\Psi$ is a product vector, i.e. $\Psi = \phi \otimes \psi$. This observation motivates the following definition:

DEFINITION 1.6: *A state of the composite system $\mathcal{B}(\mathcal{H}_1) \otimes \mathcal{B}(\mathcal{H}_2)$ is called* separable *or* classical correlated *if it can be written as*

$$\rho = \sum_{j} \lambda_j \rho_j^{(1)} \otimes \rho_j^{(2)} \tag{1.9}$$

*with states $\rho_j^{(k)} \in \mathcal{B}(\mathcal{H}_k)$ and weights $\lambda_j > 0$ . Otherwise $\rho$ is called* entangled. *The set of all separable states is denoted by $\mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ or just $\mathcal{D}$ if $\mathcal{H}_1$ and $\mathcal{H}_2$ are understood.*

## 1.3 Channels

The purpose of this section is to provide a framework for the description of all the dynamical operations on the states of a physical systems.This is done in a way that we give account not only of the effect of a physical process over a physical system, but also on the state of knowledge of this system. The basic idea is to represent each processing step by a "channel", which converts input systems, described by an observable algebra $\mathcal{A}$ into output systems described by a possibly different algebra $\mathcal{B}$. Henceforth we will call $\mathcal{A}$ the *input algebra* and $\mathcal{B}$ the *output algebra* . If we consider e.g. the free time evolution, we need quantum systems of the same type on the input and the output side, hence in this case we have $\mathcal{A} = \mathcal{B} = \mathcal{B}(\mathcal{H})$ with an appropriately chosen Hilbert space $\mathcal{H}$. If on the other hand we want to describe a measurement we have to map quantum systems (the measured system) to classical information (the measuring result). Therefore we need in this example $\mathcal{A} = \mathcal{B}(\mathcal{H})$ for the input and $\mathcal{B} = \mathcal{C}(X)$ for the output algebra, where $X$ is the set of possible outcomes of the measurement.

Our aim now is to get a mathematical object which can be used to describe a channel. To this end consider effect $A \in \mathcal{B}$ of the output system. If we invoke first a channel which transforms $\mathcal{A}$ systems into $\mathcal{B}$ systems, and measure $A$ afterward on the output systems, we end up with a measurement of an effect $T(A)$ on the input systems. Hence we get a map $T\ :\ \mathcal{E}\left(\mathcal{B}\right) \rightarrow \mathcal{E}\left(\mathcal{A}\right)$ which completely describes the *channel.* Alternatively we can look at the states and interpret a channel as a map $T^{*}\ :\ \mathcal{S}\left(\mathcal{A}\right) \rightarrow \mathcal{S}\left(\mathcal{B}\right)$ which transforms $\mathcal{A}$ systems in the state $\rho \in \mathcal{S}\left(\mathcal{A}\right)$ into $\mathcal{B}$ systems in the state $T^{*}\left(\rho\right) \in \mathcal{S}\left(\mathcal{B}\right)$. To distinguish between both maps we can say that $T$ describes the channel in the *Heisenberg picture* and $T^{*}$ in the *Schrdinger picture.* On the level of the statistical interpretation both points of view should coincide of course, i.e. the probabilities $\left(T^{*}\rho\right)\left(A\right)$ and $\rho\left(TA\right)$ to get the result "yes" during an $A$ measurement on $\mathcal{B}$ systems in the state $T^{*}\rho$, respectively a $TA$ measurement on $\mathcal{A}$ systems in the state $\rho$, should be the same. We say also that the map $T^{*}\ :\ \mathcal{B}^{*} \rightarrow \mathcal{A}^{*}$ is *dual* to $T$, i.e. $T^{*}\rho\left(A\right) = \rho\left(TA\right)$ for all $\rho \in \mathcal{B}^{*}$ and $A \in \mathcal{A}$. Since $\left(T^{*}\rho\right)\left(A\right)$ is linear in $A$ we see that $T$ must be an *affine map* i.e. $T\left(\lambda_1 A_1 + \lambda_2 A_2\right) = \lambda_1 T\left(A_1\right) + \lambda_2 T\left(A_2\right)$ for each convex linear combination $\lambda_1 A_1 + \lambda_2 A_2$ of effects in $\mathcal{B}$, and this in turn implies that $T$ can be extended naturally to a *linear map,* which we will identify in the following with the channel itself, i.e. we say that $T$ *is* the channel.

### 1.3.1 Completely positive maps

Let us change now slightly our point of view and start with a linear operator $T\ :\ \mathcal{A} \rightarrow \mathcal{B}$. To be a channel, $T$ must map effects to effects, i.e. $T$ has to be positive: $T\left(A\right) \geq 0\,\forall A \geq 0$ and bounded from above by $\mathbb{I}$, i.e. $T\left(\mathbb{I}\right) \leq \mathbb{I}$. In addition it is natural to require that two channels in parallel are again a channel. More precisely, if two channels $T\ :\ \mathcal{A}_1 \rightarrow \mathcal{B}_1$ and $S\ :\ \mathcal{A}_2 \rightarrow \mathcal{B}_2$ are given we can consider the map $T \otimes S$ which associates to each $A \otimes B \in \mathcal{A}_1 \otimes \mathcal{A}_2$ the tensor product $T\left(A\right) \otimes S\left(B\right) \in \mathcal{B}_1 \otimes \mathcal{B}_2$. It is natural to assume that $T \otimes S$ is a channel which converts composite systems of type $\mathcal{A}_1 \otimes \mathcal{A}_2$ into $\mathcal{B}_1 \otimes \mathcal{B}_2$ systems. Hence $S \otimes T$ should be positive as well.

Definition 1.7: *Consider two observable algebras $\mathcal{A}$, $\mathcal{B}$ and a linear map $T\ :\ \mathcal{A} \rightarrow \mathcal{B} \subset \mathcal{B}\left(\mathcal{H}\right)$.*

1. *$T$ is called* positive *if $T\left(A\right) \geq 0$ for all positive $A \in \mathcal{A}$.*
2. *$T$ is called* completely positive *(cp) if $T \otimes Id\ :\ \mathcal{A} \otimes \mathcal{B}\left(\mathbb{C}^{n}\right) \rightarrow \mathcal{B}\left(\mathcal{H}\right) \otimes \mathcal{B}\left(\mathbb{C}^{n}\right)$ is positive for all $n \in \mathbb{N}$. Here $Id$ denotes the identity map on $\mathcal{B}\left(\mathbb{C}^{n}\right)$.*
3. *$T$ is called* unital *if $T\left(\mathbb{I}\right) = \mathbb{I}$ holds.*

If item 2 holds only for a fixed $n \in \mathbb{N}$ the map $T$ is called *n-positive.* This is obviously a weaker condition than complete positivity. However, n-positivity implies m-positivity for all $m \leq n$, and for $\mathcal{A} = \mathcal{B}\left(\mathbb{C}^{d}\right)$ complete positivity is implied by n-positivity, provided $n \geq d$ holds.

Let us consider now the question whether a channel should be unital or not. We have already mentioned that $T\left(\mathbb{I}\right) \leq \mathbb{I}$ must hold since effects should be mapped to effects. If $T\left(\mathbb{I}\right)$ is not equal to $\mathbb{I}$ we get $\rho\left(T\mathbb{I}\right) = T^{*}\rho\left(\mathbb{I}\right) < 1$ for the probability to measure the effect $\mathbb{I}$ on systems in the state $T^{*}\rho$, but this is impossible for channels which produce an output with certainty, because $\mathbb{I}$ is the effect which measures whether we have got an output. We will assume in the future that channels are unital if nothing else is explicitly stated.

### 1.3.2 The Stinespring theorem

Consider now channels between quantum systems, i.e. $\mathcal{A} = \mathcal{B}(\mathcal{H}_1)$ and $\mathcal{B} = \mathcal{B}(\mathcal{H}_2)$. A simple example (not necessarily unital) is given in terms of an operator $V : \mathcal{H}_1 \to \mathcal{H}_2$ by $\mathcal{B}(\mathcal{H}_1) \ni A \mapsto VAV^* \in \mathcal{B}(\mathcal{H}_2)$. A second example is the restriction to a subsystem which is given by (Heisenberg picture) $\mathcal{B}(\mathcal{H}) \ni A \mapsto A \otimes \mathbb{I}_{\mathcal{K}} \in \mathcal{B}(\mathcal{H} \otimes \mathcal{K})$. Finally the composition $S \circ T = ST$ of two channels is again a channel. The following theorem, which is the most fundamental structural result about cp maps, says that each channel can be represented as a composition of these two examples.

THEOREM 1.8(STINESPRING DILATION THEOREM): *Every completely positive map $T : \mathcal{B}(\mathcal{H}_1) \mapsto \mathcal{B}(\mathcal{H}_2)$ has the form:*

$$T(A) = V^* (A \otimes \mathbb{I}_{\mathcal{K}}) V \tag{1.10}$$

*with an additional Hilbert space $\mathcal{K}$ and an operator $V : \mathcal{H}_2 \to \mathcal{H}_1 \otimes \mathcal{K}$. Both (i.e. $\mathcal{K}$ and $V$) can be chosen such that the span of all $(A \otimes \mathbb{I}_{\mathcal{K}}) V\phi$ with $A \in \mathcal{B}(\mathcal{H}_1)$ and $\phi \in \mathcal{H}_2$ is dense in $\mathcal{H}_1 \otimes \mathcal{K}$. This particular decomposition is unique (up to a unitary equivalence) and called the minimal decomposition. Also the minimal $\mathcal{K}$ satisfies $dim(\mathcal{K}) \leq [dim(\mathcal{H}_1)]^2 \, dim(\mathcal{H}_2)$.*

For a proof see [9].

### 1.3.3 Ideal and noisy channels

We can think in a channel $T : \mathcal{B}(\mathcal{H}) \mapsto \mathcal{B}(\mathcal{H})$ as an *operation* that performs the transmission of quantum information over long distances, where $T^*(\rho)$ is the quantum information which will be received when a system characterized by $\rho$ was sent. Ideally we would prefer those channels which do not affect the information at all, i.e. $T = \mathbb{I}$, or as the next best choice, a $T$ whose action can be undone by a physical device, i.e. $T$ should be invertible and $T^{-1}$ is again a channel. The Stinespring theorem (Theorem 1.8) immediately shows that this implies $T^*\rho = U\rho U^*$ with a unitary $U$. This means that the system carrying the information do not interact with the environment. We will call such a kind of channel an *ideal channel*. In real situations, however interaction with the environment, i.e. additional unobservable degrees of freedom, can not be avoided. The general structure of such a *noisy channel* is given by:

$$T^*(\rho) = Tr_{\mathcal{K}} (U (\rho \otimes \rho_o) U^*) \tag{1.11}$$

where $U : \mathcal{H} \otimes \mathcal{K} \to \mathcal{H} \otimes \mathcal{K}$ is a unitary operator describing the common evolution of the system $\mathcal{H}$ and the environment $\mathcal{K}$. Here also $\rho_o \in \mathcal{S}(\mathcal{K})$ is the initial state of the environment.

It must be remarked the connection between channel and the orthodox formulations of quantum dynamics. The connection between both formulations lies in the fact that evolution of states is given by the *Shrödinger equation* :

$$i\hbar \frac{\partial \rho}{\partial t} = [H, \rho] \tag{1.12}$$

where $H$ is the *Hamiltonian operator* of the physical system, and the solution in terms of the unitary operator

$$U = \mathbb{I} - \frac{i}{\hbar}\int_{t_{in}}^{t_{out}} dt_1 H\left(t_1\right) + \sum_{k=2}^{\infty}\frac{1}{k!}\left(\frac{-i}{\hbar}\right)^k \int_{t_{in}}^{t_{out}} dt_1 \int_{t_{in}}^{t_1} dt_2 \cdots \int_{t_{in}}^{t_{k-1}} dt_k H(t_1)\ldots H(t_k) \tag{1.13}$$

where $t_{out} > t_1 > \ldots > t_k > t_{in}$ (i.e. this is a *Dyson series*), is:

$$\rho\left(t_{out}\right) = U\rho\left(t_{in}\right)U^* \tag{1.14}$$

and this is equivalent to our ideal channel by the identification of $T^*\rho$ with $\rho\left(t_{out}\right)$ and $\rho$ with $\rho\left(t_{in}\right)$ respectively. This is nothing more than saying that the output state $T^*\rho$ is the same as the state of the system at the output time $t_{out}$ and the input state $\rho$ is the state of the system at the input time $t_{in}$.

To finish we notice that Hamiltonian operators of a quantum system are obtained by the application of *first and second quantization* on the classical Hamiltonian function of the physical system.

## 1.4 Remarks on our Informational approach

In this Chapter we have introduced the basic and fundamental concepts of quantum mechanics within an informational scheme developed by Michael Keyl in [1]. Since, Michael Keyl proposes that quantum information is a *different kind* of information than classical one, we have introduced some changes at the beginning of the sections to take a different point of view. For us the information provided by quantum mechanical phenomena is not of a different kind, but is an information that we get from a *different inference process* than the information we get from classical systems. The information we get in quantum systems is always probabilities and as in classical systems they provide us a decision-making tool of the same kind, what is different are the rules for characterize those probabilities, i.e. their inference.

The scheme gives a unifying framework for different parts of our work with quantum mechanics, first by showing how to describe preparations and test for a system, only by choosing the appropriate operator algebra, then by demonstrating that application of tensor products and partial trace over the chosen operator algebra gives account of the operations of composition and selection of subsystems, respectively. Finally, the concept of channels gives a framework for treating the information of all kind of systems under all kind of processes, being the process due to physical changes in the physical system or to selection and/or composition of the information about the system.

In Chapter two we will go further in this direction, by taking a Bayesian interpretation of probabilities and showing this way that more than just the results of quantum mechanics is just manipulation of our state of knowledge (Chapter 7), clearing the path to find the true teachings of quantum mechanics about reality.

# 2 Bayesianism: A view of probability within scientific inference.

## 2.1 Scientific method and Scientific knowledge

Science is an activity of human beings, whose purpose is the acquisition of knowledge from reality´s phenomena and the construction of tools (material or theoretical) for predicting and controlling such phenomena. This purposes are not exclusive from science, but also other philosophical systems have been proposed for making research about reality itself, some of them which also claim to go beyond the scope of science [14, 4, 10, 12]. What is particular of science is its method and because of that any statement about a phenomena can claim to be scientific, only when its truth can be analyzed within the scientific method. This is why we have first to review the steps of scientific method, as well its conceptions of truth and objectivity.

*Truth* in science is *the correspondence between concepts (or statements) and facts*[4]. Since this correspondence is demonstrated through measurement of certain quantities whose values are determined by the concepts or statements, it´s at the same time limited by the error associated to the measurement instruments. This is why scientific truth of any statement is always bounded for error and quantities value have meaning only in the significant digits. Measurement instruments are extensions of our senses, which also give a quantification of senses experiences. Because of this, scientific truth also assume human beings and its basics experiences, as well the *a priori* conditions that allow such experiences. [1]

On the other extreme is the problem of the ontology of facts, which is the link between reality and us [2]. This also is a subtle matter upon which the applicability of science depends in particular any conception of objectivity. *Objectivity* is the claim that the facts related to properties of objects are independent of the mind process of the observers and then a particularity within the ontology of facts problem.

Scientific community (i.e. the set of people who is professionally engaged with

[1] Neurophysiology and physiology use scientific method in its analysis and that´s why it would be a circular reasoning to try to justify scientific truth on their analysis, their valid claims are then only allowed in the description of the basic experiences and the phenomena associated, but never as a justification of them or its a priori conditions.

Science) has not reached a consensus about this philosophical matters that underlies any scientific experience, or which philosophical system and methods should be trusted the research in this topics. There has been recognized some implications that are essential for scientific work as well a classification of the fundamental quantities that the a priori conditions allow us to observe, experience and measure (distance, time intervals, mass, charge, etc.) is provided by the community to make the scientific results reliable and the scientific discussion meaningful. Despite this philosophical complications that certainly require more research if we attempt to have a deeper understanding of reality, science have shown until now to be the widest and most reliable source of knowledge about reality, the humanity has ever created.

This review of scientific truth and scientific objectivity has as a result the implications that *our state of knowledge is always incomplete* and *experimenters with the same state of knowledge should reach the same conclusions* respectively. Which are the most relevant implications for the present work.

The seven steps of scientific method are [11]:

1. *Identification of the Scientific Problem*
2. *Preliminary hypothesis*
3. *Acquisition of additional facts (data)*
4. *Formulation of a model (system of hypothesis)*
5. *Deduction of the consequences of the model*
6. *Experimental testing of the consequences*
7. *Development of applications*

Now we will review them shortly:

**1** ***Identification of the Scientific Problem.*** The first step comes from two possibilities: a conflict between our concepts or theory and a fact, i.e. the finding of a *false* concept or theory, or a new fact that can´t be explained from the accepted conceptions or theories about the phenomena involved, i.e. the finding of an *incomplete* conception or theory. This findings are what we understand for a *Scientific Problem* and the first step in the method is about a clear exposition of it.

**2** ***Preliminary hypothesis.*** After the identification of the scientific problem, the fact which leads to the problem and the concepts or theories involved in it must lead to some hypothesis about the phenomena. This new set of hypothesis do not necessarily consist in a theory, but should be used to realize new observations or experimentation's about the phenomena.

**3** ***Acquisition of additional facts (data).*** The observations and experimentation's implied by the preliminary hypothesis must be realized and the new data obtained from them displayed and organized.

4 ***Formulation of a model (system of hypothesis).*** Now, the new data should be analyzed and a new model or theory created to give account of all this new facts, this means the data must fit or correspond with some of the hypothesis of the model or others that are a result of the model hypothesis and combination rules. There is no rule for the creation of the model, only the requirement just outlined.

5 ***Deduction of the consequences of the model.*** After the creation of a model able to give account of all the data, the model must be examined and find new consequences from its hypothesis, rules and inner structure that predict facts that have already not being observed, this is what we know as a *prediction*.

6 ***Experimental testing of the consequences.*** The predictions of the model require the development of an experimental setup that should reproduce the necessary conditions to observe the predicted facts, if the subsequent observations are in correspondence with consequences of the model, this will validate it within its known range of applicability. Otherwise we are now in the presence of a new scientific problem and we have to start with the scientific method again, until we obtain a validated model.

7 ***Development of applications.*** When we are in possession of a validated model, this should be used as the theoretical background for the development of technology by scientists or engineers.

The goal of Bayesianism is to give an unifying procedure for making inferences, including deductive and inductive inference in the same mathematical model. This was also the original goal of probability theory, and this is why Bayesianism goes back to the roots of probability theory, reinterpreting it as extended logic. This is why Bayesianism and its statistical methods emerge as a help within scientific method by providing a powerful tool in the inference process required in steps 2, 5 and 6. Because of this, Bayesian interpretation is not only of philosophical interest, but also a source of practical developments for all disciplines interested in the use of the scientific method.

## 2.2 Bayesian approach to probabilities and statistical inference

Around the fourth century BC, Aristotle [13] recognized that deductive inference can be analyzed into repeated applications of the *strong syllogisms*:

**(a)** Major premise:{ If $A$ is true, then $B$ is true}, Minor premise: {$A$ is true} / Conclusion: {Therefore $B$ is true}

**(b)** Major premise:{ If $A$ is true, then $B$ is true}, Minor premise: {$B$ is false} / Conclusion: {Therefore $A$ is false}

Most of the scientific theories and mathematics have work on their inner structure and inferences through the Aristotelian deductive reasoning, based on this two syllogisms. They can be applied multiple times, and the conclusions will be strong as the premises are. On the other hand, Inductive or plausible inference has been proposed for the case where, because of incomplete information, Aristotelian deductive reasoning is not possible. Plausible inference is based on the *weak syllogisms*:

**(c)** Major premise:{ If $A$ is true, then $B$ is true}, Minor premise: {$B$ is true} / Conclusion: {Therefore $A$ becomes more plausible}

**(d)** Major premise:{ If $A$ is true, then $B$ is true}, Minor premise: {$A$ is false} / Conclusion: {Therefore $B$ becomes less plausible}

This extension of logic has shown problematic, due to the absence of a quantitative account of this conclusions, making obscure its meaning, their inference not reliable (or without a measure of reliability) and subject of a long history of philosophical criticism [16, 17, 15, 18]. This has been the reason why, until now, deductive reasoning has been the only unquestionable method of inference in science and mathematics, as in the words of Sir James Clerk Maxwell (1850): "*The actual science of logic is conversant at present only with things either certain, impossible or entirely doubtful, none of which (fortunately) we have to reason on.*". It must be remarked that, still this being the case, inductive reasoning is clearly used in science and mathematics in an informal way [23, 24],

The early work on probability theory by James Bernoulli (1713,[20]), Rev. Thomas Bayes (1763, [19]) and Pierre Simon Laplace (1774, [21]), viewed probability as an extension of logic, where the inclusion of plausible reasoning could be introduced by defining probability as the measure of plausibility required to give a quantitative account of the weak syllogisms. Unfortunately, Laplace failed to give convincing arguments to show why this definition of probability uniquely required the basic rules for manipulating probabilities. The frequentist definition of probability was introduced to satisfy this point, but in the process, eliminated the interpretation of probability as extended logic. So, the two conceptions of probability are:

**1)** ***Frequentist statistical inference:*** $p\,(A)$ : long-run relative frequency with which $A$ occurs in identical repeats of an experiment. "$A$" is restricted to propositions about *random variables.*

**2)** ***Bayesian statistical inference:*** $p\,(A|I)$: a real number measure of the plausibility of a proposition or hypothesis $A$, given (conditional on) the truth of the information represented by $I$. "$A$" can be any logical proposition, not restricted to propositions about random variables.

The new resurgence of Bayesian approach is due to the achievement of finding the missing arguments for the requirement of the basic rules for manipulating probabilities. This is done through a set of *desiderata* known as such, rather than *axioms,* because they do not assert anything true, but only state desirable goals.

The Desiderata of Bayesian probability theory are:

**I.** Degrees of plausibility are represented by real numbers

**II.** As new information supporting the truth of a proposition is supplied, the number which represents the plausibility will increase continuously and monotonically. Also, the deductive limit must be obtained where appropriate.

**III.** Consistency[2]

**(a)** ***Structural consistency:*** If a conclusion can be reasoned out in more than one way, every possible way must lead to the same result.

**(b)** ***Propriety:*** The theory must take account of all information, provided it is relevant to the question.

**(c)** ***Jaynes consistency:*** Equivalent states of knowledge must be represented by equivalent plausibility assignments.

From this desiderata is it possible to develop probability theory as an extension of logic [7, 8]. This was done through the works of G. Polya [24], R.T. Cox [25] and E.T. Jaynes (most remarkably in the book *Probability Theory-The Logic of Science* [7]). Their greatest achievement was the deduction of the basic rules for manipulating probabilities: *Sum rule* and *Product rule*, just by requiring to a monotonous, continuous function $0 \leq p\left(\cdot\right) \leq 1$ of the degree of plausibility $\left(A|I\right)$ of an statement or hypothesis $A$ (assuming information $I$), to satisfy the previously stated desiderata. This is:

**(1)** ***Sum Rule:***

$$p\left(A|I\right) + p\left(\bar{A}|I\right) = 1$$

**(2)** ***Product Rule:***

$$p\left(A,B|I\right) = p\left(A|I\right)p\left(B|A,I\right) = p\left(B|I\right)p\left(A|B,I\right)$$

Where $A,B$ states for the conjunction of statements $A$ and $B$ and $\bar{A}$ is the negation of $A$. From both rules is it possible to derive the *extended sum rule*:

$$p\left(A+B|I\right) = p\left(A|I\right) + p\left(B|I\right) - p\left(A,B|I\right)$$

Here $A+B$ is the disjunction of statements $A$ and $B$. In the special case of $A$ and $B$ mutually exclusive, we have the *generalized sum rule*:

$$p\left(A+B|I\right) = p\left(A|I\right) + p\left(B|I\right)$$

And in particular, from the product rule (a rearrangement of the two sides of the equation) follows the *Bayes´ Theorem*:

$$p\left(H_i|D,I\right) = \frac{p\left(H_i|I\right)p\left(D|H_i,I\right)}{p\left(D|I\right)} \tag{2.1}$$

[2] In my opinion this *consistency* desiderata express more a desire of *equivalency* between statements rather than requiring logical consistency. Of course, they are required to ensure consistency, but this one is assumed only in the construction of the sum rule.

We have rewrite the variables with the purpose to explain in more detail their usual meaning in the Bayesian analysis. This is: $H_i \equiv$ "proposition assenting the truth of a hypothesis of interest", $I \equiv$ "proposition representing our prior information", $D \equiv$ "proposition representing data", $p\left(D|H_i, I\right) =$ "probability of obtaining data $D$, if $H_i$ and $I$ are true", also known as the *likelihood function* $\mathcal{L}\left(H_i\right)$, $p\left(H_i|I\right) =$ "prior probability of $H_i$", $p\left(H_i|D, I\right) =$ "posterior probability of $H_i$", $p\left(D|I\right) = \sum_i p\left(H_i|I\right) p\left(D|H_i, I\right)$, normalization factor which ensures $\sum_i p\left(H_i|D, I\right) = 1$ , also known as total probability factor.

Now we are in position to show, how Bayesian inference includes all kinds of syllogism. First we start with the strong syllogisms (a) and (b). The major premise translated to Boolean algebra (i.e. the algebra of Aristotelian logic, [22]) is $A, B = A$ and the minor premise is "$A$ is true". Then by writing down the product rule for (a) :

$$\begin{aligned} p\left(A, B|I\right) = & \; p\left(A|I\right) p\left(B|A, I\right) \\ \rightarrow \quad p\left(B|A, I\right) = & \; \frac{p\left(A, B|I\right)}{p\left(A|I\right)} \end{aligned}$$

Using the prior information $I \equiv$ "$A, B = A$"

$$\begin{aligned} \rightarrow p\left(A, B|I\right) &= p\left(A|I\right) \\ \rightarrow p\left(B|A, I\right) &= 1 \end{aligned}$$

which means that $B$ is certain when $A$ is true under the major premise. Since minor premise assure us that $A$ is true, then $B$ must be certain. Now for (b) we have the same major premise as in (a) and minor premise "$B$ is false".

$$\begin{aligned} p\left(A, \bar{B}|I\right) = & \; p\left(\bar{B}|I\right) p\left(A|\bar{B}, I\right) \\ \rightarrow \quad p\left(A|\bar{B}, I\right) = & \; \frac{p\left(A, \bar{B}|I\right)}{p\left(\bar{B}|I\right)} \end{aligned}$$

Using the prior information $I \equiv$ "$A, B = A$"

$$\begin{aligned} \rightarrow p\left(A, \bar{B}|I\right) &= 0 \\ \rightarrow p\left(A|\bar{B}, I\right) &= 0 \end{aligned}$$

which means that $A$ is impossible when $\bar{B}$ (the negation of $B$) is true under the major premise. Since minor premise assure us that $\bar{B}$ is true, then $A$ must be impossible. For the weak syllogism (c) we start with the Bayes´ theorem:

$$p\left(A|B, I\right) = \frac{p\left(A|I\right) p\left(B|A, I\right)}{p\left(B|I\right)}$$

Using the prior information $I \equiv$ "$A, B = A$"

$$\rightarrow p\left(B|A,I\right)=1$$

Also, by the definition of plausibility function $p\left(\cdot\right)$ we have

$$p\left(B|I\right)\leq 1$$

Then, substituting into Bayes´ theorem gives:

$$p\left(A|B,I\right)\geq p\left(A|I\right)$$

which means that $A$ is more plausible when $B$ is true under major premise, than in its prior plausibility. Since minor premise assure that $B$ is certain, the plausibility of $A$ is increased. For the weak syllogism (d) we start again with the Bayes´ theorem:

$$p\left(B|\bar{A},I\right)=\frac{p\left(B|I\right)p\left(\bar{A}|B,I\right)}{p\left(\bar{A}|I\right)}$$

Based on the same prior information syllogism (c) gives $p\left(A|B,I\right)\geq p\left(A|I\right)$ so using the sum rule we have:

$$\begin{aligned}\rightarrow 1-p\left(\bar{A}|B,I\right) &\geq 1-p\left(\bar{A}|I\right)\\ \rightarrow p\left(\bar{A}|B,I\right) &\leq p\left(\bar{A}|I\right)\\ \rightarrow \frac{p\left(\bar{A}|B,I\right)}{p\left(\bar{A}|I\right)} &\leq 1\end{aligned}$$

Substituting into Bayes´ theorem:

$$p\left(B|\bar{A},I\right)\leq p\left(B|I\right)$$

which means that $B$ is less plausible when $\bar{A}$ is true under the major premise, than in its prior plausibility. Since the minor premise states that $\bar{A}$ is true, we have that the plausibility of $B$ decreases. This shows how Bayesianism includes deductive, as well plausible reasoning in its plausibility model, unifying in this way the different inference procedures.

To finish the section we remark that in Bayesianism, a probability is a representation of our state of knowledge of the real world. A frequency is a factual property of the real world that we measure or estimate. One of the strengths of Bayesian inference is the ability to incorporate relevant prior information in the analysis. In Bayesian inference, we can readily incorporate frequency information using Bayes´ theorem and by treating it as data. In general, probabilities change when we change our state of knowledge; frequencies do not.

## 2.3 Examples of Bayes theorem application

This examples have been taken from the chapter one of the book, “Bayesian Logical Data Analysis For The Physical Sciences - A Comparative Approach With Mathematica” [8].

### 2.3.1 Model selection

Here we analyze a simple model comparison problem using Bayes theorem. We start by stating our prior information $I$ and the new data $D$.

$I$ stands for:

1. Model $M_1$ predicts a star´s distance, $d_1 = 100$ light years (ly).

2. Model $M_2$ predicts a star´s distance, $d_2 = 200$ ly.

3. The uncertainty $e$ in distance measurements is described by a Gaussian distribution of the form:

$$p\left(e\,|\,I\right) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{e^2}{2\sigma^2}\right)$$

   where $\sigma = 40$ ly.

4. There is no current basis for preferring $M_1$ over $M_2$, so we set $p\left(M_1|I\right) = p\left(M_2|I\right) = 1/2$ .

$D \equiv$ “The measured distance is $d = 120$ ly.”

The prior information tells us that the hypothesis space of interest consist of models (hypotheses) $M_1$ and $M_2$. We proceed by writing down Bayes theorem for each hypothesis, e.g.,

$$\begin{aligned} p\left(M_1|D,I\right) &= \frac{p\left(M_1|I\right)p\left(D|M_1,I\right)}{p\left(D|I\right)} \\ p\left(M_2|D,I\right) &= \frac{p\left(M_2|I\right)p\left(D|M_2,I\right)}{p\left(D|I\right)} \end{aligned}$$

Since we are interested in comparing the two models, we will compute the *odds ratio*, equal to the ratio of the posterior probabilities of the two models. We will abbreviate the odds ratio of model $M_1$ to model $M_2$ by the symbol $O_{12}$.

$$O_{12} = \frac{p\left(M_1|D,I\right)}{p\left(M_2|D,I\right)} = \frac{p\left(M_1|I\right)p\left(D|M_1,I\right)/p\left(D|I\right)}{p\left(M_2|I\right)p\left(D|M_2,I\right)/p\left(D|I\right)} = \frac{p\left(D|M_1,I\right)}{p\left(D|M_2,I\right)} \tag{2.2}$$

where the equal terms cancel each other. Then, to evaluate the likelihood $p\left(D|M_1,I\right)$ , we note that in this case we are assuming $M_1$ is true. In that case, the only reason the measured $d$ can differ from the prediction $d_1$ is because of measurement uncertainties $e$. We can thus write $d = d_1 + e$ or $e = d - d_1$. Since $d_1$ is determined by the model, it it certain and so the probability $p\left(D|M_1, I\right)$of obtaining the measured distance is equal to the probability of the error $e\left\{d_1\right\}$. Thus we can write:

$$p\left(D|M_1, I\right) = p\left(e\left\{d_1\right\}|I\right) = \frac{1}{\sqrt{2\pi\sigma}}\exp\left(-\frac{\left(d-d_1\right)^2}{2\sigma^2}\right) = 0.00880 \tag{2.3}$$

Similarly we can write for model $M_2$:

$$p\left(D|M_2, I\right) = p\left(e\left\{d_2\right\}|I\right) = \frac{1}{\sqrt{2\pi\sigma}}\exp\left(-\frac{\left(d-d_2\right)^2}{2\sigma^2}\right) = 0.00135 \tag{2.4}$$

The relative likelihood of the two models is proportional to the heights of the two Gaussian probability distributions at the location of the measured distance. Substitution of equations 2.3 and 2.4 into 2.2 gives us an odd ratio of 6.52 in favor of model $M_1$.

### 2.3.2 Incorporating frequency information

A 1996 newspaper article reported that doctors in Toronto were concerned about a company selling an unapproved mail-order HIV saliva test. According to laboratory tests, the false positive rate for this test was 2.3% and the false negative rate was 1.4% (i.e., 98.6% reliable based on testing of people who actually have the disease).

In this example, suppose a new deadly disease is discovered for which there is no known cause but a saliva test is available with the above specifications. We will refer to this disease by the abbreviation UD, for unknown disease. You have no reason to suspect you have UD but decide to take the test anyway and test positive. What is the probability that you really have the disease? Here is a Bayesian analysis of this situation. For the purpose of this analysis, we will assume that the incidence of the disease in a random sample of the region is $10^{-4}$ .

Let $H \equiv$ "You have UD", $\overline{H} \equiv$ "You do not have UD", $D_1 \equiv$ "You test positive for UD", $I_1 \equiv$ "No known cause for the UD, $p\left(D_1|H, I_1\right) = 0.986$ , $p\left(D_1|\overline{H}, I_1\right) = 0.023$, incidence of UD in population is $10^{-4}$."

The starting point for any Bayesian analysis is to write down Bayes theorem,

$$p\left(H|D_1, I_1\right) = \frac{p\left(H|I_1\right)p\left(D_1|H, I_1\right)}{p\left(D_1|I_1\right)} \tag{2.5}$$

Since $p\left(D_1|I_1\right)$ is a normalization factor, which ensures $\sum_i p\left(H_i|D_1, I_1\right) = 1$, we can write:

$$p\left(D_1|I_1\right) = p\left(H|I_1\right)p\left(D_1|H, I_1\right) + p\left(\overline{H}|I_1\right)p\left(D_1|\overline{H}, I_1\right) \tag{2.6}$$

In words, this latter equation stands for

$$\begin{aligned}\begin{pmatrix}\text{prob. of a}\\ \text{+ test}\end{pmatrix} &= \begin{pmatrix}\text{prob. you}\\ \text{have UD}\end{pmatrix}\times\begin{pmatrix}\text{prob. of a + test,}\\ \text{when you have UD}\end{pmatrix}\\ &\quad+\begin{pmatrix}\text{prob. you don´t}\\ \text{have UD}\end{pmatrix}\times\begin{pmatrix}\text{prob. of a + test}\\ \text{when you don´t have UD}\end{pmatrix}\\ &= \begin{pmatrix}\text{Incidence of UD}\\ \text{in population}\end{pmatrix}\times\begin{pmatrix}\text{reliability}\\ \text{of test}\end{pmatrix}+\begin{pmatrix}\text{1-incidence}\\ \text{of UD}\end{pmatrix}\times\begin{pmatrix}\text{false positive}\\ \text{rate}\end{pmatrix}\end{aligned}$$

and then,

$$p\left(H|D_1,I_1\right)=\frac{10^{-4}\times 0.986}{10^{-4}\times 0.986+0.9999\times 0.023}=0.0042 \tag{2.7}$$

Thus, the probability you have the disease is 0.4% and not 98.6%.

## 2.4 Advantages of the Bayesian approach

Some advantages that can be deduced from our presentation are:

1. Provides a simple and rational approach for answering any scientific question involving plausible reasoning for a given state of information.

2. Calculates probabilities of hypothesis directly

3. Incorporate relevant prior information through Bayes theorem. This is one of the great strengths of Bayesian analysis. For data with a high signal-to-noise ratio, a Bayesian analysis can frequently yield many orders of magnitude improvement in model parameter estimation, through the incorporation of relevant prior information about the signal model.

Other advantages that we will not review here are:

1. Provides a way of eliminating *nuisance parameters* (i.e. parameters that are unimportant or uninteresting for the analysis) through the procedure of *marginalization.* For some problems the marginalization can be performed analytically, permitting certain calculations to become computationally tractable [8].

2. Provides a factor in model comparison which automatically quantifies *Occam´s razor.* In any given model. this quantitative Occam´s razor helps us to identify and eliminate those variables that are not really needed to explain the phenomenon [8].

3. Provides a way for incorporating the effects of *systematic errors* arising from both the measurement operation and theoretical model predictions [8].

## 2.5 Implication for quantum mechanics

In Chapter one we developed an informational scheme to quantum mechanics that is general enough to include other statistical models, and in particular classical probability theory. This chapter goes one step forward by giving those statistical results the Bayesian interpretation and this has deep implications over our scheme and quantum mechanics. If our scheme include classical probability, and this last one is an extension of logic, the scheme must also be an inference tool of the same kind. We have seen that in the informational scheme quantum mechanics comes out from the consideration of the underlying operator algebra $\mathcal{B}\left(\mathcal{H}\right)$ which is richer in structure (and includes also) than the classical algebra $\mathcal{C}\left(X\right)$ of complex valued functions on a finite set. From this we conclude that quantum mechanics should be a *new* extension of our logic inference tool. This changes our point of view about quantum mechanics to a *Instrumentalistic* one, changing the question of *how classical world comes out from quantum rules?*, to *how we extend classical inference rules to a quantum inference rules?*. In Chapter 7 we will present this new approach of quantum mechanics known as *Quantum Bayesianism* and also our contributions in this line of research.

# 3 State discrimination

The purpose of this chapter is to give a brief introduction to the topic of quantum state discrimination [1] and the contributions of the thesis author within this topic.

Quantum state discrimination (QSD) is the following problem: given a quantum system known to be prepared in one of a finite number of possible states $\{\rho_i\}_{i=1}^{N}$ with *a priori* probabilities $\{p_i\}_{i=1}^{N}$, *what is the best measurement to determine the actual state in which the system was prepared?*.

Perfect discrimination is only possible when the states are mutually orthogonal. Just in this case is possible to find an observable operator $A = \sum_{\lambda\in\sigma(A)} \lambda P_\lambda$ (as it was described in subsection 1.1) whose spectral measure $P = (P_\lambda)_{\lambda\in\sigma(A)}$ elements match with the states to be discriminated (i.e. $\sigma(A) = \{\rho_i\}_{i=1}^{N}$ ) because eigenstates are always orthogonal sets, hence this spectral measure gives the desired PVM.

In the general (i.e. non-orthogonal set of states $\{\rho_i\}$) case there are several figures of merit which may be optimized, each leading to a different strategy to solve the problem. QSD in particular is relevant to quantum key distribution (QKD) [27]; security of QKD relies on the existence of states which cannot be perfectly discriminated by an eavesdropper. Any quantum information protocol has a read-out stage, where the user wishes to obtain some classical information about the result of the quantum information task. This may be thought of as a problem in QSD.

## 3.1 Minimum error discrimination

The first criteria we are going to study for the problem of QSD is to minimize the probability of making an error in identifying the state. We begin with the special case where the state is known to be one of two possible pure states, $|\psi_o\rangle, |\psi_1\rangle$, with associated probabilities $p_o$, $p_1 = 1 - p_o$. If outcome $i = 0, 1$, associated with the effect $\Pi_i$ is taken to indicate that the state was $|\psi_i\rangle$, the probability of making an error in determining

[1] We use as a reference for the introduction to QSD the article [26]

the state is given by

$$\begin{aligned} P_{err} &= P\left(\psi_o\right) P\left(1|\psi_o\right) + P\left(\psi_1\right) P\left(1|\psi_1\right) \\ &= p_o \left\langle\psi_o\right| \Pi_1 \left|\psi_o\right\rangle + p_1 \left\langle\psi_1\right| \Pi_o \left|\psi_1\right\rangle \\ &= p_o - Tr\left[\left(p_o \left|\psi_o\right\rangle\left\langle\psi_o\right| - p_1 \left|\psi_1\right\rangle\left\langle\psi_1\right|\right)\Pi_o\right] \end{aligned} \tag{3.1}$$

where in the last line we have used the completeness relation $\Pi_o + \Pi_1 = \mathbb{I}$. This expression takes its minimum value when the second term reaches a maximum, which in turns is achieved if $\Pi_o$ is a projector onto the positive eigenstate of the operator $p_o \left|\psi_o\right\rangle\left\langle\psi_o\right| - p_1 \left|\psi_1\right\rangle\left\langle\psi_1\right|$. Note that two pure states define a two dimensional space, without loss of generality we can choose an orthogonal basis $\{|0\rangle, |1\rangle\}$ of this space such that the components of each state in this basis are real. Thus we can express $\left|\psi_o\right\rangle, \left|\psi_1\right\rangle$ as follows:

$$\begin{aligned} \left|\psi_o\right\rangle &= \cos\left(\theta\right)|0\rangle + \operatorname{sen}\left(\theta\right)|1\rangle \\ \left|\psi_1\right\rangle &= \cos\left(\theta\right)|0\rangle - \operatorname{sen}\left(\theta\right)|1\rangle \end{aligned} \tag{3.2}$$

and in this basis the operator $p_o \left|\psi_o\right\rangle\left\langle\psi_o\right| - p_1 \left|\psi_1\right\rangle\left\langle\psi_1\right|$ has (in the computational basis $\{|0\rangle, |1\rangle\}$) the following matrix representation:

$$\begin{aligned} p_o \left|\psi_o\right\rangle\left\langle\psi_o\right| - p_1 \left|\psi_1\right\rangle\left\langle\psi_1\right| &= \begin{pmatrix} \left(p_o - p_1\right)\cos^2\left(\theta\right) & \left(p_o + p_1\right)\cos\left(\theta\right)\operatorname{sen}\left(\theta\right) \\ \left(p_o + p_1\right)\cos\left(\theta\right)\operatorname{sen}\left(\theta\right) & \left(p_o - p_1\right)\operatorname{sen}^2\left(\theta\right) \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2}\left(p_o - p_1\right)\left(1 + \cos\left(2\theta\right)\right) & \frac{1}{2}\operatorname{sen}\left(2\theta\right) \\ \frac{1}{2}\operatorname{sen}\left(2\theta\right) & \frac{1}{2}\left(p_o - p_1\right)\left(1 - \cos\left(2\theta\right)\right) \end{pmatrix} \end{aligned} \tag{3.3}$$

The eigenvalues of $p_o \left|\psi_o\right\rangle\left\langle\psi_o\right| - p_1 \left|\psi_1\right\rangle\left\langle\psi_1\right|$ can be calculated directly from equation (3.3), whose characteristic equation:

$$\left[\frac{1}{2}\left(p_o - p_1\right)\left(1 + \cos\left(2\theta\right)\right) - \lambda\right]\left[\frac{1}{2}\left(p_o - p_1\right)\left(1 - \cos\left(2\theta\right)\right) - \lambda\right] - \frac{1}{4}\operatorname{sen}^2\left(2\theta\right) = 0$$

has solutions:

$$\lambda_\pm = \frac{1}{2}\left(p_o - p_1 \pm \sqrt{1 - 4p_o p_1 \cos^2\left(2\theta\right)}\right)$$

From this, the minimum probability of making an error is then given by the so called *Helstrom bound* [30]:

$$P_{err} = p_o - \lambda_+ = \frac{1}{2}\left(1 - \sqrt{1 - 4p_o p_1 \left|\left\langle\psi_o|\psi_1\right\rangle\right|^2}\right) \tag{3.4}$$

and the optimal measurement is simply a PVM on the eigenvectors of (3.3).

### 3.1.1 Minimum error conditions

In the general case of $N$ possible states $\{\rho_i\}$ with associated *a priori* probabilities $\{p_i\}$, the aim is to minimize the expression:

$$P_{err} = \sum_{i=1}^{N} p_i \sum_{j \neq i} Tr\left(\rho_i \Pi_j\right) \tag{3.5}$$

or equivalently to maximize

$$P_{corr} = 1 - P_{err} = \sum_{i} p_i Tr\left(\rho_i \Pi_i\right) \tag{3.6}$$

Necessary and sufficient conditions for realizing a minimum error measurement were originally given by Holevo [28], and by Yuen, Kennedy and Lax [29], see also [30, 31]. One way of proving the conditions is by using semi-definite programming techniques. The conditions for minimum error are:

$$\begin{aligned} &\Pi_j\left(p_j\rho_j - p_k\rho_k\right)\Pi_k = 0\,\forall j,\,k \\ &\sum_i p_i\rho_i\Pi_i - p_j\rho_j \geq 0\,\forall j \end{aligned} \tag{3.7}$$

For any set of states and preparation probabilities there will exist at least one minimum error measurement with effects satisfying this conditions. These conditions allow us to verify whether a given measurement is optimal for discriminating a given set of states. Unfortunately they do not give the optimal measurement in an arbitrary case, which is not known in general. However, minimum error discrimination is a semi-definite program problem and may be solved efficiently numerically. Further, several bounds on the minimum probability of error are known.

## 3.2 Unambiguous state discrimination

Suppose again that we wish to discriminate between the two pure states given by equation (3.2), occurring with *a priori* probabilities $p_o$, $p_1$. Consider the PVM:

$$\begin{aligned} \Pi_{\mho} &= \left|\psi_1\right\rangle\left\langle\psi_1\right| \\ \Pi_o &= \left(\text{sen}\left(\theta\right)\left|0\right\rangle + \cos\left(\theta\right)\left|1\right\rangle\right)\left(\text{sen}\left(\theta\right)\left\langle 0\right| + \cos\left(\theta\right)\left\langle 1\right|\right) \end{aligned} \tag{3.8}$$

If outcome $\mho$, associated with the operator $\Pi_{\mho}$ is realized, we cannot say for sure what state was prepared. However, note that $\left\langle\psi_1\right|\Pi_o\left|\psi_1\right\rangle = 0$, and thus when outcome 0 corresponding to the POVM element $\Pi_o$, is realized, we can say for certain that the state was $\left|\psi_o\right\rangle$. Thus, by allowing for measurement outcome $\mho$, which does not lead us to identify any state, we can construct a measurement which sometimes allows us to determine unambiguously which state was prepared, that´s why this method is known as *unambiguous state discrimination* (USD). This measurement however only ever identifies state $|\psi_o\rangle$, ideally we would like to design a measurement which can identify either state unambiguously, at the cost of sometimes giving an inconclusive result. This kind of measurement is indeed possible and the generalized formalism was developed by Ivanovic [32], Diecks [33] and Peres [34].

Consider therefore the operators

$$\begin{aligned}\Pi_o =& a_o \left(\mathrm{sen}\left(\theta\right)|0\rangle + \cos\left(\theta\right)|1\rangle\right)\left(\mathrm{sen}\left(\theta\right)\langle 0| + \cos\left(\theta\right)\langle 1|\right)\\ \Pi_1 =& a_1 \left(\mathrm{sen}\left(\theta\right)|0\rangle - \cos\left(\theta\right)|1\rangle\right)\left(\mathrm{sen}\left(\theta\right)\langle 0| - \cos\left(\theta\right)\langle 1|\right)\end{aligned} \tag{3.9}$$

chosen such that $\langle\psi_o|\,\Pi_1\,|\psi_o\rangle = \langle\psi_1|\,\Pi_o\,|\psi_1\rangle = 0$, and where $0 \leq a_o,\, a_1 \leq 1$. Thus when outcome $i = 0, 1$ is realized, we can say for sure that the corresponding state was $|\psi_i\rangle$ with certainty. But, unless $|\psi_o\rangle\,,\,|\psi_1\rangle$ are orthogonal, there is no choice of $a_o$, $a_1$ such that these form a complete measurement, and thus an inconclusive outcome is needed, associated with the operator:

$$\Pi_{\mho} = \mathbb{I} - \Pi_o - \Pi_1 \tag{3.10}$$

The probability of occurrence of the inconclusive result is given by

$$\begin{aligned}P\left(\mho\right) &= p_o\,\langle\psi_o|\,\Pi_{\mho}\,|\psi_o\rangle + p_1\,\langle\psi_1|\,\Pi_{\mho}\,|\psi_1\rangle\\ &= p_o\left(1 - \langle\psi_o|\,\Pi_o\,|\psi_o\rangle\right) + p_1\left(1 - \langle\psi_1|\,\Pi_1\,|\psi_1\rangle\right)\\ &= 1 - \mathrm{sen}^2\left(2\theta\right)\left(p_o a_o + p_1 a_1\right)\end{aligned} \tag{3.11}$$

and the unambiguous discrimination strategy may be further optimized by minimizing this probability, subject to the constraints $a_o,\, a_1 \geq 0$, $\Pi_{\mho} \geq 0$. $P\left(\mho\right)$ is a monotonically decreasing function of $a_o$ and $a_1$, thus the minimum value lies at the boundary of the allowed domain, defined by $\Pi_{\mho} \geq 0$. In the computational basis, $\Pi_{\mho}$ has the matrix representation:

$$\Pi_{\mho} = \begin{pmatrix} 1 - \left(a_o + a_1\right)\mathrm{sen}^2\theta & \left(a_1 - a_o\right)\cos\theta\,\mathrm{sen}\theta \\ \left(a_1 - a_o\right)\cos\theta\,\mathrm{sen}\theta & 1 - \left(a_o + a_1\right)\cos^2\theta \end{pmatrix} \tag{3.12}$$

The optimal $P\left(\mho\right)$ was first given by Jaeger and Shimony [35]. The optimal measurement is given by equations (3.9, 3.10) with:

$$\begin{aligned} a_o =& \frac{1 - \sqrt{\frac{p_1}{p_o}}\cos\left(2\theta\right)}{\mathrm{sen}^2\left(\theta\right)}\\ a_1 =& \frac{1 - \sqrt{\frac{p_o}{p_1}}\cos\left(2\theta\right)}{\mathrm{sen}^2\left(\theta\right)}\end{aligned} \tag{3.13}$$

giving

$$P\left(\mho\right) = 2\sqrt{p_o p_1}\cos\left(2\theta\right) \tag{3.14}$$

without loss of generality we can choose $p_o > p_1$, so when $\sqrt{\frac{p_o}{p_1}}\cos(2\theta) > 1$ (3.9) and (3.10) no longer defines a physical measurement; the optimal measurement then is simply the PVM given by (3.8). In this case $|\psi_1\rangle$ always gives the inconclusive result, and the probability of failure is $P(\mho) = p_o\,|\langle\psi_o|\psi_1\rangle|^2 + p_1$. Thus for $p_o$ much bigger than $p_1$, the optimal strategy is the one which rules out the less probable state, in contrast to the minimum error measurement, which in this regime approximately identifies or rules out the more probable state.

### 3.2.1 Ancillary USD

A different USD scheme can be implemented by the use of ancillary systems. Consider again the states given by equation (3.2), occurring with a priori probabilities $p_o$, $p_1$ and an ancillary two dimensional system of orthonormal basis $\{|\mu\rangle, |\nu\rangle\}$. We consider the ancillary system in a known state $|\Lambda\rangle$, then by coupling and applying a unitary $U$ over the composed system:

$$\begin{aligned} |\Psi_o\rangle_{ab} &= U\left(|\psi_o\rangle_a |\Lambda\rangle_b\right) = &\cos(\alpha_o)\,|0\rangle_a |\mu\rangle_b + \text{sen}\,(\alpha_o)\,|\phi\rangle_a |\nu\rangle_b \\ |\Psi_1\rangle_{ab} &= U\left(|\psi_1\rangle_a |\Lambda\rangle_b\right) = &\cos(\alpha_1)\,|1\rangle_a |\mu\rangle_b + \text{sen}\,(\alpha_1)\,|\phi\rangle_a |\nu\rangle_b \end{aligned} \tag{3.15}$$

Where $|\phi\rangle$ is a state in the first Hilbert space and its components depend on the states under discrimination and the unitary operation. First we notice that since $U$ is a unitary operator, inner product must be preserved:

$$\cos(2\theta) = \langle\psi_o|\psi_1\rangle_a \langle\Lambda|\Lambda\rangle_b = \langle\Psi_o|\Psi_1\rangle_{ab} = \text{sen}\,(\alpha_o)\,\text{sen}\,(\alpha_1)$$

$$\cos(2\theta) = \text{sen}\,(\alpha_o)\,\text{sen}\,(\alpha_1) \tag{3.16}$$

The USD scheme must discriminate between states $|\Psi_o\rangle_{ab}, |\Psi_1\rangle_{ab}$ because unitary $U$ transforms $|\psi_k\rangle_a |\Lambda\rangle_b$ into $|\Psi_k\rangle_{ab}$, so finding state $|\Psi_k\rangle_{ab}$ on the composite system will imply state $|\psi_k\rangle_a$ on the first system. We apply the PVM:

$$\begin{aligned} \Pi_\mu &= \mathbb{I} \otimes |\mu\rangle\langle\mu| \\ \Pi_\nu &= \mathbb{I} \otimes |\nu\rangle\langle\nu| \end{aligned} \tag{3.17}$$

If effect $\Pi_\mu$ gives the "yes" answer, then we apply the PVM:

$$\begin{aligned} \Pi_o &= |0\rangle\langle 0| \otimes \mathbb{I} \\ \Pi_1 &= |1\rangle\langle 1| \otimes \mathbb{I} \end{aligned} \tag{3.18}$$

Where effect $\Pi_k$ allow us to identify state $|\Psi_k\rangle_{ab}$ because just this state has projections on state $|k\rangle_a$. Since effect $\Pi_\nu$ gives the inconclusive result the probability of failure is:

$$\begin{aligned} P(\mho) &= p_o \langle\Psi_o|\,\Pi_\nu\,|\Psi_o\rangle + p_1 \langle\Psi_1|\,\Pi_\nu\,|\Psi_1\rangle \\ &= p_o \text{sen}^2(\alpha_o) + p_1 \text{sen}^2(\alpha_1) \end{aligned} \tag{3.19}$$

optimizing $P(\mho)$ under the constraint (3.16) we find that the optimum $\alpha_o$, $\alpha_1$ are given by:

$$\begin{aligned}
\operatorname{sen}(\alpha_o) &= \sqrt{\cos(2\theta)}\sqrt[4]{\frac{p_1}{p_o}} \\
\operatorname{sen}(\alpha_1) &= \sqrt{\cos(2\theta)}\sqrt[4]{\frac{p_o}{p_1}}
\end{aligned} \tag{3.20}$$

with the probability of failure

$$P(\mho) = 2\sqrt{p_o p_1}\cos(2\theta) \tag{3.21}$$

The method can be resume as follows: 1) we couple a known two dimensional ancillary system, 2) we apply a unitary operation such that each state has a projection on a orthonormal basis that uniquely determines it and a common inconclusive projection, 3) we apply a PVM (3.17) that discriminates between conclusive and inconclusive spaces and 4) we apply a PVM (3.18) that discriminates among the corresponding orthonormal basis. Also we can mix both PVM in one POVM:

$$\begin{aligned}
\Pi_o &= |0\rangle\langle 0| \otimes |\mu\rangle\langle\mu| \\
\Pi_1 &= |1\rangle\langle 1| \otimes |\mu\rangle\langle\mu| \\
\Pi_\nu &= \mathbb{I} \otimes |\nu\rangle\langle\nu|
\end{aligned} \tag{3.22}$$

## 3.3 Equidistant states

In our paper *Conclusive discrimination among N equidistant states* [38] we introduce a novel set of states suitable for theoretical research (due to the few parameters involved in its definition). This set have been the starting point for research of different authors [39, 40], being a useful theoretical tool in QSD, Quantum Cloning and Quantum Tomography. In [38] the states where introduced showing their usefulness for QSD, but they will show to be of importance also for one of the last research proposals of this thesis (Chapter 6).

Following [38, 39], a set of *equidistant states* $A_N(\alpha)$ is a set of $N$ normalized pure states such that:

$$A_N(\alpha) \equiv \left\{ |\alpha_1\rangle, |\alpha_2\rangle, \ldots, |\alpha_N\rangle \;:\; \langle\alpha_k|\alpha_{k'}\rangle = \alpha \,\forall\, k < k' \right\} \tag{3.23}$$

Since the inner product between pure states is a measure of the distinguishability between this states, we say that two pairs of states with the same inner product are equally *distant* [2] and since our set is characterized for having the same inner product pairwise is that we call them a set of equidistant states.

[2] *Here the right concept is distinguishability and not distance, but we mint the term equidistant because it keeps the intuitive idea of the set while being more economic than equidistinguishable.*

But distinguishability for a set of more than two states is not completely characterized by their inner products, but for their linear independence [37, 36]. The set $A_N\left(\alpha\right)$ is linearly independent (LI) iff the equation:

$$\sum_{k=1}^{N} A_k \left|\alpha_k\right\rangle = \mathbf{0} \tag{3.24}$$

implies that all the $N$ coefficient $A_k = 0$, otherwise is a linearly dependent (LD) set of states. If we apply each $\left\langle\alpha_j\right|$ on equation (3.24) we get a system of equations:

$$\sum_{k=1}^{N} A_k \left\langle\alpha_j|\alpha_k\right\rangle = 0 \, \forall j = 1,\ldots,N \tag{3.25}$$

in the $A_k$. If this system is invertible then all $A_k = 0$, this is so when the Gram determinant:

$$\begin{aligned}\det\left(\mathcal{D}_{N\times N}\right) = \det & \begin{pmatrix} 1 & \alpha & \alpha & \cdots & \alpha \\ \alpha^* & 1 & \alpha & \cdots & \alpha \\ \alpha^* & \alpha^* & 1 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \alpha \\ \alpha^* & \alpha^* & \alpha^* & \cdots & 1 \end{pmatrix}_{N\times N} \\ & = \frac{\alpha\left(1-\alpha^*\right)^N - \alpha^*\left(1-\alpha\right)^N}{\alpha-\alpha^*}\end{aligned} \tag{3.26}$$

is non zero ($\det\left(\mathcal{D}_{N\times N}\right)$ is always higher or equal to zero). So if $\det\left(\mathcal{D}_{N\times N}\right) > 0$, $A_N\left(\alpha\right)$ is LI and if $\det\left(\mathcal{D}_{N\times N}\right) = 0$, $A_N\left(\alpha\right)$ is LD. In general the overlap $\alpha = \left|\alpha\right| e^{i\theta}$ of $N$ LI equidistant states must satisfy the constraint:

$$0 \leq \left|\alpha\right| < \left|\alpha_\theta\right| \tag{3.27}$$

with

$$\left|\alpha_\theta\right| = \frac{\mathrm{sen}\left(\frac{\pi-\theta}{N}\right)}{\mathrm{sen}\left(\theta + \frac{\pi-\theta}{N}\right)} \tag{3.28}$$

Here the angle $\theta$ must be evaluated within the interval $[0, 2\pi[$ and we have the results of the case $\alpha \in \mathbb{R}$ taken the limits $\theta = 0$ and $\theta = \pi$, where the states $A_N\left(\alpha\right)$ are L.I. iff:

$$-\frac{1}{N-1} < \alpha < 1 \tag{3.29}$$

For $\alpha = 1$ all the states are the same and when $\alpha = -1/(N-1)$ they are LD. For the general case the contour defined by $|\alpha| = |\alpha_\theta|$ the set of states is LD and within it (i.e. the convex region defined by the contour) is LI.

In [38] a constructive method for representing the equidistant states associated with an overlap $\alpha = |\alpha| e^{i\theta}$ was presented, but in [39] a more compact representation was introduced and it will be stated as the *canonical representation* of equidistant states. This is:

$$|\alpha_j\rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \sqrt{\lambda_k} \left(\omega_k^j\right)^* |k\rangle \tag{3.30}$$

The coefficients $\lambda_k$ are eigenvalues of the matrix $\mathcal{D}_{N\times N}$ and are given by:

$$\lambda_k = 1 - |\alpha| \frac{\operatorname{sen}\left(\theta + \frac{k\pi - \theta}{N}\right)}{\operatorname{sen}\left(\frac{k\pi-\theta}{N}\right)} \tag{3.31}$$

and fulfill the identity,

$$\sum_{k=0}^{N-1} \lambda_k = N \tag{3.32}$$

Also, the complex phases $\omega_k$ are the coefficients entering in the matrix which diagonalizes $\mathcal{D}_{N\times N}$ . These phases are given by

$$\omega_k = e^{\frac{2i}{N}(\theta - k\pi)} \tag{3.33}$$

In [38] we studied the USD of $N$ equidistant states, finding the probability of success in the case of equal *a priori* probabilities for each state preparation by an ancillary USD. The success probability under the ancillary USD $P_{usd}$ in this case is:

$$P_{usd} = 1 - |\alpha| \frac{\operatorname{sen}\left(\frac{\pi-\theta}{N}\right)}{\operatorname{sen}\left(\theta + \frac{\pi-\theta}{N}\right)} \tag{3.34}$$

We found also that the inconclusive effect associated with such USD measurement projects the set of LI equidistant states into the set of LD equidistant states with an overlap of the same phase $\theta$ ($\neq 0$) as that of the LI set. Hence, we can prepare a minimum error discrimination protocol for the LD equidistant states in the case of an inconclusive result on the USD measurement. The probability of success under minimal error protocol $P_{me}$ for the inconclusive space is:

$$P_{me} = \frac{1}{1 + (N-1)\left(\frac{\operatorname{sen}\left(\frac{\pi-\theta}{N}\right)}{\operatorname{sen}\left(\theta + \frac{\pi-\theta}{N}\right)}\right)^2} \tag{3.35}$$

Taking both into account, the success probability $P_{cs}$ for the complete scheme is:

$$\begin{aligned} P_{cs} &= P_{usd} + (1 - P_{usd})\, P_{me} \\ P_{cs} &= 1 - \frac{|\alpha|\,(N-1)\left(\frac{\operatorname{sen}\left(\frac{\pi-\theta}{N}\right)}{\operatorname{sen}\left(\theta+\frac{\pi-\theta}{N}\right)}\right)^{3}}{1+(N-1)\left(\frac{\operatorname{sen}\left(\frac{\pi-\theta}{N}\right)}{\operatorname{sen}\left(\theta+\frac{\pi-\theta}{N}\right)}\right)^{2}} \end{aligned}$$

In this way, for $N$ equidistant LI pure states with overlap phase different from zero and equal a priori preparations probabilities we apply a *complete states discrimination* obtaining all the possible information about the prepared state of the system of interest.

## 3.4 Remarks on our results in QSD

In this thesis our contributions to QSD are:

- The characterization of a new set of states (the equidistant states), useful for QSD and any research area where a set of non-orthogonal states is required, because of their simplicity.

- The joint application of Ancillary USD and Minimum error protocols in their QSD of the equidistant states.

# 4 State Tomography

State tomography is the process of constructing the state of a physical system from the measurements of observables of the system. In particular we are interested in quantum state tomography i.e. the state tomography of quantum systems. This is a fundamental tool of not just quantum information, but of quantum mechanics in general, since on quantum tomography relies all the knowledge and predictions we can get from a quantum system. In this chapter we briefly review the methods used to make the state assignment from the measurements and then we explain a novel method that is other of our contributions to this field.

## 4.1 Linear inversion

The basic idea of this method consist in the use of Born´s rule to state a linear system of equations relating the probabilities outcomes from the measurement of observables with the coefficients of the state density matrix associated with the quantum system and then solve the system by inversion.

Born´s rule:

$$p_i = Tr\left(E_i \rho\right) \tag{4.1}$$

gives, as remarked on Chapter 1, the probability $p_i$ of a "yes" outcome when an effect $E_i$ is measured on a quantum system described by the state $\rho$. Thus from a POVM $\{E_1, E_2, \ldots, E_n\}$ on such a system follows:

$$A\vec{\rho} = \begin{pmatrix} Tr\left(E_1\rho\right) \\ Tr\left(E_2\rho\right) \\ \vdots \\ Tr\left(E_n\rho\right) \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} = \vec{p} \tag{4.2}$$

Here, $\vec{\rho}$ is a vector representation of the coefficients of the density matrix associated with the state $\rho$. Fixing $\vec{\rho}$ we have $A$ as the matrix elements of the system 4.2. If we apply the same vector representation as in $\vec{\rho}$ to the POVM elements, we have:

$$A = \begin{pmatrix} \vec{E}_1^\dagger \\ \vec{E}_2^\dagger \\ \vdots \\ \vec{E}_n^\dagger \end{pmatrix} \tag{4.3}$$

Since $A$ is not square in general, to solve the system we need to multiply by its transpose $A^t$ and hence we have:

$$\vec{\rho} = \left(A^t A\right)^{-1} A^t \vec{p} \tag{4.4}$$

as solution to the system. This is possible only when the POVM is chosen such that $\left(A^t A\right)$ is invertible, in this case we say that the POVM is *informational complete* (IC-POVM). Hence when a IC-POVM is measured on the system the tomographic construction of the state is given by 4.4.

## 4.2 State estimation

This method consist in searching for a density matrix $\hat{\rho}$ within a subset $\mathcal{S}' \subseteq \mathcal{S}(\mathcal{H})$ such that optimizes a certain *likelihood function* $\mathcal{L}(\hat{\rho})$ i.e. a function that measures how close a state is to characterize a given (from the experiments) probability distribution. The parametrization of $\mathcal{S}'$ must assure that the *estimator* $\hat{\rho}$ is always within an open ball that is also a subset of $S(\mathcal{H})$. The most popular likelihood function is the probability that would be assigned to the observed results if the system is characterized by the estimator of the state. If we have a POVM $\{P_1, P_2, \dots, P_n\}$ and each effect $P_j$ has been observed with frequency $f_j$ the likelihood function we should maximize is:

$$\mathcal{L}(\hat{\rho}) = \prod_{j=1}^{n} \left[Tr\left(P_j \hat{\rho}\right)\right]^{f_j} \tag{4.5}$$

Also is very common to minimize its negative logarithm $-\log \mathcal{L}(\hat{\rho})$. Since $\mathcal{L}(\hat{\rho})$ usually have local maxims, is also usual to demand the simultaneous optimization of other function, like the Von Neumann entropy:

$$S(\hat{\rho}) = -Tr\left[\hat{\rho} \log\left(\hat{\rho}\right)\right] \tag{4.6}$$

to get one global maxim as is done in [43, 42]. One also can improve the method by choosing the POVM such that minimizes the Hilbert-Schmidt distance between the state given by linear inversion and the estimator, this will be analyzed in chapter 5.

As an example of this method we will show the tomography of a single qubit [41]. The parametrization of $\mathcal{S}'$ will be:

$$\hat{\rho} = \frac{T^\dagger T}{Tr\left(T^\dagger T\right)} \tag{4.7}$$

with $T$ :

$$T\left(t\right)=\begin{pmatrix} t_1 & 0 \\ t_3+it_4 & t_2 \end{pmatrix} \tag{4.8}$$

and $t=\left(t_1,\ldots,t_4\right)$. By choosing the parametrization (4.7) we ensure that $\hat{\rho}$ will satisfy all the requirements of a density matrix, while (4.8) is tridiagonal because this is useful to be able to invert relation (4.7) and it has 4 parameters (being only 3 for a qubit) because is better in order to fit the intensity of the data.

The likelihood function, will in general depend on the specific measurement apparatus used and the physical implementation of the qubit as these will determine the statistical distribution of counts, and therefore their relative weightings, so to keep with this example we will assume both Gaussian counting statistics and that each of our measurements is taken for the same amount of time, then we can provide a suitable likelihood function.

Let $n_q$ be the $q$-th measurement, out of a total of $K$ measurements. The expected values for these measurements on an uncharacterized system are given by $\bar{n}_q=\mathcal{N}Tr\left[P_q\hat{\rho}\right]$. Here $\mathcal{N}$ is a normalization parameter corresponding to the total size per measurement of the ensemble and so we have for the probability of the outcome $p_q=n_q/\mathcal{N}$. It is not always possible to know the size of a measured ensemble, and so the counts rather than the probabilities are used in the likelihood function. Given these definitions, the probability of obtaining the observed experimental counts $n_q$ from the density matrix $\hat{\rho}$ is:

$$P\left(n_1,\ldots,n_K\right)=\frac{1}{\mathcal{N}_{\mathcal{G}}}\prod_q \exp\left[-\frac{\left(\bar{n}_q-n_q\right)^2}{2\sigma_q^2}\right] \tag{4.9}$$

where $\sigma_q$ is the standard deviation of the $q$-th measurement and $\mathcal{N}_{\mathcal{G}}$ is the normalization constant of the distribution. For our estimator $\hat{\rho}_e\left(t\right)$ density matrix the number of counts expected for the $q$-th measurement is:

$$\bar{n}_q\left(t\right)=\mathcal{N}Tr\left[P_q\hat{\rho}_e\left(t\right)\right] \tag{4.10}$$

Thus the likelihood that the matrix $\hat{\rho}_e\left(t\right)$ could produce the measured data $\left\{n_1,\ldots,n_K\right\}$ is:

$$P\left(n_1,\ldots,n_K\right)=\frac{1}{\mathcal{N}_{\mathcal{G}}}\prod_q \exp\left[-\frac{\left(\mathcal{N}Tr\left[P_q\hat{\rho}_e\left(t\right)\right]-n_q\right)^2}{2\mathcal{N}Tr\left[P_q\hat{\rho}_e\left(t\right)\right]}\right] \tag{4.11}$$

Here we approximated $\sigma_q\approx\sqrt{\bar{n}_q}$ and assumed that $\mathcal{N}$ is the same for each measurement (for simplicity, since in practice this may not necessarily be the case). Then, rather than find the maximum value of $P\left(t\right)$, it is numerically simpler to find the maximum of its logarithm which is equivalent since this is a monotonically increasing function. Also, because $\mathcal{N}$ is unknown we absorb it into the $\hat{T}$ matrix, by setting:

$$t_i' = \mathcal{N} t_i \tag{4.12}$$

Thus the optimization problem reduces to finding the minimum of the following function:

$$\mathcal{L}\left(t'\right) = \sum_q \frac{\left(Tr\left[P_q \hat{\rho}_e\left(t'\right)\right] - n_q\right)^2}{2Tr\left[P_q \hat{\rho}_e\left(t'\right)\right]} \tag{4.13}$$

The final part of the maximum likelihood technique is an optimization routine, of which there are many available in the literature [41, 42].

## 4.3 Quantum state tomography by quantum state discrimination

A different tomographic scheme using QSD is proposed by us in our paper *Quantum tomography via unambiguous state discrimination* [44]. We show that the inverse process of Ancillary USD, i.e. the mapping from orthogonal states onto linearly independent non-orthogonal states, which is implemented by concatenating a unitary transformation acting jointly onto the system of interest with an ancillary and then a PVM on the ancillary system, can be employed to implement quantum tomography. Here we develop in detail the single qubit case, because this way we simply focus on the main features of this new method. We also generalize the method to single qudits and in the next section we will outlined briefly since there is nothing conceptually different from the qubit case.

Let us start by considering a density matrix $\hat{\rho}^{(s)}$ of a two-dimensional *uncharacterized* [1] quantum system $s$ . This is given by:

$$\hat{\rho}^{(s)} = \sum_{i,j=0,1} \rho_{ij} \left|i\right\rangle_s \left\langle j\right| \tag{4.14}$$

where the states $\left|0\right\rangle_s$ and $\left|1\right\rangle_s$ form an orthonormal base of the Hilbert space of system $s$. To determine these coefficients we resort to the unitary transformation $U$ whose action onto states $\left|i\right\rangle_s$ is defined by

$$U\left(\left|i\right\rangle_s \left|\Lambda\right\rangle_a\right) = \sqrt{p_i}\left|\beta_i\right\rangle_s \left|0\right\rangle_a + \sqrt{1-p_i}\left|\gamma_i\right\rangle_s \left|1\right\rangle_a \tag{4.15}$$

[1] Since we still need to perform the tomography to make the corresponding assignments to the coefficients of $\hat{\rho}^{(s)}$. This is usually called a *unknown* density matrix. We will avoid this term because if quantum states correspond to a state of knowledge, talking of an unknown state is an oxymoron. This problem and the justification of actual tomographic methods find a reasonably solution through Quantum de Finetti´s Theorem, a more detailed exposition of this point is given in Chapter 7.

with $i = 0, 1$. In order to implement this transformation an ancillary system $a$ in the initial arbitrary state $\left|\Lambda\right\rangle_a$ is necessary. Pure states of this system are spanned by the orthogonal states $\left|0\right\rangle_a$ and $\left|1\right\rangle_a$. A projection of system $a$ onto the state $\left|0\right\rangle_a$ transform states $\{\left|0\right\rangle_s, \left|1\right\rangle_s\}$ into states $\{\left|\beta_o\right\rangle_s, \left|\beta_1\right\rangle_s\}$ with probabilities $p_1$ and $p_2$ respectively. Similarly, a projection of system $a$ onto the state $\left|1\right\rangle_a$ transform states $\{\left|0\right\rangle_s, \left|1\right\rangle_s\}$ into states $\{\left|\gamma_o\right\rangle_s, \left|\gamma_1\right\rangle_s\}$ with probabilities $1 - p_1$ and $1 - p_2$, correspondingly. Thereby the transformation $U$ together with a projective measurement onto the ancillary allow a map of the initially orthogonal states onto states with a non-vanishing inner product.

In general, states $\left|\beta_i\right\rangle_s$ and $\left|\gamma_i\right\rangle_s$do not need to be mutually orthogonal as far as the unitary $U$ preserves the orthogonality of the states $\left|0\right\rangle_s$ and $\left|1\right\rangle_s$, that is

$$\sqrt{p_o p_1}\left\langle\beta_o|\beta_1\right\rangle + \sqrt{\left(1-p_o\right)\left(1-p_1\right)}\left\langle\gamma_o|\gamma_1\right\rangle = 0 \tag{4.16}$$

For the sake of simplicity we will assume $p_o = p_1 = p$. This, together with the polar decompositions $\left\langle\beta_o|\beta_1\right\rangle = |\beta|e^{\theta_\beta}$ and $\left\langle\gamma_o|\gamma_1\right\rangle = |\gamma|e^{\theta_\gamma}$ leads to the value of $p$ as a function of the absolute value of the inner products

$$p = \frac{|\gamma|}{|\beta| + |\gamma|} \tag{4.17}$$

with the constraint $\theta_\beta - \theta_\gamma = \pm\pi$. Let us note that the probability $p$ does not depend on $\theta_\beta$.

We now apply the transformation $U$ onto the uncharacterized system (described by density matrix $\hat{\rho}^{(s)}$) and the ancillary system $a$, which is in the initial arbitrary state $\left|\Lambda\right\rangle_a\left\langle\Lambda\right|$, to generate the new density matrix $\hat{\sigma}^{(sa)}$ given by

$$\hat{\sigma}^{(sa)} = U\left(\hat{\rho}^{(s)}\left|\Lambda\right\rangle_a\left\langle\Lambda\right|\right)U^{\dagger} \tag{4.18}$$

This new density matrix describes the state of the four-dimensional bipartite system $\mathcal{H}_s \otimes \mathcal{H}_a$. The state $\hat{\sigma}^{(sa)}$ can be cast in the form:

$$\hat{\sigma}^{(sa)} = \sum_{i,j}\hat{\sigma}_{ij}^{(s)}\left|i\right\rangle_a\left\langle j\right| \tag{4.19}$$

The elements $\hat{\sigma}_{ij}^{(s)}$ are operators acting onto the Hilbert space of system $s$ given the four expressions:

$$\hat{\sigma}_{\mu\nu}^{(s)} = \sqrt{p\left(\mu\right)p\left(\nu\right)}\sum_{i,j}\rho_{ij}\left|\mu_i\right\rangle_s\left\langle\nu_j\right| \tag{4.20}$$

where

$$\begin{gathered} p\left(\zeta\right) = \begin{cases} p & \zeta = 0 \\ 1-p & \zeta = 1 \end{cases}, \zeta = \{\mu, \nu\} \\ \zeta_i = \begin{cases} \beta_i & \zeta = 0 \\ \gamma_i & \zeta = 1 \end{cases}, \zeta_i = \{\mu_i, \nu_i\} \end{gathered} \tag{4.21}$$

After applying the unitary transformation $U$ the probability $P_o$ of projecting the ancillary system $a$ onto state $|0\rangle_a$ becomes:

$$P_o = p\left[\rho_{00} + \rho_{11} + 2|\beta||\rho_{01}|\cos\left(\theta_{01} - \theta_\beta\right)\right] \tag{4.22}$$

where we have used the polar decomposition $\rho_{01} = |\rho_{01}|e^{\theta_{01}}$. Considering the normalization condition of the density matrix $\rho$ we obtain

$$P_o = p\left[1 + 2|\beta||\rho_{01}|\cos\left(\theta_{01} - \theta_\beta\right)\right] \tag{4.23}$$

Analogously, the probability $P_1$ of projecting the ancillary system $a$ onto state $|1\rangle_a$ is:

$$P_1 = (1-p)\left[1 - 2|\gamma||\rho_{01}|\cos\left(\theta_{01} - \theta_\beta\right)\right] \tag{4.24}$$

Each one of the latter two equations allows us to determine directly the real part of the coefficient $\rho_{01}$ with the choice $\theta_\beta = 0$, that is:

$$\begin{aligned} P_o =& p\left[1 + 2|\beta||\rho_{01}|\cos\left(\theta_{01}\right)\right] \\ P_1 =& (1-p)\left[1 - 2|\gamma||\rho_{01}|\cos\left(\theta_{01}\right)\right] \end{aligned} \tag{4.25}$$

The imaginary part of $\rho_{01}$ is obtained with a similar procedure. In this case we need to consider a second unitary transformation $\tilde{U}$ analogous to $U$ but with the choice $\theta_\beta = \pi/2$, but otherwise the same as in $U$. Thereby, the value of $p$ stays unchanged. After this second transformation onto the uncharacterized system (density matrix $\hat{\rho}^{(s)}$ ) the probabilities for projecting onto states $|0\rangle_a$ and $|1\rangle_a$ the ancillary system are:

$$\begin{aligned} \tilde{P}_o =& p\left[1 + 2|\beta||\rho_{01}|\mathrm{sen}\left(\theta_{01}\right)\right] \\ \tilde{P}_1 =& (1-p)\left[1 - 2|\gamma||\rho_{01}|\mathrm{sen}\left(\theta_{01}\right)\right] \end{aligned} \tag{4.26}$$

respectively. In a more compact form:

$$\rho_{01} = \frac{1}{2p|\beta|}\left[\left(P_o - p\right) + i\left(\tilde{P}_o - p\right)\right] \tag{4.27}$$

Since the knowledge of the non-diagonal coefficient $\rho_{01}$ univocally defines a two-dimensional pure quantum state, this result is enough to characterize pure quantum system.

The reconstruction of an uncharacterized mixed quantum system requires additionally the determination of the diagonal coefficients $\rho_{00}$ and $\rho_{11}$. This can be carried out by applying onto the uncharacterized system $\hat{\rho}^{(s)}$ the transformation $\hat{\sigma}_y$ defined by

$$\begin{aligned} \hat{\sigma}_y\left|0\right\rangle_s =& \frac{1}{\sqrt{2}}\left(\left|0\right\rangle_s + i\left|1\right\rangle_s\right) \\ \hat{\sigma}_y\left|1\right\rangle_s =& \frac{1}{\sqrt{2}}\left(\left|0\right\rangle_s - i\left|1\right\rangle_s\right) \end{aligned} \tag{4.28}$$

The new density matrix $\hat{\varrho}^{(s)} = \hat{\sigma}_y \hat{\rho}^{(s)} \hat{\sigma}_y^{\dagger}$ becomes:

$$\hat{\varrho}^{(s)} = \frac{1}{2} \begin{pmatrix} \rho_{00} + \rho_{11} + 2|\rho_{01}| \cos(\theta_{01}) & i\left[\rho_{00} - \rho_{11}\right] - 2|\rho_{01}| \mathrm{sen}(\theta_{01}) \\ -i\left[\rho_{00} - \rho_{11}\right] - 2|\rho_{01}| \mathrm{sen}(\theta_{01}) & \rho_{00} + \rho_{11} - 2|\rho_{01}| \cos(\theta_{01}) \end{pmatrix} \quad (4.29)$$

Onto this new density matrix we apply the procedure described in the previous paragraphs based on the transformation $\tilde{U}$. This way we determine the imaginary part of the coefficient $\varrho_{01}$ which is proportional to $\rho_{00} - \rho_{11}$. This, together with the normalization condition allow us to obtain the value of the coefficients $\rho_{00}$ and $\rho_{11}$.

An important feature of the scheme here proposed is the fact that all the information needed for the reconstruction is obtained through the ancillary system. Consequently, the system $s$ does not undergo a measurement process. it seems thus feasible to recover the initial state $\hat{\rho}^{(s)}$. After the application of transformation $U$ and a projection of the ancillary system onto state $|0\rangle_a$ normalized post-measurement density matrix $\hat{\sigma}_{00}^{(s)}$ of system $s$ is given by

$$\hat{\sigma}_{00}^{(s)} = \sum_{i,j} \frac{\rho_{ij}}{N_o} |\beta_i\rangle_s \langle\beta_j| \quad (4.30)$$

with

$$N_o = 1 + 2|\beta||\rho_{01}| \cos(\theta_{01}) \quad (4.31)$$

Now, to transform $\hat{\sigma}_{00}^{(s)}$ into $\hat{\rho}^{(s)}$ we must undo the quantum state separation process applied on system $s$, that is the transformation of states $|\beta_i\rangle_s$ into states $|i\rangle_s$. For this purpose we resort to optimal state discrimination and apply it to the states $|\beta_i\rangle_s$. This discrimination strategy is based in the following transformation

$$D_\beta\left(|\beta_i\rangle_s |\Lambda_\beta\rangle_a\right) = \sqrt{m_\beta}\, |i\rangle_s |0\rangle_a + \sqrt{1 - m_\beta}\, |\phi_i\rangle_s |1\rangle_a \quad (4.32)$$

where $|\Lambda_\beta\rangle_a$ is an arbitrary state of the ancillary system and $m_\beta$ is the optimal success discrimination given by the Ivanovic-Dieks-Peres limit [32, 33, 34]:

$$1 - |\beta| \quad (4.33)$$

This limit is obtained assuming that the states $|\beta_i\rangle_s$ are generated with the same a priori probabilities and states $|\phi_i\rangle_s$ are linearly dependent. A projection of the ancillary system onto state $|0\rangle_a$ maps states $|\beta_i\rangle_s$ onto $|i\rangle_s$ with probability $m_\beta$. Otherwise, states $|\beta_i\rangle_s$ are mapped onto $|\phi_i\rangle_s$.

We see that transformation $D_\beta$ works in the desired way, since its application to the state $\hat{\sigma}_{00}^{(s)}$ gives:

$$D_\beta\left(\hat{\sigma}_{00}^{(s)} |\Lambda_\beta\rangle_a \langle\Lambda_\beta|\right) D_\beta^{\dagger} = m_\beta \sum_{i,j} \frac{\rho_{ij}}{N_o} |i\rangle_s \langle j| \, |0\rangle_a \langle 0| + \ldots$$

as guessed, the projection of ancillary system onto state $|0\rangle_a$ maps system $s$ onto state $\hat{\rho}^{(s)}$ with probability

$$P_\beta = \frac{m_\beta}{N_o} = \frac{1-|\beta|}{1+2|\beta||\rho_{01}|\cos(\theta_{01})} \tag{4.34}$$

Which can be large as required by minimizing the value of $|\beta|$.

Analogously, it is also possible to recover the state, it is also possible to recover the state $\hat{\rho}^{(s)}$ from the state $\hat{\sigma}_{11}^{(s)}$, which is obtained after applying transformation $U$ and projecting the ancillary system onto state $|1\rangle_a$. This state is given by

$$\hat{\sigma}_{11}^{(s)} = \sum_{i,j} \frac{\rho_{ij}}{N_1} |\gamma_i\rangle_s \langle\gamma_j| \tag{4.35}$$

with

$$N_1 = 1 - 2|\gamma||\rho_{01}|\cos(\theta_{01}) \tag{4.36}$$

The recovery of $\hat{\rho}^{(s)}$ succeed in this case with the help of the transformation

$$D_\gamma\left(|\gamma_i\rangle_s |\Lambda_\gamma\rangle_a\right) = \sqrt{m_\gamma}\,|i\rangle_s |0\rangle_a + \sqrt{1-m_\gamma}\,|\phi_i\rangle_s |1\rangle_a \tag{4.37}$$

where all quantities are defined in analogy to equation (4.32) and $m_\gamma = 1 - |\gamma|$.

In this case the probability $P_\gamma$ of recovering the state $\hat{\rho}^{(s)}$ from the state $\hat{\sigma}_{11}^{(s)}$ is given by

$$P_\gamma = \frac{m_\gamma}{N_1} = \frac{1-|\gamma|}{1-2|\gamma||\rho_{01}|\cos(\theta_{01})} \tag{4.38}$$

We can now compute the total probability $P_{rec}$ of recovering state $\hat{\rho}^{(s)}$ from the post-measurement states $\hat{\sigma}_{00}^{(s)}$ and $\hat{\sigma}_{11}^{(s)}$.This is given by

$$P_{rec} = P_o P_\beta + P_1 P_\gamma = p m_\beta + (1-p)\, m_\gamma \tag{4.39}$$

or equivalently

$$P_{rec} = 1 - 2\frac{|\beta||\gamma|}{|\beta| + |\gamma|} \tag{4.40}$$

This probability turns out to be a symmetric function of the inner products $|\beta|$ and $|\gamma|$, which define the transformation $U$, and does not depend on the particular state $\hat{\rho}^{(s)}$ to be recovered. This probability approach the maximum value as $|\beta|$ or $|\gamma|$ vanish. In this limit the contribution of the real part of $\rho_{01}$ to probabilities $P_o$ and $P_1$ diminishes since this real part is multiplied by one of the inner products $|\beta|$ or $|\gamma|$ respectively. Thereby, exist a trade-off between the difficulty to estimate the real part of $\rho_{01}$ and the probability of retrieving the initial state $\hat{\rho}^{(s)}$ after this process.

The state $\hat{\rho}^{(s)}$ can also be recovered from the states of system $s$ generated by our procedure to estimate the imaginary part of $\rho_{01}$ and the diagonal coefficients $\rho_{00}$ and $\rho_{11}$. In each of this stages the probability of recovering $\hat{\rho}^{(s)}$ is given by probability $P_{rec}$.

We can now consider to concatenate each one of the three stages of the tomographic scheme with a recovery stage and use the output state of one stage as the input state for the next stage. The combined probability $P_{suc}$ of successfully recovering the state $\hat{\rho}^{(s)}$ after the three successive stages of the tomographic scheme is thus $P_{suc} = (P_{rec})^3$. A simpler expression for $P_{rec}$ is found for the case $|\beta| = |\gamma|$. In this case $p = 1/2$ and $P_{rec} = 1 - |\beta|$, which is the Ivanovic-Dieks-Peres limit. Thus, our capability for recovering the initially uncharacterized system in the state $\hat{\rho}^{(s)}$ of the preparation after the application of the tomographic scheme is limited by our ability to unambiguously discriminate among non-orthogonal states.

## 4.4 The Qudit case

The generalization to the qudit case goes simply by several uses of ancillary systems with unitary transformations as 4.15, whose action is over all the possible pairs of the canonical basis, i.e. taking $i = \{k, n\}$ with $k, n \in \{0, \ldots, d\}$. The $d(d-1)/2$ different combinations of this tomography scheme for $i = \{k, n\}$ gave us the non-diagonal coefficients.

For a mixed state we should apply the analogous transformation of $\sigma_y$ with action over $\{|k\rangle_s, |n\rangle_s\}$ at least $(d-1)$ times, this with the normalization condition allow for the calculation of the diagonal coefficients. If we want to recover the state from the post-measurement states, the unitary transformations analogous to $D_\beta$ and $D_\gamma$ should be performed.

This method requires $2d^2 - d$ projectors in the general case and $5d - 4$ projectors for a pure state. This requires $d$ order of projectors less than standard tomography and more projectors than MUB state tomography but of the same order in $d$.

## 4.5 Remarks on our method

This new tomographic scheme we have proposed has three main features:

- The coefficients defining the quantum state are obtained directly from the measured quantities, no inversion procedure is required.
- The information about the coefficients is obtained through transition probabilities arising from a PVM carried out onto the ancillary system.
- The system described by the state does not undergo a measurement process. Therefore, it is possible to recover the initial state of this system from the post-measurement states.

The third feature prevent us from changing the quantum state of the system while still measuring, something that has been though impossible.

This is achieved with the help of an optimal unambiguous state discrimination process, albeit with a certain probability which can be large. The unitary transformations required in our scheme have been successfully performed experimentally via polarization-dependent absorption in a fiber [45] or with the help of polarizing beam-splitters [46].

# 5

# Tight Informational complete POVMs

In the previous chapter we showed the construction of a quantum state from its measurement statistics. In this chapter we will investigate a class of IC-POVM which share a particularly simple tomographic formula. To this end, first we introduce the basic concepts of frame theory and then we rewrite IC-POVM definition in this context [1]. Finally we review the contributions of the author in the research of one of the more relevant of this IC-POVMs; the existence conditions of the so called SIC-POVM.

## 5.1 Super operator algebra

Frame theory [47, 48, 49] provides a useful setting for the study of IC-POVMs [75]. In this section we will introduce some of the basic concepts of super operator algebra and frames.

Following [59] we will write a linear operator $A$ in vector notation as $|A)$. The vector space of all such operators, $\mathcal{B}\left(\mathbb{C}^d\right) \cong \mathbb{C}^{d^2}$, equipped with the Hilbert-Schmidt (H-S) inner product $(A|B) \equiv Tr(A^\dagger B)$ is a Hilbert space. The usefulness of this notation becomes clear when we consider linear maps on operators, i.e. superoperators. given an orthonormal operator basis $\{M_k\}_{k=1}^{d^2} \subset \mathcal{B}(\mathbb{C}^d)$, $(M_i|M_j) = \delta_{ij}$, a superoperator $S \in \mathcal{B}\left(\mathcal{B}(\mathbb{C}^d)\right) \cong \mathbb{C}^{d^4}$ may be written in two different ways:

$$S = \sum_{j,k} s_{jk} M_j \odot M_k^\dagger = \sum_{j,k} s_{jk} |M_j)(M_k| \;\; (s_{jk} \in \mathbb{C}) \tag{5.1}$$

The first representation illustrates the *ordinary* action of the superoperator,

$$S(A) \equiv \sum_{j,k} s_{jk} M_j A M_k^\dagger \tag{5.2}$$

[1] The introduction to frame theory has been taken from [65]

which amounts to inserting $A$ into the location of the $\odot$ symbol. The second reflects the *left-right* action,

$$S|A) \equiv \sum_{j,k} s_{jk}|M_j)(M_k|A) = \sum_{j,k} s_{jk}|M_j)Tr(M_k^{\dagger}A) \tag{5.3}$$

where the superoperator acts on operators just like an operator on vectors. The identity superoperators relative to the ordinary and left-right actions are, respectively $\mathcal{I} \equiv I \odot I$ and $\mathbf{I} \equiv \sum_k |M_k)(M_k|$. Further results on superoperators in the current notation can be found in [65].

Frames generalize the notion of bases. We call a countable family of operators $\{A_x\}_{x\in\mathcal{K}} \subset \mathcal{B}\left(\mathbb{C}^d\right)$ an *operator frame* if there exist constants $0 < a \leq b < \infty$ such that

$$a\,(C|C) \leq \sum_{x\in\mathcal{K}} |(A_x|C)|^2 \leq b\,(C|C) \quad \forall C \in \mathcal{B}\left(\mathbb{C}^d\right) \tag{5.4}$$

When $a = b$ the frame is called *tight* [50]. Tight frames which are most like orthonormal bases [62]. An operator frame with cardinality $|\mathcal{K}| = d^2$, i.e. an operator basis, is tight iff is an orthonormal basis. For every frame $\{A_x\}_{x\in\mathcal{K}}$ there is a *dual frame* $\{B_x\}_{x\in\mathcal{K}}$, such that

$$\sum_{x\in\mathcal{K}} |B_x)(A_x| = \mathbf{I} \tag{5.5}$$

When $|\mathcal{K}| > d^2$ there is different choices for the dual frame [63], but there is one known as the *canonical dual frame* $\left\{\tilde{A}_x\right\}_{x\in\mathcal{K}}$:

$$|\tilde{A}_x) \equiv \mathcal{A}^{-1}|A_x) \tag{5.6}$$

,

where $\mathcal{A}$ is the *frame superoperator* :

$$\mathcal{A} \equiv \sum_{x\in\mathcal{K}} |A_x)(A_x| \tag{5.7}$$

Note that the inverse of $\mathcal{A}$ is taken with respect to left-right action and exist whenever $\{A_x\}_{x\in\mathcal{K}}$ is an operator frame. A tight operator frame is one with $\mathcal{A} = a\mathbf{I}$ i.e. a resolution of unity, and thus $|\tilde{A}_x) \equiv \frac{1}{a}|A_x)$. In general, however, inverting the frame superoperator will be a difficult analytical task. Also, because tight frames are by themselves a resolution of unity and the following inequality [62]:

Theorem 5.1: *Let $\{A_x\}_{x\in\mathcal{K}} \subseteq \mathcal{B}\left(\mathbb{C}^d\right)$ be an operator frame. Then*

$$\sum_{x,y\in\mathcal{K}} |(A_x|A_y)|^2 \geq \frac{[Tr\,(\mathcal{A})]^2}{d^2} \tag{5.8}$$

*with equality iff $\{A_x\}_{x\in\mathcal{K}}$ is a tight operator frame.*

which is called the frame bound, that they are as close as possible to orthonormal bases. Theorem 5.1 shows that tight frames are those which minimize the average correlation amongst the frame elements. In fact, explicit examples of tight *unitary* operator frames are known for all $d$ and $|\mathcal{K}| \geq d^2$ [64].

## 5.2 Informational complete POVMs

In this section we are going to relate the concept of IC-POVMs with those of frame theory.

As explained in Chapter 4 an informational complete quantum measurement $\{E_k\}$ is one that uniquely determines each quantum state $\rho \in \mathcal{B}^*(\mathcal{H})$ from its measurement statistics $p_k = Tr[E_k\rho]$. Consequently, given multiple copies of an uncharacterized system, a sequence of measurements will give an estimate of the statistics, and hence, allow us to make a state assignment for the corresponding preparation of the system. The measurement $\{E_k\}$ is then called an informational complete POVM (IC-POVM).

Definition 5.2: *A* POVM $\{E_k\}$*is called* informational complete *if for each pair of distinct quantum states* $\rho \neq \sigma \in \mathcal{B}^*(\mathcal{H})$
there exist an effect $E_p$ such that $Tr(E_p\rho) \neq Tr(E_p\sigma)$.

For an arbitrary POVM $\{F_x\}_{x\in\mathcal{K}}$, define the superoperator:

$$\mathcal{F} = \sum_{x\in\mathcal{K}} [Tr(F_x)]^{-1} |F_x)(F_x| \tag{5.9}$$

This superoperator is positive an bounded under left-right action:

$$0 \leq (A|\mathcal{F}|A) \leq d(A|A) \;\; \forall A \in \mathcal{B}\left(\mathbb{C}^d\right) \tag{5.10}$$

Also we have the following result:

Proposition 5.3: *Let* $\{F_x\}_{x\in\mathcal{K}}$ *be a POVM. Then* $\{F_x\}_{x\in\mathcal{K}}$ *is informational complete iff there exist a constant* $a > 0$ *such that* $(A|\mathcal{F}|A) \geq a(A|A)$ *for all* $A \in \mathcal{B}\left(\mathbb{C}^d\right)$.

so, this means that $\{F_x\}_{x\in\mathcal{K}}$ is a IC-POVM when $\mathcal{F}$ has full rank with respect to the left-right action. If we take as an operator frame the *normalization* $|P_x) = [Tr(F_x)]^{-1}|F_x)$ of some IC-POVM $\{F_x\}_{x\in\mathcal{K}}$, then the canonical dual frame defines a *reconstruction operator*:

$$|R_x) \equiv \mathcal{F}^{-1}|P_x) \tag{5.11}$$

where the inverse of $\mathcal{F}$, which we now call the *POVM superoperator*, is taken with respect to the left-right action.

The identity

$$\sum_{x\in\mathcal{K}} [Tr(F_x)] |R_x)(P_x| = \sum_{x\in\mathcal{K}} [Tr(F_x)] \mathcal{F}^{-1}|P_x)(P_x| = \mathcal{F}^{-1}\mathcal{F} = \mathbf{I} \tag{5.12}$$

then allows state reconstruction in terms of the measurement statistics:

$$\rho = \sum_{x\in\mathcal{K}} \left[Tr\left(F_x\right)\right] \left|R_x\right) \left(P_x|\rho\right) = \sum_{x\in\mathcal{K}} \left(F_x|\rho\right) \left|R_x\right) = \sum_{x\in\mathcal{K}} p_x R_x \tag{5.13}$$

Notice that we need $|\mathcal{K}| \geq d^2$ for $\{F_x\}_{x\in\mathcal{K}}$ to be informational complete, since otherwise $\mathcal{F}$ could not have full rank. An IC-POVM with $|\mathcal{K}| = d^2$ is called *minimal.* In this case the reconstruction operator frame is unique.

## 5.3 Tight IC-POVMs

For the POVM superoperator we have the decomposition:

$$\mathcal{F} = \frac{\mathcal{I}}{d} + \sum_{x\in\mathcal{K}} \left[Tr\left(F_x\right)\right] \left|P_x - I/d\right) \left(P_x - I/d\right| \tag{5.14}$$

The superoperator $\mathcal{I}/d = |I)\,(I|\,/d$ is in fact an eigenprojector. It left-right projects onto the subspace spanned by the identity, whose orthogonal complement $\mathbb{I}^{\perp}\left(\mathbb{C}^d\right) \cong \mathbb{R}^{d^2-1}$, the subspace of traceless operators, is $\mathcal{F}$-invariant. Define $\Xi \equiv \mathbf{I} - \mathcal{I}/d$ , which left-right projects onto this latter subspace. The action of $\Xi$ on a quantum state then realizes the above embedding into $\mathbb{I}^{\perp}\left(\mathbb{C}^d\right)$:

$$\Xi\,|\rho) = |\rho - I/d) \tag{5.15}$$

Let $\mathbf{I}_{\Xi}$ denote the identity superoperator for $\mathbb{I}^{\perp}\left(\mathbb{C}^d\right)$ under left-right action, we are now ready to define a *tight IC-POVM*:

Definition 5.4: *Let $\{F_x\}_{x\in\mathcal{K}}$ be a POVM, with POVM superoperator $\mathcal{F}$. Then $\{F_x\}_{x\in\mathcal{K}}$ is called* tight IC-POVM *if the embedding of its normalization $\{P_x\}_{x\in\mathcal{K}}$ POVM, $\Xi\,|P_x) = |P_x - I/d)$ forms a tight operator frame in $\mathbb{I}^{\perp}\left(\mathbb{C}^d\right)$, i.e.:*

$$\sum_{x\in\mathcal{K}} \left[Tr\left(F_x\right)\right] \left|P_x - I/d\right) \left(P_x - I/d\right| = a\mathbf{I}_{\Xi} \tag{5.16}$$

*or equivalently $\Xi\mathcal{F}\Xi = a\Xi$ for some constant $a > 0$.*

The constant $a$ can be found by taking the superoperator trace of 5.16:

$$\begin{aligned} a =& \frac{1}{d^2-1} \sum_{x\in\mathcal{K}} \left[Tr\left(F_x\right)\right] \left(P_x - I/d|P_x - I/d\right) \\ =& \frac{1}{d^2-1} \left( \sum_{x\in\mathcal{K}} \left[Tr\left(F_x\right)\right] \left(P_x|P_x\right) - 1 \right) \end{aligned} \tag{5.17}$$

The POVM superoperator of a tight IC-POVM satisfies the identity:

$$\mathcal{F} = \frac{\mathcal{I}}{d} + a\Xi = a\mathbf{I} + \frac{1-a}{d}\mathcal{I} \tag{5.18}$$

Since $a > 0$ by definition, this superoperator has full rank and its inverse is:

$$\mathcal{F}^{-1} = \frac{1}{a}\mathbf{I} - \frac{1-a}{ad}\mathcal{I} \tag{5.19}$$

and thus the reconstruction operator takes the form:

$$R_x = \frac{1}{a}P_x - \frac{1-a}{ad}I \tag{5.20}$$

A tight IC-POVM then has a simple state-reconstruction formula[2]:

$$\rho = \frac{1}{a}\sum_{x\in\mathcal{K}} p_x P_x - \frac{1-a}{ad}I \tag{5.21}$$

The above formula simplify further in the important special case of a tight rank-one IC-POVM. The frame constant then takes its maximum possible value:

$$a = \frac{1}{1+d} \tag{5.22}$$

which gives for a tight rank-one IC-POVM the superoperator:

$$\mathcal{F} = \frac{\mathbf{I}+\mathcal{I}}{d+1} \tag{5.23}$$

and the state-reconstruction formula for a tight rank-one IC-POVM also takes an elegant form:

$$\rho = (d+1)\sum_{x\in\mathcal{K}} p_x \Pi_x - I \tag{5.24}$$

where we have set the POVMs to $P = \Pi$, to keep the notation most used in the literature.

A tight rank-one IC-POVM which is minimal i.e. $|\mathcal{K}| = d^2$, is defined by the property [51]:

$$(\Pi_x|\Pi_y) = |\langle x|y\rangle|^2 = \frac{1+d\delta_{xy}}{1+d} \tag{5.25}$$

[2] Remember that from the Born rule $p_x = (F_x|\rho)$, because $P_x$ is only the normalization of the POVM element $F_x$.

Since their overlap is the same between each one of the propositions is that they are known as *symmetric* IC-POVM (SIC-POVM). Although analytical constructions are known only for $d \leq 10$, $d = 12, 13, 19$ [51, 54, 55, 56, 57, 58] and numerically until $d \leq 67$ [70], SIC-POVMs are conjectured to exist in all dimensions [51, 54].

Another relevant example of a tight rank-one IC-POVM is a complete set of mutually unbiased bases (MUBs) [52, 53]. That is, a set of $d+1$ orthonormal bases for $\mathbb{C}^d$ with the same overlap of $1/d$ between elements of different bases:

$$\left(\Pi_j^l|\Pi_k^m\right) = \left|\left\langle e_j^l|e_k^m\right\rangle\right|^2 = \left\{\begin{array}{ll} \delta_{jk}, & l = m \\ 1/d, & l \neq m \end{array}\right. \tag{5.26}$$

Such IC-POVMs allow state determination via orthogonal measurements. Although constructions are known for prime-power dimensions [52, 53], a complete set of MUBs is unlikely to exist in all dimensions.

## 5.4 SIC-POVMs existence conditions

To get a useful parametrization of the state space first we represent $\mathcal{B}^*\left(\mathbb{C}^d\right)$ as a decomposition between what is expanded by the identity and its orthocomplement:

$$\mathcal{B}^*\left(\mathbb{C}^d\right) = \mathbb{I} \oplus \mathbb{I}^{\perp}\left(\mathbb{C}^d\right) \tag{5.27}$$

Since $Tr\left(\rho\right) = 1$ the component of states in $\mathbb{I}\left(\mathbb{C}^d\right)$ is $I/d$, and for $\mathbb{I}^{\perp}\left(\mathbb{C}^d\right)$ we choose a linear combination of the $d^2 - 1$ generators $T_b$ of the *su(d)* algebra. In the following we consider a traceless Hermitian representation of the generators $T_b$, such as the generalized Gell-Mann basis [69], given by:

$$T_b = \sqrt{\frac{2}{b\left(b+1\right)}}\left(\sum_{j=1}^{b}\left|j\right\rangle\left\langle j\right| - b\left|b+1\right\rangle\left\langle b+1\right|\right) \tag{5.28}$$

with $b = 1, \ldots, d-1$,

$$T_b = \left|k\right\rangle\left\langle m\right| + \left|m\right\rangle\left\langle k\right| \tag{5.29}$$

with $b = d, \ldots, \left(d^2 + d - 2\right)/2$, $k = 2, \ldots, d$ and $m = 1, \ldots, k-1$,

$$T_b = \left|-ik\right\rangle\left\langle m\right| + i\left|m\right\rangle\left\langle k\right| \tag{5.30}$$

with $b = \left(d^2 + d\right)/2, \ldots, d^2 - 1$, where $k = 2, \ldots, d$ and $m = 1, \ldots, k-1$. Note that the indexes $(k, m)$ are connected with the index $b$ by the equation $b = k^2/2 + 3k/2 + m + 1$.

Thereby, an arbitrary state $\rho$ can be cast in the form

$$\rho = \frac{I}{d} + \sqrt{\frac{d-1}{2d}} \sum_{b=1}^{d^2-1} r_b T_b \tag{5.31}$$

Furthermore, considering Hermitian generators $T_b$ the $\left(d^2-1\right)$-dimensional vector $\vec{r}$ has real coefficients, that is $\vec{r} \in \mathbb{R}^{d^2-1}$. Consider again the Hilbert-Schmidt product $\langle \rho, \sigma \rangle = Tr\left(\rho^* \sigma\right)$, but now on $\mathcal{B}^*\left(\mathbb{C}^d\right)$. Using the representation 5.31 the H-S product between states takes the form:

$$Tr\left(\rho^* \sigma\right) = Tr\left(\rho \sigma\right) = \frac{1}{d} + \left(\frac{d-1}{d}\right) \vec{r}_\rho \cdot \vec{r}_\sigma \tag{5.32}$$

Thus, the H-S product of states is given by the scalar product between the real vectors representing the states. In particular the case $\rho = \sigma$ allow us to study the purity of the operators:

$$Tr\left(\rho^2\right) = \frac{1}{d} + \left(1 - \frac{1}{d}\right) \left|\vec{r}_\rho\right|^2 \tag{5.33}$$

clearly $Tr\left(\rho^2\right) \leq 1 \Rightarrow \left|\vec{r}_\rho\right|^2 \leq 1$. Consequently pure states, which satisfy $Tr\left(\rho^2\right) = 1$, are on the surface of a unitary hypersphere and mixed states (i.e. $Tr\left(\rho^2\right) < 1$) are within the hypersphere. This hypersphere is known as the *Bloch sphere* and this is why the $\vec{r}_\rho$ representation of a $d$-dimensional state $\rho$ is known as *Bloch representation*. This maps states into a ball in $\mathbb{I}^{\perp}\left(\mathbb{C}^d\right)$, such a map is surjective when $d = 2$, but is otherwise only injective.

By the same procedure we can achieve a Bloch representation of operators in $\mathcal{B}\left(\mathbb{C}^d\right)$, such that rank-one operators are in the surface of the Bloch sphere and higher ranks are inside. In this picture equations 5.25 and 5.26 give us the geometrical structure of a SIC-POVM and MUBs respectively. The elements of a SIC-POVM correspond to unitary vectors pointing to the vertices of a regular simplex, while each MUB correspond to unitary vectors pointing to the vertices of a regular simplex in the $(d-1)$-dimensional subspace which they span and a complete set of MUBs correspond to a maximal set of $d+1$ mutually orthogonal MUB subspaces.

In our paper *Constructing symmetric informationally complete positive-operator-valued measures in Bloch space* [66] the problem of finding the necessary and sufficient conditions for the existence of SIC-POVM in arbitrary finite dimensions was solved. Since in general the map induced by the Bloch representation is only injective (i.e. some points into the Bloch sphere do not correspond to positive semidefinite operators in 5.31) we need other conditions to that vectors in the Bloch sphere must satisfy in order to represent states or operators. It was found that to ensure operators to be

Hermitian rank-one positive semidefinite we just need to satisfy the conditions:

$$\begin{aligned} i)\; Tr\left(\Pi\right) &= 1 \\ ii)\; Tr\left(\Pi^2\right) &= 1 \\ iii)\; Tr\left(\Pi^3\right) &= 1 \end{aligned} \tag{5.34}$$

Conditions i) and ii) are satisfied by any unitary vector in the bloch sphere. Hence, a SIC-POVM is a set of $d^2$ rank-one POVMs satisfying 5.25 i.e. unitary vectors pointing to the vertex of a regular simplex in the Bloch sphere that also satisfy iii).

First we notice that generators $T_b$ , $iT_b$, $I$ and $iI$ together span $\mathcal{M}_{d\times d}\left(\mathbb{C}\right)$, it follows that we have a multiplication law of the type [67]:

$$T_iT_j = \frac{2}{d}I\delta_{ij} + \sum_{k=1}^{d^2-1}\left\{g_{ijk} + if_{ijk}\right\}T_k \tag{5.35}$$

were $g_{ijk}$ are the symmetric structure constant and $f_{ijk}$ the antisymmetric structure constant, of the algebra of generators. Now, condition iii) in the Bloch sphere is:

$$\sum_{ijk} g_{ijk}r_ir_jr_k = \frac{d-2}{d}\sqrt{\frac{2d}{d-1}} \tag{5.36}$$

if we define $n\left(d\right) = \left[\left(d-2\right)/d\right]\sqrt{2d/\left(d-1\right)}$, then we also can define the vectors:

$$q_i\left(\vec{r}\right) = \frac{1}{n\left(d\right)}\sum_{jk} g_{ijk}r_jr_k \tag{5.37}$$

and iii) can be rewritten as:

$$\vec{r}\cdot\vec{q}\left(\vec{r}\right) = 1 \tag{5.38}$$

Now, from 5.34 we have $\rho = \rho^2$, $Tr(\rho^n) = 1$ and in particular $Tr(\rho^4) = 1$:

$$Tr(\rho^4) = \frac{1}{d} + 2\left\{\frac{4}{d^2}\left(\frac{d-1}{2d}\right)\left|\vec{r}\right|^2 + \frac{4}{d}\left(\frac{d-1}{2d}\right)\left(\frac{d-2}{d}\right)\vec{r}\cdot\vec{q}\left(\vec{r}\right) + \left(\frac{d-1}{2d}\right)\left(\frac{d-2}{d}\right)^2\left|\vec{q}\left(\vec{r}\right)\right|^2\right\} = 1 \tag{5.39}$$

Thus replacement of condition ii) and iii) i.e. $\left|\vec{r}\right|^2 = 1$ and $\vec{r}\cdot\vec{q}\left(\vec{r}\right) = 1$ in 5.39 implies:

$$\left|\vec{q}\left(\vec{r}\right)\right|^2 = 1 \tag{5.40}$$

Because $\vec{r}$ and $\vec{q}$ are unitary vectors, equation 5.38 forces them to be parallel, and then:

$$\vec{r} = \vec{q}\,(\vec{r}) \tag{5.41}$$

This shows that under assumptions i) and ii) condition 5.38 is equivalent to 5.41, because for the converse $\vec{r} \cdot \vec{q}\,(\vec{r}) = \vec{r} \cdot \vec{r} = |\vec{r}|^2 = 1$. In this way 5.41 defines the geometry of our projectors[3].

In the case of SIC-POVM we must require:

$$\vec{r}_\rho \cdot \vec{r}_\sigma = \frac{d^2 \delta_{\rho\sigma} - 1}{d^2 - 1} \tag{5.42}$$

This defines a system of quadratic and cubic polynomial equations in the components of the Bloch vectors representing SIC-POVMs and thus the existence of SIC-POVMs in a finite dimension $d$ is defined by the existence of solutions to the corresponding system of equations:

$$\begin{aligned} \vec{r}_\rho \cdot \vec{q}\,(\vec{r}_\rho) &= 1 \\ \vec{r}_\rho \cdot \vec{r}_\sigma &= \frac{d^2 \delta_{\rho\sigma} - 1}{d^2 - 1} \end{aligned} \tag{5.43}$$

for $\rho, \sigma = 1, \ldots, d^2$ or

$$\begin{aligned} \vec{r}_\rho &= \vec{q}\,(\vec{r}_\rho) \\ \vec{r}_\rho \cdot \vec{r}_\sigma &= \frac{d^2 \delta_{\rho\sigma} - 1}{d^2 - 1} \end{aligned} \tag{5.44}$$

In [66] a solution for $d = 3$ was obtained by solving a system of the form 5.44 and leads to projectors $\Pi_{i,k} = |\psi_{i,k}\rangle\langle\psi_{i,k}|$ onto the following normalized pure states

$$\begin{aligned} |\psi_{0,k}\rangle &= \frac{1}{\sqrt{2}}(|1\rangle + e^{-i(\theta_0 + \frac{2\pi k}{3})}|2\rangle), \\ |\psi_{1,k}\rangle &= \frac{1}{\sqrt{2}}(|1\rangle + e^{-i(\theta_1 + \frac{2\pi k}{3})}|3\rangle), \\ |\psi_{2,k}\rangle &= \frac{1}{\sqrt{2}}(|2\rangle + e^{-i(\theta_2 + \frac{2\pi k}{3})}|3\rangle), \end{aligned} \tag{5.45}$$

with $k = 0, 1, 2$ and where each angle $\theta_i$ with $i = 1, 2, 3$ belongs to the interval $[0, 2\pi)$.

But a generalization of the procedure has been difficult because the analytic value of the symmetric structure constant $g_{ijk}$. By changing the Gell-Mann diagonal matrices for those of the *Cartan subalgebra* generators, i.e. :

$$\tilde{T_b} = |b\rangle\langle b| - |b+1\rangle\langle b+1| \quad b = 1, \ldots d-1 \tag{5.46}$$

we solve that problem keeping all the other results and this development will be explored in future research.

[3] Is clear that the same calculations hold for pure states and thus we have also characterized the geometry of pure states in the Bloch sphere.

## 5.5 Characterization of SIC-POVMs using MUBs

In this section we reproduce our main results presented in our article *Characterization of fiducial states in prime dimensions via mutually unbiased bases* [68]. In our research we explore the existence conditions for SIC-POVMs in the particular case of prime dimensions, by applying the generators of the Weyl-Heisenberg group onto a special *fiducial* pure state.

### 5.5.1 Displacement operators

For every $k, s = 0, \ldots, d^2 - 1$. Considering the projecting directions we can study the problem to find $d^2$ pure quantum states $|\phi_k\rangle$ instead of $d^2$ SIC-POVM projectors. That is, assuming that $\Pi_k = |\phi_k\rangle\langle\phi_k|$ then the inner product between two arbitrary states must fulfill the property

$$|\langle\phi_k|\phi_r\rangle|^2 = \frac{d\delta_{k,s} + 1}{d+1} \quad \forall\; k, s = 0, \ldots, d^2 - 1. \tag{5.47}$$

There is a strong conjecture that simplifies the way to construct SIC-POVMs. In order to introduce this conjecture let us define first the displacement operators in finite dimension, given by

$$D_{\mathbf{r}} = \tau^{r_1 r_2} X^{r_1} Z^{r_2}, \tag{5.48}$$

where $\mathbf{r} = (r_1, r_2) \in \mathbb{Z}_d^2$ and $\tau = -e^{i\pi/d}$. The operators $X$ and $Z$ are the *shift* and *phase* operators, defined by

$$X|k\rangle = |k+1\rangle, \;\; Z|k\rangle = \omega^k|k\rangle, \tag{5.49}$$

where $\omega = e^{2\pi i/d}$, $k = 0, \ldots, d-1$ and $\{|k\rangle\}$ is the canonical (computational) base. If $d = 2$, the displacement operators are the Pauli matrices plus the identity. These operators form, up to a multiplicative constant factor, the generalized Pauli group or Weyl-Heisenberg group. Their commutation rule is given by

$$D_{\mathbf{r}} D_{\mathbf{q}} = \tau^{\langle\mathbf{r},\mathbf{q}\rangle - \langle\mathbf{q},\mathbf{r}\rangle} D_{\mathbf{q}} D_{\mathbf{r}}, \tag{5.50}$$

where $\langle\mathbf{r},\mathbf{q}\rangle = r_2 q_1 - q_2 r_1$ is a symplectic form. As we can see in the above equation, two displacement operators commute if and only if $\langle\mathbf{r},\mathbf{q}\rangle = \langle\mathbf{q},\mathbf{r}\rangle$. Let us now assume that $d$ is a prime number. It is easy to show that

$$D_{\mathbf{r}} = \begin{cases} \tau^{r_2 - r_1 r_2}(D_{\tilde{\mathbf{r}}})^{r_1} & \text{when } r_1 \neq 0, \\ (D_{(0,1)})^{r_2} & \text{when } r_1 = 0, \end{cases} \tag{5.51}$$

where $\tilde{\mathbf{r}} = (1, r_2 r_1^{d-2})$ and all operations are modulo $d$. Notice that the $d^2$ displacement operators can be written as a function of $d+1$ of them. Of course, the functions must be non linear, because all the operators are linearly independent. The eigenvector basis of the displacement operators in prime dimensions are a maximal set of $d+1$ MU bases. [71].

Zauner's conjecture [54] states that in every dimension $d$ there exists a *fiducial* state $|\phi\rangle$ such that $\{D_{\mathbf{r}}|\phi\rangle\}$ determines a SIC-POVM. That is,

$$|\langle\phi|D_{\mathbf{r}}|\phi\rangle|^2 = \frac{d\delta_{\mathbf{r},\vec{0}}+1}{d+1} \quad \forall\, \mathbf{r} \in \mathbb{Z}_d^2. \tag{5.52}$$

If a pure quantum state $|\phi\rangle$ satisfies the previous condition, then the SIC-POVM is given by $\{D_{\mathbf{r}}|\phi\rangle,\ \mathbf{r} \in \mathbb{Z}_d^2\}$. If such a construction is possible we say that the SIC-POVM is covariant under Weyl-Heisenberg group.

### 5.5.2 Fiducial states in MU bases decomposition

Any quantum state $\rho$ acting on a prime dimensional Hilbert space $\mathcal{H}$ can be written as a linear combination of rank-one projectors related to a complete set of $d+1$ mutually unbiased bases. That is,

$$\rho = \sum_{j=0}^{d}\sum_{k=0}^{d-1}\left(p_k^j - \frac{1}{d+1}\right)\Pi_k^j, \tag{5.53}$$

where

$$p_k^j = \mathrm{Tr}(\rho\Pi_k^j), \tag{5.54}$$

with $j$ the index of the MU bases family and $k$ the index within the family. We consider that $j=d$ corresponds to the canonical base (eigenvectors of $Z$). Sometimes, this base is denoted with the index $j=\infty$ (see [72]) and this choice is justified by arguments about the discrete affine plane picture [73]. Every operator $\Pi_k^j$ is a rank one projector, namely

$$\Pi_k^j = |\varphi_k^j\rangle\langle\varphi_k^j|, \tag{5.55}$$

where $\{|\varphi_k^j\rangle\}$ satisfy the relationship

$$|\langle\varphi_k^j|\varphi_s^l\rangle|^2 = \left\{\begin{array}{ll} \frac{1}{d} & \text{when } j \neq l, \\ \delta_{k,s} & \text{when } j = l, \end{array}\right. \tag{5.56}$$

that is, they form a complete set of $d+1$ MU bases. It is also possible to decompose a quantum state as a linear combination of the displacement operators

$$\rho = \sum_{\mathbf{r}\in\mathbb{Z}_d^2} a_{\mathbf{r}} D_{\mathbf{r}}, \tag{5.57}$$

where $a_{\mathbf{r}} \in \mathbb{C}$ and $a_{\mathbf{0}} = 1/d$. Zauner's conjecture given in Equation (5.52) can be cast now in the form

$$|\mathrm{Tr}(\rho D_{\mathbf{r}})|^2 = \frac{d\delta_{\mathbf{r},\vec{0}} + 1}{d+1} \quad \forall\, \mathbf{r} \in \mathbb{Z}_d^2. \tag{5.58}$$

Considering Equations (5.57) and (5.58) and the fact that set $\{D_{\mathbf{r}}\}$ of displacement operators form an orthogonal base we obtain

$$a_{\mathbf{r}} = \begin{cases} \frac{1}{d} & \text{when } \mathbf{r} = \mathbf{0}, \\ \frac{1}{d\sqrt{d+1}} w^{\beta_{\mathbf{r}}} & \text{when } \mathbf{r} \neq \mathbf{0}. \end{cases} \tag{5.59}$$

for a given set of real parameters $\beta_{\mathbf{r}} \in [0, d)$. It is easy to show that the following completeness relation holds

$$\sum_{r \in \mathbb{Z}_d^2} \frac{1}{d} D_{\mathbf{r}} \rho D_{\mathbf{r}}^{\dagger} = \mathbb{I}, \tag{5.60}$$

for any set $\{\beta_{\mathbf{r}}\}$. In order to deduce a relationship between the set of coefficients $\{p_k^j\}$ given by Equation (5.53) and the set of coefficients $\{a_{\mathbf{r}}\}$ given by Equation (5.57) we need an expression of the displacement operators as a function of the MU bases projectors. Taking into account that every operator $D_{\mathbf{r}}$ has the same set of eigenvalues $\{\omega^k\}$ (with $k = 0, \ldots, d-1$) we obtain

$$D_{\mathbf{r}} = \begin{cases} \tau^{r_2 - r_1 r_2} \sum_{k=0}^{d-1} \omega^{k r_1} \Pi_k^{\tilde{r}_2} & \text{when } r_1 \neq 0, \\ \sum_{k=0}^{d-1} \omega^{k r_2} \Pi_k^0 & \text{when } r_1 = 0. \end{cases} \tag{5.61}$$

Putting Equations (5.61) into Equation (5.57) and considering Equation (5.59) we have

$$\begin{aligned} \rho &= \sum_{r_2=1}^{d} \sum_{k=0}^{d-1} \frac{1}{d(d+1)} \Pi_k^{r_2} \\ &+ \sum_{r_2=1}^{d} \sum_{k=0}^{d-1} \sum_{r_1=1}^{d-1} a_{(r_1, r_1 r_2)} \tau^{r_1 r_2 - r_1^2 r_2} \omega^{k r_1} \Pi_k^{r_2} \\ &+ \sum_{k=0}^{d-1} \left( \frac{1}{d(d+1)} + \sum_{r_2=1}^{d-1} a_r \omega^{k r_2} \right) \Pi_k^0. \end{aligned} \tag{5.62}$$

From the last result and considering Equation (5.59), we have the following relationship

$$p_k^j = \frac{1}{d} + \frac{1}{d\sqrt{d+1}} \sum_{s=1}^{d-1} \omega^{\alpha_s^j + ks}, \tag{5.63}$$

for every $j = 0, \dots, d$ and $k = 0, \dots, d-1$, where the $d^2 - 1$ phases have the form

$$\omega^{\alpha_s^j} = \begin{cases} \omega^{\beta_{(s,js)}} \tau^{js - js^2} & \text{when } j \neq 0, \\ \omega^{\beta_{(0,s)}} & \text{when } j = 0, \end{cases} \tag{5.64}$$

Given that every probability $p_k^j$ is a real number and considering that $d$ is an odd prime number we find the symmetry

$$\alpha_s^j = -\alpha_{d-s}^j. \tag{5.65}$$

for every $j = 0, \dots, d$ and $s = 1, \dots, d-1$. Therefore, Equation (5.63) is reduced to

$$p_k^j = \frac{1}{d} + \frac{2}{d}\sqrt{\frac{1}{d+1}} \sum_{s=1}^{(d-1)/2} \cos(\alpha_s^j + 2\pi k r/d). \tag{5.66}$$

This characterization of the probability distributions for fiducial operators in MU bases decomposition is our first result in the corresponding article.

### 5.5.3 Existence conditions for a fiducial state in prime dimensions

Let us note that if $\mathrm{Tr}(\rho) = 1$ and $\mathrm{Tr}(\rho^2) = 1$ the only way to have a positive semidefinite operator $\rho$ is that $\mathrm{Tr}(\rho^3) = 1$. If this does not hold, then $\rho$ has one negative eigenvalue, at least. This fact is easy to understand because the only quantum states having $\mathrm{Tr}(\rho^2) = 1$ are the pure states. In our case, $\mathrm{Tr}(\rho) = 1$ is implicit in the MUBs decomposition given in Equation (5.53). We can also easily deduce from Equation (5.63) that

$$\sum_{j,k} (p_k^j)^2 = 2, \tag{5.67}$$

for every value of the phases $\{e^{i\alpha_r^j}\}$. On the other hand, it is known a general property relating probabilities in MUBs decomposition with the purity of a quantum state [74], namely:

$$\sum_{j,k} (p_k^j)^2 = \mathrm{Tr}(\rho^2) + 1. \tag{5.68}$$

Combining the last two equations we obtain

$$\mathrm{Tr}(\rho^2) = 1. \tag{5.69}$$

Therefore, we only need to impose the condition

$$\mathrm{Tr}(\rho^3) = 1, \tag{5.70}$$

in order to have a pure state. Taking into account Equation (5.70) and the MU bases decomposition given in Equation (5.62) we obtain

$$\sum_{\substack{j\ l\ s\\ k\ m\ n}} p_k^j\, p_m^l\, p_n^s \,\mathrm{Tr}\,(\Pi_k^j\, \Pi_m^l\, \Pi_n^s) = \mathrm{Tr}\,(\rho^3) + d + 6, \tag{5.71}$$

which must be satisfied by the probability distributions $\{p_k^j\}$ (depending on $\{\alpha_k^j\}$) to guarantee that they represent a physically admissible fiducial pure state. Equation (5.71) indicates that the problem of the existence of Weyl-Heisenberg covariant SIC-POVMs is equivalent to demonstrate that the function

$$F(\{p_k^j\}) = \sum_{\substack{j\ l\ s\\ k\ m\ n}} p_k^j\, p_m^l\, p_n^s \,\mathrm{Tr}\,(\Pi_k^j\, \Pi_m^l\, \Pi_n^s) \tag{5.72}$$

reaches a maximum value of $d+7$ for some $\rho$. We can affirm that the maximum value belongs to the interval $[d+5, d+7]$, but so far we have not been able to proof that $F = d+7$ is reached for every prime dimension $d$. Numerical evidence tells us that this maximal number is reached in every dimension $d \leq 67$ [70] and, consequently, we have a strong evidence of its existence in every prime dimension. Let us establish an upper bound for the probabilities $p_k^j$. From Equation (5.66) we can bound the cosines, obtaining

$$p_k^j \leq \frac{1}{d} + \frac{d-1}{d}\sqrt{\frac{1}{d+1}}, \tag{5.73}$$

for every $k = 0, \ldots, d-1$ and $j = 0, \ldots, d$. This bound is highly non-trivial and it may be useful to reduce computational time in order to find numerical solutions in higher dimensions.

## 5.6 Remarks on our results concerning SIC-POVM

Our achievements with our two strategies to prove the existence of SIC-POVM in finite dimensions are:

- Complete characterization of the necessary and sufficient conditions for SIC-POVM existence in a Geometrical style, for all finite dimension.
- Construction of a solution in $d = 3$ using our geometrical conditions.
- Characterization of fiducial state necessary and sufficient conditions to generate a SIC-POVM in prime dimensions
- A non-trivial upper bound for fiducial state parameters in prime dimensions

# 6

# Generalization of IC-POVMs

In this chapter we show in what sence tight rank one IC-POVMs are optimal for quantum tomography and as well a generalization of them that keeps the same sence of optimality. Also we propose a generalization of MUBs making use of equidistant states in the construction.

## 6.1 Optimal linear QST

Optimal QST is achieved by choosing a IC-POVM that minimizes the statistical errors for the tomography procedure. In Chapter 5 we introduced dual frames and the induced recosntruction operators $R_x$ which gives rise to the tomography reconstruction:

$$\rho = \sum_{x\in\mathcal{K}} p_x R_x \tag{6.1}$$

If now we proceed to make an estimation $\hat{p}_x$ of the probabilities $p_x$ this give us a estimation of the state $\rho$ as:

$$\hat{\rho} = \sum_{x\in\mathcal{K}} \hat{p}_x R_x \tag{6.2}$$

we will call $\hat{\rho}$ the *estimator* of the state $\rho$, for a given reconstruction operator set. In Chapter 4 we introduced the likelihood function for an estimator, that in this case we can take as:

$$\mathcal{L}\left(\hat{\rho}\right) = \prod_{j\in\mathcal{K}} \hat{p}_j^{n_j} \tag{6.3}$$

with $n_j = \sum_{x\in\mathcal{K}} \delta_{jx}$ . If $N$ is the total number of counts, the normalization constrain only, gives by maximum-likelihood the estimation $\hat{p}_j = n_j/N = f_j$; which is the relative frequency. Since $\hat{p}_j$ is a linear function of the counts $n_j$, this estimator is known as a *linear* estimator. A optimal QST using this particular linear estimator is also an optimal linear QST because the normalization constrain is a linear constrain.

The statistical error measure for this optimization will be the average over the square of the norm induced by the H-S product, of the difference between the reconstructed state and the linear estimator:

$$E\left(\hat{\rho}\right)=\left\langle \left\Vert \rho-\hat{\rho}\right\Vert ^{2}\right\rangle \tag{6.4}$$

a direct calculation leads to:

$$\begin{aligned}
\left\langle \left\Vert \rho-\hat{\rho}\right\Vert ^{2}\right\rangle &= \sum_{x,y\in\mathcal{K}}\left\langle \left(p_{x}-\hat{p}_{x}\right)\left(p_{y}-\hat{p}_{y}\right)\right\rangle \left(R_{x}|R_{y}\right)\\
&= \frac{1}{N}\sum_{x,y\in\mathcal{K}}\left(p_{x}\delta_{xy}-p_{x}p_{y}\right)\left(R_{x}|R_{y}\right)\\
&= \frac{1}{N}\left(\sum_{x\in\mathcal{K}}p_{x}\left(R_{x}|R_{x}\right)-Tr\left(\rho^{2}\right)\right)\\
&\equiv \frac{1}{N}\left(\Delta_{p}\left(R\right)-Tr\left(\rho^{2}\right)\right)
\end{aligned}$$

The error that we should minimize is:

$$E\left(\hat{\rho}\right)=\frac{1}{N}\left(\Delta_{p}\left(R\right)-Tr\left(\rho^{2}\right)\right) \tag{6.5}$$

Since we have no control over the purity of $\rho$, it is the quantity $\Delta_{p}\left(R\right)$ in equation (6.5) which is now of interest. The IC-POVM which minimizes $\Delta_{p}\left(R\right)$ , and hence the error, will in general depend on the quantum state under examination. We thus average over all unitarily equivalent states by seting $\rho=\rho\left(\sigma,U\right)=U\sigma U^{\dagger}$, and removing this dependence taking the Haar average over all $U\in U\left(d\right)$:

$$\begin{aligned}
\int_{U(d)}d\mu_{H}\left(U\right)\Delta_{p}\left(R\right) &= \int_{U(d)}d\mu_{H}\left(U\right)\sum_{x\in\mathcal{K}}Tr\left(F_{x}U\sigma U^{\dagger}\right)\left(R_{x}|R_{x}\right)\\
&= \frac{1}{d}\sum_{x\in\mathcal{K}}Tr\left(F_{x}\right)Tr\left(\sigma\right)\left(R_{x}|R_{x}\right)\\
&= \frac{1}{d}\sum_{x\in\mathcal{K}}Tr\left(F_{x}\right)\left(R_{x}|R_{x}\right)\\
&\equiv \frac{1}{d}\Delta_{F}\left(R\right)
\end{aligned}$$

In the second step of this calculation we evaluate the integral:

$$\Upsilon\left(\sigma\right)=\int_{U(d)}d\mu_{H}\left(U\right)U\sigma U^{\dagger} \tag{6.6}$$

To evaluate this integral we notice first that for any projector $P$ of rank one, $\Upsilon(P) = C$ with $C$ the same constant operator, because all projectors are unitarily equivalent. Also we have that $\sum_{i=1}^{d} P_i = I$ , then

$$dC = \sum_{i=1}^{d} \Upsilon(P_i) = I$$

and we have $C = \frac{1}{d} I$. Therefore for $A = \sum_{i=1}^{d} \lambda_i P_i$ we have,

$$\Upsilon(A) = \sum_{i=1}^{d} \lambda_i \Upsilon(P_i) = \sum_{i=1}^{d} \lambda_i C = \frac{1}{d} Tr(A) I$$

So, the evaluation of the integral for the required case is

$$\Upsilon(\sigma) = \frac{1}{d} Tr(\sigma) I \tag{6.7}$$

We will now minimize $\Delta_F(R)$ over all choices for $R$, while keeping the IC-POVM $F$ fixed. Our only constraint is that $\{R_x\}_{x\in\mathcal{K}}$ remains a dual frame to $\{P_x\}_{x\in\mathcal{K}}$ so that the reconstruction formula (6.1) remains valid for all $\rho$. The following Theorem shows that the reconstruction defined in (6.1) is the optimal choice for the dual frame.

THEOREM 6.1: *Let $\{A_x\}_{x\in\mathcal{K}} \subseteq \mathcal{B}(\mathbb{C}^d)$ be an operator frame. Then for all dual frames $\{B_x\}_{x\in\mathcal{K}}$,*

$$\Delta_F(B) \equiv \sum_{x\in\mathcal{K}} Tr(F_x)(B_x|B_x) \geq \sum_{x\in\mathcal{K}} Tr(F_x)\left(\tilde{A}_x|\tilde{A}_x\right) \equiv \Delta_F\left(\tilde{A}\right) \tag{6.8}$$

*with equality only if $B \equiv \tilde{A}$, where $\left\{\tilde{A}_x\right\}_{x\in\mathcal{K}}$ is the canonical dual frame.*

To prove it we define $D = B - \tilde{A}$, which satisfies:

$$\begin{aligned}
\sum_{x\in\mathcal{K}} Tr(F_x)|\tilde{A}_x)(D_x| &= \sum_{x\in\mathcal{K}} Tr(F_x)|\tilde{A}_x)(B_x| - \sum_{x\in\mathcal{K}} Tr(F_x)|\tilde{A}_x)(\tilde{A}_x| \\
&= \sum_{x\in\mathcal{K}} Tr(F_x)\mathcal{A}^{-1}|A_x)(B_x| - \sum_{x\in\mathcal{K}} Tr(F_x)\mathcal{A}^{-1}|A_x)(A_x|\mathcal{A}^{-1} \\
&= \mathcal{A}^{-1}\mathbf{I} - \mathcal{A}^{-1}\mathcal{A}\mathcal{A}^{-1} \\
&= 0
\end{aligned}$$

when $\{B_x\}_{x\in\mathcal{K}}$ is the dual frame to $\{A_x\}_{x\in\mathcal{K}}$ and $\left\{\tilde{A}_x\right\}_{x\in\mathcal{K}}$ is the canonical dual frame. Thus,

$$\sum_{x\in\mathcal{K}} Tr(F_x)\left(\tilde{A}_x|D_x\right) = 0 \tag{6.9}$$

Now, since $B = D + \tilde{A}$,

$$\begin{aligned}
\sum_{x\in\mathcal{K}} Tr\left(F_x\right)\left(B_x|B_x\right) &= \sum_{x\in\mathcal{K}} Tr\left(F_x\right)\left(\tilde{A}_x|\tilde{A}_x\right) + \sum_{x\in\mathcal{K}} Tr\left(F_x\right)\left(\tilde{A}_x|D_x\right) \\
&+ \sum_{x\in\mathcal{K}} Tr\left(F_x\right)\left(D_x|\tilde{A}_x\right) + \sum_{x\in\mathcal{K}} Tr\left(F_x\right)\left(D_x|D_x\right) \\
&= \sum_{x\in\mathcal{K}} Tr\left(F_x\right)\left(\tilde{A}_x|\tilde{A}_x\right) + \sum_{x\in\mathcal{K}} Tr\left(F_x\right)\left(D_x|D_x\right) \\
&\geq \sum_{x\in\mathcal{K}} Tr\left(F_x\right)\left(\tilde{A}_x|\tilde{A}_x\right)
\end{aligned}$$

with equality only if $D \equiv 0$. □
We also can see that:

$$\sum_{x\in\mathcal{K}} Tr\left(F_x\right)\left|R_x\right)\left(R_x\right| = \sum_{x\in\mathcal{K}} Tr\left(F_x\right)\mathcal{F}^{-1}\left|P_x\right)\left(P_x\right|\mathcal{F}^{-1} = \mathcal{F}^{-1}$$

and then $\Delta_F\left(R\right) = Tr\left(\mathcal{F}^{-1}\right)$, This quantity will now be minimized over all IC-POVMs. For this we will prove the following theorem:

THEOREM 6.2: *Let* $\{F_x\}_{x\in\mathcal{K}} \subseteq \mathcal{B}\left(\mathbb{C}^d\right)$ *be an IC-POVM. Then*

$$Tr\left(\mathcal{F}^{-1}\right) \geq d\left(d\left(d+1\right)-1\right) \tag{6.10}$$

*with equality iff* $F$ *is a tight rank-one IC-POVM.*
We will minimize the quantity

$$Tr\left(\mathcal{F}^{-1}\right) = \sum_{k=1}^{d^2}\frac{1}{\lambda_k} \tag{6.11}$$

where $\lambda_1, \ldots, \lambda_{d^2} > 0$ denote the left-right eigenvalues of $\mathcal{F}$. These eigenvalues satisfy the constraint

$$\sum_{k=1}^{d^2}\lambda_k = Tr\left(\mathcal{F}\right) = \sum_{x\in\mathcal{K}} Tr\left(F_x\right)\left(P_x|P_x\right) \leq \sum_{x\in\mathcal{K}} Tr\left(F_x\right) = d$$

since $Tr\left(P^2\right) \leq 1$. We know, however, that the identity operator is always a left-right eigenvector of $\mathcal{F}$ with unit eigenvalue. Thus we in fact have $\lambda_1 = 1$ say, and then $\sum_{k=2}^{d^2}\lambda_k \leq d-1$. Under this latter constraint it is straightforward to show that the RHS of equation (6.11) takes its minimum value if and only if $\lambda_2 = \cdots = \lambda_{d^2} = \left(d-1\right)/\left(d^2-1\right) = 1/\left(d+1\right)$, or equivalently,

$$\mathcal{F} = 1\frac{\mathcal{I}}{d} + \frac{1}{d+1}\Lambda_0 = \frac{\mathbf{I}+\mathcal{I}}{d+1} \tag{6.12}$$

with $\Lambda_0$ the projector into the subspace of traceless operators. From what we have seen in Chapter 5, this is also the case iff $F$ is a tight rank-one IC-POVM. So the minimum of $Tr\left(\mathcal{F}^{-1}\right)$ is attained iff $F$ is a tight rank-one IC-POVM $\square$.

We have thus confirmed that it is optimal to use a tight rank-one IC-POVM for quantum state tomography. Suppose now that we have a finite IC-POVM with $|\mathcal{K}| = d^2 + K$ possible measurement outcomes. We know that every POVM satisfies the normalization constraint, $\sum_{x\in\mathcal{K}} F_x = I$, which implies normalization of the statistics: $\sum_{x\in\mathcal{K}} p_x = 1$. Our previous estimate satisfies this constraint. It does not, however, incorporate any additional constraints specific to the particular choice of IC-POVM. Embedding the POVM in $\mathbb{I}^{\perp}\left(\mathbb{C}^d\right)$ shows that there will be a further $K$ linear constraints of the form:

$$\sum_{x\in\mathcal{K}} c_x^k F_x = 0 \;\rightarrow\; \sum_{x\in\mathcal{K}} c_x^k p_x = 0, \qquad \left(c_x^k \in \mathbb{R},\, k = 1,\ldots,K\right) \tag{6.13}$$

The intersection of the probability simplex in $\mathbb{R}^{d^2+K}$ with the subspace perpendicular to the $K$ vectors $\left\{c_x^k\right\}_{x\in\mathcal{K}}$ forms the subset of statistics which are isomorphic, under the mapping $p = Tr\,(FA) \rightarrow A$, to the normalized Hermitian operators in $\mathcal{B}\left(\mathbb{C}^d\right)$. Then, our estimator gives a full mapping only in the case of minimal IC-POVM, i.e. when $K = 0$ and no other constraints than the normalization are needed. For this reason minimal IC-POVMs (like the SIC-POVM) should be preferred over other IC-POVMs for a linear estimation. Under both the normalization and additional constraints this nonlinear optimization problem becomes difficult to solve analytically. One exception is an IC-POVM consisting of $d+1$ MUBs, in which case the $K = d$ additional constraints $\left[c^k\left(e_j^l\right) = (d+1)\,\delta_{kl} - 1\right]$ single out $\hat{p}\left(e_j^l\right) = n\left(e_j^l\right)/\left[(d+1)\sum_{k=1}^{d} n\left(e_j^l\right)\right]$ [53] for the maximum-likelihood estimate, and define a non-linear estimator when replaced in equation (6.2).

## 6.2 Conditional SIC-POVMs

We start decomposing $\mathcal{B}\left(\mathbb{C}^d\right)$ to three orthogonal subspaces:

$$\mathcal{B}\left(\mathbb{C}^d\right) = \mathbb{C}I \oplus \Theta_0^k \oplus \Theta_0^u \tag{6.14}$$

where $\mathbb{C}I = \{\alpha I : \alpha \in \mathbb{C}\}$ and $\Theta_0^k$, $\Theta_0^u$ are the a priori known and unknown traceless subspaces. This means that for a $\Lambda_k \in \Theta_0^k$ and a $\Lambda_u \in \Theta_0^u$ we have:

$$\rho = \frac{1}{d} I + \Lambda_k \rho + \Lambda_u \rho \tag{6.15}$$

being $\Lambda_k \rho$ the known traceless part of $\rho$ and $\Lambda_u \rho$ the unknown traceless part of $\rho$. We use the notation $\rho_* = \rho - \Lambda_k \rho$. Then, our goal is the tomography of $\rho_*$. If the dimension of $\Theta_0^k$ is $m$, then the dimension of $\Theta_0^u$ is $d^2 - m - 1$. For the state estimation we have to use a POVM with at least $D = d^2 - m$ elements. To get a unique solution we will use a POVM with exactly $D$ elements $\{F_1, \ldots, F_D\}$. For obtaining the optimal POVM we propose a reconstruction of the form:

$$\hat{\rho}_* = \sum_{x \in \mathcal{K}} \hat{p}_{x*} R_x \tag{6.16}$$

the terms defined in analogy to those of the previous section. Also the optimization is carried on the quantity $E\left(\hat{\rho}_*\right) = \left\langle \|\rho_* - \hat{\rho}_*\|^2 \right\rangle$ . All quantities of the last section on optimal linear QST can be *-defined (when needed), and the same calculations carried on in exactly the same way until the stage of evaluating $Tr\left(\mathcal{F}^{-1}\right)$ where the eigenvalues of $\mathcal{F}$ takes the values, $\lambda_{1*} = 1$ and $\lambda_{2*} = \cdots = \lambda_{d^2*} = (d-1)/(D-1)$ because now $|\mathcal{K}| = D$.

Now, this POVM is such that:

$$\mathcal{F} = \frac{\mathcal{I}}{d} + \frac{d-1}{D-1}\Lambda_u \tag{6.17}$$

From this:

$$\sum_{x \in \mathcal{K}} \left[Tr\left(F_x\right)\right] |P_x)\,(P_x| = \frac{\mathcal{I}}{d} + \frac{d-1}{D-1}\Lambda_u$$

If we define $|Q_y) = |P_y) - \mu|I)$ , $\mu \in \mathbb{R}$ we have:

$$\sum_{x \in \mathcal{K}} \left[Tr\left(F_x\right)\right] (Q_y|P_x)\,(P_x|Q_y) = (Q_y|\,\frac{\mathcal{I}}{d} + \frac{d-1}{D-1}\Lambda_u\,|Q_y) \tag{6.18}$$

Since $(P_x|Q_y) = Tr\left(P_x P_y\right) - \mu$ the left hand side of (6.18) becomes:

$$\sum_{x \in \mathcal{K}} \left[Tr\left(F_x\right)\right] (Q_y|P_x)\,(P_x|Q_y) = \left[Tr\left(F_y\right)\right](1-\mu)^2 + \sum_{x \neq y} \left[Tr\left(F_x\right)\right]\left(Tr\left(P_x P_y\right) - \mu\right)^2$$

and the right hand side:

$$\begin{aligned} \frac{\mathcal{I}}{d}\left[|P_y) - \mu|I)\right] = \frac{\mathcal{I}}{d}|P_y) - \mu|I) = {} & \tfrac{\mathcal{I}}{d}\left[|P_y) - \tfrac{1}{d}|I)\right] + \tfrac{1}{d}|I) - \mu|I) = & \left(\frac{1}{d} - \mu\right)|I), \\ \rightarrow (Q_y|\,\frac{\mathcal{I}}{d}\,|Q_y) = {} & \left(\tfrac{1}{d} - \mu\right) Tr\left(P_y - \mu I\right) = & d\left(\frac{1}{d} - \mu\right)^2 \end{aligned}$$

Also in Chapter 5 we introduced a Bloch representation for states and normalized operators. Using the Bloch representation of $P_y$ we have:

$$\Lambda_u\,|Q_y) = \sqrt{\frac{d-1}{2d}} \sum_{T_b \in \Theta_0^u} r_b T_b \;\rightarrow (Q_y|\,\Lambda_u\,|Q_y) = \left(\frac{d-1}{d}\right) \sum_{T_b \in \Theta_0^u} r_b^2$$

So, (6.18) becomes:

$$[Tr\,(F_y)]\,(1-\mu)^2+\sum_{x\neq y}[Tr\,(F_x)]\,(Tr\,(P_xP_y)-\mu)^2 = d\left(\frac{1}{d}-\mu\right)^2+\left(\frac{d-1}{D-1}\right)\left(\frac{d-1}{d}\right)\sum_{T_b\in\Theta_0^u} r_b^2 \tag{6.19}$$

but in our Bloch representation $\sum_{T_b\in\Theta_0^u} r_b^2 \leq 1$ , this implies

$$[Tr\,(F_y)]\,(1-\mu)^2 \leq d\left(\frac{1}{d}-\mu\right)^2+\left(\frac{d-1}{D-1}\right)\left(\frac{d-1}{d}\right) \tag{6.20}$$

which is true for every value of $\mu$, so

$$[Tr\,(F_y)] \leq \min_{\mu}\left\{\frac{d\left(\frac{1}{d}-\mu\right)^2+\left(\frac{d-1}{D-1}\right)\left(\frac{d-1}{d}\right)}{(1-\mu)^2}\right\}$$

By differentiating we obtain that the right hand side is minimal if:

$$\mu = \frac{D-d}{d\,(D-1)} = \frac{\tilde{D}-1}{D-1}$$

with $\tilde{D} = D/d$. Then we get,

$$[Tr\,(F_y)] \leq \frac{1}{\tilde{D}}, \quad \forall y \in \mathcal{K}$$

From $\sum_{x\in\mathcal{K}}[Tr\,(F_x)] = d$, we have $Tr\,(F_1) = Tr\,(F_2) = \ldots = Tr\,(F_D) = 1/\tilde{D}$. By replacement of this in (6.20) and an appropriate rearrangement of the terms, we have:

$$\left(\frac{d-1}{D-1}\right)\left(\frac{d-1}{d}\right)\left[1-\sum_{T_b\in\Theta_0^u} r_b^2\right]+\frac{1}{\tilde{D}}\sum_{x\neq y}\left(Tr\,(P_xP_y)-\frac{\tilde{D}-1}{D-1}\right)^2 = 0 \tag{6.21}$$

Because $d>1$, $D>1$ and all terms on the left hand are strictly non-negative, the only possible solution is when each term is zero. This implies that,

$$\sum_{T_b\in\Theta_0^u} r_b^2 = 1 \quad \Rightarrow \quad r_b = 0, \text{if}\, T_b\in\Theta_0^k \quad \Rightarrow \quad Tr\,(T_bP_x) = 0, \text{if}\, T_b\in\Theta_0^k$$

Also that,

$$Tr\,(P_xP_y) = \frac{\tilde{D}-1}{D-1} \quad \text{if}\, x\neq y$$

So the optimal IC-POVM for tomography, when there is a conditioning a priori known part $\Lambda_k\rho$ of the density matrix is a tight rank one IC-POVM $\{F_1, \ldots, F_D\}$ such that[1],

$$F_i = \frac{1}{\tilde{D}}\Pi_i, \quad Tr\left(\Pi_i\Pi_j\right) = \frac{\tilde{D}-1}{D-1} \quad \text{if}\, i \neq j, \quad Tr\left(T_b\Pi_i\right) = 0 \quad \text{if}\, T_b \in \Theta_0^k \tag{6.22}$$

We have used the symbol $\Pi_i$ as in the case of SIC-POVM because this conditional IC-POVM has the same symmetry of SIC-POVMs in the $\Theta_0^u$ subspace; also they generalize SIC-POVMs, since in the case $\Theta_0^k = \mathbf{null}$ we have $D = d^2$, $\tilde{D} = d$ and in consequence

$$Tr\left(\Pi_i\Pi_j\right) = \frac{1}{d+1} \quad \text{if}\, i \neq j$$

we have the SIC-POVMs as a special case. For this reason we call this conditional IC-POVM, conditional SIC-POVM or CSI-POVM for shorter.

## 6.3 Existence of conditional SIC-POVM

In this section we examine some particular cases where is possible to prove the existence of CSI-POVMs. As we stated at the end of the previous section, when $\Theta_0^k = \mathbf{null}$ the CSI-POVMs existence is reduced to that of the SIC-POVMs, whose existence has been proved analytically or numerically only for $d \leq 67$ and further research is going in this field as we explained in Chapter 5.

Other option is to assume $\Theta_0^k$ the off-diagonal elements of $\rho$, and we want to estimate the diagonal entries ($m = d^2 - d$, $D = d$), then it follows that the CSI-POVM has the properties:

$$F_i = \Pi_i, \quad Tr\left(\Pi_i\Pi_j\right) = 0 \quad \text{if}\, i \neq j, \quad \Pi_i \quad \text{is diagonal.}$$

So, the diagonal matrix projectors $H_k$ form the desired CSI-POVM and they exist for all dimensions.

Take now the special case when $\Theta_0^k$ are the diagonal terms of the density matrix, this is:

$$\Lambda_k\rho = \sum_{k=0}^{d-1}\left(p_k - \frac{1}{d}\right) H_k \tag{6.23}$$

In this case $m = d-1$ and $D = d^2 - d + 1$, $\mu = \frac{d-1}{d^2}$.

In what follows we reproduce the proof of Petz et. al. [76] of the following theorem:

[1] A common abuse of notation in literature is to identify the normalization $\Pi_i$ of a SIC-POVM or CSI-POVM element with the element itself. While the calculations and tomographic schemes be expressed in terms of the $\Pi_i$ one should not forget that only the $F_i$ are the proper POVM elements, since only them satisfy in general the completeness relation.

THEOREM 6.3: *A CSI-POVM exist with respect to the diagonal part ofa density matrix if $d-1$ is a power of a prime.*

First we start with the fiducial state:

$$|\phi\rangle = \frac{1}{\sqrt{d}}\sum_{j=0}^{d-1} |j\rangle$$

And the unitary operators:

$$X = \sum_{j=0}^{d-1} |j+1\rangle\langle j|\,, \quad U = \sum_{j=0}^{d-1} \omega^{\alpha_j} |j\rangle\langle j| \qquad \omega = e^{2\pi i/D}$$

were $\alpha_j \in \mathbb{Z}_D$. A direct calculation shows that: $X|\phi\rangle = X^\dagger|\phi\rangle = |\phi\rangle$ and $\left|\left\langle U^k\phi, j\right\rangle\right|^2 = |\langle\phi,0\rangle|^2$. From this:

$$Tr\left\{\left|U^k\phi\right\rangle\left\langle U^k\phi\right|(|j\rangle\langle j| - I/d)\right\} = 0$$

And thus the $D$ projectors $\Pi_{k+d} = \left|U^k\phi\right\rangle\left\langle U^k\phi\right|$ are orthogonal to the diagonal projections $H_k$ ( and in consequence to $\sigma_k$ ).

Now we prove that they satisfy a completeness relation:

$$\sum_k \left\langle i, U^k\phi\right\rangle\left\langle U^k\phi, j\right\rangle = \frac{1}{d}\sum_k \omega^{-\alpha_i k}\omega^{\alpha_j k} = \frac{1}{d}\sum_k \omega^{(\alpha_j-\alpha_i)k} = \tilde{D}\delta_{ij}$$

Then:

$$\sum_k \left|U^k\phi\right\rangle\left\langle U^k\phi\right| = \tilde{D}I \tag{6.24}$$

Now we have to choose the numbers $\alpha_0, \alpha_1, \ldots, \alpha_{d-1}$ such that:

$$Tr\{\Pi_i\Pi_j\} = \left|\left\langle U^{i-d}\phi, U^{j-d}\phi\right\rangle\right|^2 = \frac{1}{d^2}\left|\sum_m \omega^{(j-i)\alpha_m}\right|^2 = \frac{\beta}{d^2}$$

but we also have:

$$\tilde{D} = Tr\left(\Pi_i \sum_j \Pi_j\right) = \sum_j Tr\{\Pi_i\Pi_j\} = 1 + (D-1)\frac{\beta}{d^2}$$

This fixes $\beta = d-1$. Let $G := \{0, 1, \ldots, D-1\}$ be an additive group mod $D$ and the subset $W := \{\alpha_0, \alpha_1, \ldots, \alpha_{d-1}\}$ a difference set with parameters $(D, d, \lambda)$when the difference set $\alpha_j - \alpha_i$ contains all non null elements of $G$ exactly $\lambda$ times.

When this happens we have:

$$\left|\sum_{m=0}^{d-1}\omega^{n\alpha_m}\right|^2=\sum_{i,j=0}^{d-1}\omega^{n(\alpha_j-\alpha_i)}=d+\sum_{s=1}^{D-1}\lambda\omega^s=d-\lambda$$

in our case we have $\lambda=1$. Then, if the appropriate difference set exist, the corresponding CSI-POVM also exist. The problem of existence of difference sets with parameters $(D,d,1)$ is well known in the literature [78, 79]. Its known that a sufficient condition for the existence of a difference set with parameters $\left(d^2-d+1,d,1\right)$ is that $d-1$ be the power of a prime $p^n$ [79]□.

From this we see that CSI-POVM tomographic scheme is possible when the apriori information is the diagonal elements of the density matrix and the dimension of the state is $d=p^n+1$.

## 6.4 Generalization of MUBs

In this section we use the previous results of this chapter, Chapter 3 and Chapter 5 to propose a generalization of MUBs, which is still an open problem and a new research line. This research proposal of us, is on the draft stage and we expect to submit for publication soon.

### 6.4.1 Complementary decomposition

Let $\mathcal{A}$ and $\mathcal{B}$ be two subalgebras of $M_d(\mathbb{C})$, they are never orthogonal because identity is element of both subalgebras. They are called quasi-orthogonal if their traceless part are orthogonal:

$$\mathcal{A}^-\perp\mathcal{B}^-,\quad where\quad \mathcal{A}^-=\{A\in\mathcal{A}\mid A-\frac{I}{d}\cdot Tr(A)\}. \tag{6.25}$$

We call them *complementary subalgebras* [76, 77]. A *complementary decomposition* to subalgebras is defined by a set of algebras $\mathcal{A}_i$, $(i=1,2,\ldots,m)$ such that $\mathcal{A}_i$ and $\mathcal{A}_j$ are quasi-orthogonal for all $i\neq j$ $(\mathcal{A}_i^-\perp\mathcal{A}_j^-)$ and the set satisfies:

$$M_d(\mathbb{C})=\mathbb{C}I\oplus\mathcal{A}_1^-\oplus\mathcal{A}_2^-\oplus\ldots\oplus\mathcal{A}_m^- \tag{6.26}$$

If we have a state $\rho$ then we denote its projection on to subalgebra $\mathcal{A}_i$ with $\rho_i$, and its traceless part with $\rho_i^-=\rho_i-I/d$. Then, we can decompose a state with respect to the complementary decomposition:

$$\rho=\frac{I}{d}+\rho_1^-+\rho_2^-+\cdots+\rho_m^-. \tag{6.27}$$

Let us assume, that we want to estimate the state only on subalgebra $\mathcal{A}_l$. If $\dim\mathcal{A}_l=d_l$, then $\{F_1,F_2,\ldots F_{d_l}\}$ produce the optimal linear estimation on $\rho_l$ according to section (6.2), if it satisfies the following conditions:

$$F_i = \frac{d}{d_l}\Pi_i, \quad Tr\left(\Pi_i \Pi_j\right) = \frac{d_l - d}{d\left(d_l - 1\right)}\ \left(i \neq j\right), \quad \Pi_i \in \mathcal{A}_i$$

Which is the CSI-POVM just studied in this chapter, but proposed for a tomography scheme on each of the subspaces defined for the corresponding subalgebra.

### 6.4.2 Mutually unbiased equidistant bases

At the end of Chapter 3 we defined the set of Equidistant states as the pure states such that their vectorial representation $\{|\alpha_1\rangle, |\alpha_2\rangle, \ldots, |\alpha_d\rangle\} \in \mathbb{C}^d$ has the property:

$$\langle\alpha_i|\alpha_j\rangle = \alpha \quad (i \neq j)\ \alpha \in \mathbb{C}$$

Based on previous research [38, 39], we also shown that just from the knowledge of $\alpha$ is possible to construct the set of equidistant states in a closed analytic form and also the restrictions under which the equidistant states are a L.I. set and then a base for $\mathbb{C}^d$.

Now we assume the existence of a decomposition of $M_d(\mathbb{C})$ like in (6.26), but such that all the quasi-orthogonal subalgebras are maximally abelian, i.e. they satisfy the isomorfism:

$$\mathcal{A}_l \cong \mathbb{C}^{d_l}$$

now we assign a set of complex numbers $\alpha_l$ to each $\mathcal{A}_l$ such that the equidistant states associated with each number form a base in the corresponding $\mathbb{C}^{d_l}$. We will call the bases constructed in this way mutually unbiased equidistant bases (MUEBs). When all the $\alpha_l$ of a MUEBs satisfy

$$\|\alpha_l\|^2 = \frac{d_l - d}{d\left(d_l - 1\right)} \tag{6.28}$$

then the set of MUEBs are equivalent to a CSI-POVM for the decomposition in maximal abelian subalgebras. Because of this, when the set of MUEBs corresponding to the decomposition satisfy (6.28) then they are an optimal POVM for linear estimation. It must be notice that the allowed values of $\alpha_l$ are restricted in the way we exposed in Chapter 3, as well because of (6.28) we must take care to choose $d_l \geq d$.

We will call the MUEBs constructed in this way, the canonical MUEBs optimal POVM.

Ordinary MUBs are a special case of our canonical MUEBs, because they are the canonical MUEBs when a subalgebra decomposition can be choosen as $\mathcal{A}_1 \cong \mathcal{A}_2 \cong \ldots \cong \mathcal{A}_{d+1} \cong \mathbb{C}^d$. Is in this sense that MUEBs are an optimal POVM that generalize that of MUBs.

Something interesting about the relation between canonical MUEBs and MUBs is that the Bloch representation (previously exposed in Chapter 5) of canonical MUEBs are orthogonal sets of vectors each of them in a real subspace of dimention $d_l - 1$ such that the scalar product between the unitary vectors is:

$$\mathbf{r}_i^l \cdot \mathbf{r}_j^l = -\frac{1}{d_l - 1} \quad (i \neq j)$$

forming a regular simplex in their subspaces and in this way they have the same symmetry of MUBs in each subspace. We still work on the construction of a particular example of a MUEB, searching for appropriate decompositions in quasiorthogonal maximally abelian subalgebras.

### 6.4.3 Conditional MUEBs

A next step in generalization is to define MUEBs in the case when we need to make a estimation in a set of quasiorthogonal subalgebras that are not necessarily a complete decomposition. So in this case we assume we have a maximal abelian subalgebra (or a set of them) quasiorthogonal to the rest of the decomposition of $M_d(\mathbb{C})$ (let´s say $\mathcal{A}_l \cong \mathbb{C}^{d_l}$) and for this subalgebra we apply the same kind of construction for a MUEB and a canonical MUEB, being the canonical MUEB optimal and equivalent to a CSI-POVM.

## 6.5 Remarks on this Chapter

In this chapter we have:

- Review in detail the proofs of optimality for the IC-POVMs known as SIC-POVM and conditional SIC-POVM (CSI-POVM).
- Proposed a new generalization of MUBs that we named MUEBs and its conditional form.

We are actually developing the MUEBs tomography, but their versatility and similar optimal features to MUBs suggest that their use in QST will be in dimensions where we lack a complete set of MUBs. One interesting example that we could research in the future is dimension $d = 6$, where only three MUBs are known, which means that a measurement in this bases fixes three complementary subalgebras each isomorphic to $\mathbb{C}^6$, and to complete the tomography, at least we should find a complementary subalgebra $\mathcal{A}_4 \cong \mathbb{C}^{21}$ which would define a conditional MUEB.

We conjecture that MUEBs optimize a nonlinear estimator of the same structure as that of MUBs. Is likely that this tomography improves CSI-POVM tomography because this ones are implemented to optimize a linear estimator while MUEBs will optimize a nonlinear estimator, which in the case of MUBs improves the tomography. We expect a tomographic scheme with efficiency between that of SIC-POVM and MUBs.

# 7 Quantum Bayesianism

In this Chapter we discuss the fundamental results of Quantum Bayesianism in the field of QST, first the de Finetti theorem and then the SIC-POVM probability assignment formula. After this we show our contributions in the same line of research.

## 7.1 The meaning of Quantum states

The outcomes of Quantum Mechanics are the probabilities to get some experimental results every time a measurement is performed over a physical system. Quantum Bayesianism consist in a reformulation program of quantum mechanics that is consequence of giving to its probabilities outcomes a Bayesian interpretation [80, 88].

As we explained in Chapter 2 Bayesian interpretation of probabilities says that probability is a measure of plausibility. As in formal logic that can not assure any set of truth values of the propositions it manipulates, but instead only show whether various truth values are inconsistent, In the same way the desiderata of consistency from which the sum and product rules come out we have as a consequence that probability theory is of the character of formal logic. A probability assignment is a tool an agent uses to make decisions and inferences.

Quantum states, through the born rule, can be used to calculate probabilities and because of this if one assigns probabilities for the outcomes of a well-selected set of measurements (As we have stated in Chapters 4 and 5 this means a IC-POVM), then this is mathematically equivalent to making the quantum-state assignment itself. Thus, if probabilities are personal in the Bayesian sense, then quantum states must be too.

This last statement has consequences for our informational approach, since terms as *unknown* quantum state make no sense at all. Because of this an answer is required to questions like: *Which is the operational meaning of quantum state tomography and estimation?*. In Chapter 4 and 5 we took care to avoid the concept of unknown quantum state and always talk of an state *assignment* to a quantum *uncharacterized* system, but the idea that we are trying to discover a *true* quantum state for each preparation may have been implicitly assumed, specially in Chapters 5 and 6 when we talk of reconstruction operators and make distinction between the state $\rho$ and its estimator $\hat{\rho}$. Now we clarify this possible misunderstanding by exposing one of the results of

Quantum Bayesianism: the Quantum de Finetti theorem.

### 7.1.1 Classical de Finetti theorem

In classical probability de Finetti theorem is a result that also explain the use of the concept of *true* probabilities for the determination of probability distributions in repeated experiments [86]. Here the individual trials are described for example by $N$ discrete random variables $x_n$ that assume $k$ discrete values, i.e. $x_n \in \{1, 2, \ldots, k\}$, $n = 1, \ldots, N$. In a frequentist theory, such an experiment has a standard formulation in which the probability in the multi-trial hypothesis space is given by an independent, identically distributed (i.i.d.) distribution:

$$p\left(x_1, \ldots, x_N\right) = p_{x_1} \cdots p_{x_N} = p_1^{n_1} \cdots p_k^{n_k} \tag{7.1}$$

The number $p_j, \quad j = 1 \ldots k$ describes the *true* probability that the result of a single experiment will be $j$, while the variable $n_j$ is the number of times outcome $j$ is listed in the outcome vector $(x_1, \ldots, x_N)$. To the experimentalist, the *true* probabilities $p_1 \cdots p_k$ will in general be *unknown* at the outset. Thus, a frequentist has the duty to estimate the unknown probabilities by a statistical analysis of the experiment´s outcomes.

In the Bayesian approach, it does not make sense to talk about estimating a true probability. Instead, a Bayesian assigns a prior probability distribution $p\left(x_1, \ldots, x_N\right)$ on the multi-trial hypothesis space, which is generally not a i.i.d., and then uses Bayes´ theorem to update the distribution in the light of measurement results [7, 8]. As mentioned in Chapter 2, the task of choosing a prior distribution is not always trivial and a field of actual research [82]. The de Finetti representation theorem makes also this task more tractable.

In the present case, the key feature is contained in the assumption that an arbitrary number of repeated trials are equivalent. This means that one has no reason to believe there will be a difference between one trial and the next. In this case, the prior distribution is judged to have a sort of permutation symmetry which de Finetti [86] called *exchangeability*. The definition of exchangeability proceeds in two stages:

1. A probability distribution $p\left(x_1, \ldots, x_N\right)$ is said to be *finitely exchangeable* (f.e.) if is invariant under permutations of its arguments, i.e., if:

$$p\left(x_{\pi(1)}, \ldots, x_{\pi(N)}\right) = p\left(x_1, \ldots, x_N\right) \tag{7.2}$$

for any permutation $\pi$ of the set$\{1, \ldots, N\}$.

2. The distribution $p\left(x_1, \ldots, x_N\right)$ is called *exchangeable* if it is f.e. and if for any integer $M > 0$, there is a f.e. distribution $p_{(N+M)}\left(x_1, \ldots, x_N, x_{N+1}, \ldots, x_{N+M}\right)$ such that

$$p\left(x_1, \ldots, x_N\right) = \sum_{x_{N+1}, \ldots, x_{N+M}} p_{(N+M)}\left(x_1, \ldots, x_N, x_{N+1}, \ldots, x_{N+M}\right) \tag{7.3}$$

This last statement means the distribution $p$ can be extended to a f.e. distribution of arbitrary many random variables.

We are now prepared to enunciate the de Finetti theorem:

THEOREM 7.1 (DE FINETTI): *If a probability distribution $p\left(x_1,\ldots,x_N\right)$ is exchangeable, then it can be written uniquely in the form:*

$$p\left(x_1,\ldots,x_N\right)=\int_{S_k}P\left(\mathbf{p}\right)p_{x_1}\cdots p_{x_N}d\mathbf{p}=\int_{S_k}P\left(\mathbf{p}\right)p_1^{n_1}\cdots p_k^{n_k}d\mathbf{p} \tag{7.4}$$

*where $\mathbf{p}=\left(p_1,\ldots,p_k\right)$ and $P\left(\mathbf{p}\right)$ is a probability density function on the probability simplex $S_k$.*

Equation (7.4) comprises the classical de Finetti representation theorem for discrete random variables, for a proof we suggest the references [85, 86, 87].

The content of this result is that an agent, making solely the judgment of exchangeability for a sequence of random variables $x_j$, can proceed *as if* his state of knowledge had instead come about through ignorance of an *unknown*, but objectively existent set of probabilities $\mathbf{p}$. His precise ignorance of $\mathbf{p}$ is captured by the probability density $P\left(\mathbf{p}\right)$. This is in direct analogy to what we desire of a solution to the problem of the unknown quantum state in quantum state tomography.

### 7.1.2 Quantum de Finetti theorem

In QST the equivalent element to an i.i.d. distribution of probabilities is the ensemble of identically prepared systems, over which the different measurements take place. The relevant point here is that there is no distinction between the systems the device is preparing. In operational terms, this is the judgment that *all the systems are and will be the same as far as observational predictions are concerned.* Because of this, if the experimenter judges a collection of $N$ of the device´s outputs to have an overall quantum state $\rho^{(N)}$, he will also judge any permutation of those outputs to have an overall quantum state $\rho^{(N)}$ , Moreover, he will do this no matter how large the number $N$ is. This, complemented only by the consistency condition that for any $N$ the state $\rho^{(N)}$ be derivable from $\rho^{(N+1)}$, makes possible a representation analogous to that of the classical de Finetti theorem.

From this comes the definition of the quantum version of exchangeability which is closely analogous to the classical definition. Again, the definition proceeds in two stages,

1. A joint state $\rho^{(N)}$ of $N$ systems is said to be *finitely exchangeable* (f.e.) if it is invariant under any permutation of the systems. If we expand $\rho^{(N)}$ with respect to any orthonormal tensor product basis on $\mathcal{H}_d^{\otimes N}$,

   $$\rho^{(N)}=\sum_{i_1,\ldots,i_N;j_1,\ldots,j_N}\varrho^{(N)}_{i_1,\ldots,i_N;j_1,\ldots,j_N}\left|i_1\right\rangle\cdots\left|i_N\right\rangle\left\langle j_1\right|\cdots\left\langle j_N\right| \tag{7.5}$$

   Then the f.e. condition for $\rho^{(N)}$ means:

   $$\varrho^{(N)}_{\pi(i_1),\ldots,\pi(i_N);\pi(j_1),\ldots,\pi(j_N)}=\varrho^{(N)}_{i_1,\ldots,i_N;j_1,\ldots,j_N} \tag{7.6}$$

   for any permutation $\pi$ of the set$\{1,\ldots,N\}$.

2. The state $\rho^{(N)}$ is said to be *exchangeable* if it is f.e. and if, for any $M > 0$, there is a f.e. state $\rho^{(N+M)}$ of $N+M$ systems such that the marginal density operator for $N$ systems is $\rho^{(N)}$, i.e.,

$$\rho^{(N)} = Tr_{\mathcal{H}_d^{\otimes M}}\left(\rho^{(N+M)}\right) \tag{7.7}$$

   where the trace is taken over the additional $M$ systems.

Before the explicit statement of the quantum de Finetti theorem, we point out that the above definition of exchangeable applies for the overall quantum state of the ensemble of identical preparations for a QST in the light of the primarily given reasoning. Now we have,

Theorem 7.2 (Quantum de Finetti): *If a state $\rho^{(N)}$ of $N$ systems is exchangeable, then it can be written uniquely in the form:*

$$\rho^{(N)} = \int_{\mathcal{H}_d^{\otimes N}} P(\rho)\, \rho^{\otimes N} d\rho \tag{7.8}$$

*where $d\rho$ being a suitable measure on density operator space $\mathcal{H}_d^{\otimes N}$ , $P(\rho)$ is a probability distribution over the density operators.*

For a proof see [85]. Now our Bayesian view of QST allow us to act *as if* $\rho^{(N)} = \rho^{\otimes N}$ encoding our ignorance about $\rho$ in $P(\rho)$. In this way our procedure for QST start by assigning a prior quantum state to the joint system composed of the $N$ systems, reflecting our prior state of knowledge, then we can update our density operator to reflect information gathered from measurements by a quantum version of Bayes´ theorem. Specifically, if measurements on $K$ systems yield results $D_K$, then the state of additional systems is constructed as in equation (7.8), but using an updated probability on density operators given by,

$$P(\rho|D_K) = \frac{P(\rho)\, P(D_K|\rho)}{P(D_K)} \tag{7.9}$$

Here $P(D_K|\rho)$ is the probability for the measurement results $D_K$ given the state $\rho$ and $P(D_K)$ a normalization constant as in the classical case. We can think now the process of QST as an updating of our knowledge of the whole joint system $\rho^{(N)}$ . In the hope that different experimenters with different a priori knowledge achieve an agreement over future probabilistic predictions for a sufficiently large data $D_K$ input. This is such that for different priors $P_i(\rho)$ $i = 1, 2$, the measurement results force a common state of knowledge in which any number $N$ of additional systems are assigned the product state $\rho_{D_K}^{\otimes N}$, i.e.,

$$\int_{\mathcal{H}_d^{\otimes N}} P_i(\rho|D_K)\, \rho^{\otimes N} d\rho \rightarrow \rho_{D_K}^{\otimes N} \tag{7.10}$$

independent of $i$, for $K$ sufficiently large.

Accordingly, quantum de Finetti theorem shows how our new view of quantum states gives a new insight into QST process, that justifies all previous procedures and tecniques for reconstruction and estimation, but also includes the full power of Bayesian update of probabilities and prior information knowledge in the field (as we notice in Chapter 2 this implies an improvement in prediction power for data with a high signal-to-noise ratio.) within a consistent approach.

## 7.2 Seeking a new Bayes probability law

Since Bayes´ theorem is of fundamental relevance in Bayesian analysis, an analogous representation for the probability assignments of quantum mechanics will strengthen the Quantum Bayesianism approach by the same reasons as in classical Bayesianism.

In this same line of thought, it was noticed by C. Caves [81] that SIC-POVM provide a simple tomographic reconstruction formula:

$$\rho = \sum_{k=1}^{d^2} \left[d(d+1)p_k - 1\right] F_k, \tag{7.11}$$

where $\{p_k\}$ is the set of probabilities associated to the elements of the SIC-POVM $F_k = \{\Pi_k/d\}$ acting on a quantum system described by the state $\rho$. This latter expression together with the Born's rule allow us to cast the set of probabilities $\{q_m\}$ associated to a different POVM $\{E_m\}$ as an affine mapping (or rescaling) of the Bayes law of total probability, that is

$$Q(E_m) = \sum_{k=1}^{d^2} \left[d(d+1)P(F_k) - 1\right] P(E_m|F_k), \tag{7.12}$$

were we have rewritten $p_k$ as $P(F_k)$, $q_m$ as $Q(E_m)$ and $P(E_m|F_k) = Tr(E_m F_k)$ is the conditional probability of a measurement result associated to the operator $E_m$ given the measurement result associated to the operator $F_k = \Pi_k/d$. Quantum bayesianism proposes Eq. (7.12) as the quantum rule for assigning probabilities, replacing Born's rule, and changing quantum states and operators for sets of probabilities $\{P(E_k)\}$ and sets of conditional probabilities $\{P(E_m|F_k)\}$. From what we have discussed it comes out that such a prescription for calculating the probabilities of measurement processes will be equivalent to the orthodox formulation of quantum mechanics every time our rule to specify the outcomes that give rise to probabilities $P(F_k)$ gives the same outcomes as a SIC-POVM.

Also, in this scheme prior distributions are introduced directly in the tomography by taking $P\left(F_k\right) = P\left(F_k|I\right)$ and $P\left(E_m|F_k\right) = P\left(E_m|F_k, I\right)$ (here $I$ is the prior information as described in Chapter 2) which is one of the most debated issues in the quantum tomography researchers community [82, 83, 84]. In the present work we provide a way to specify such outcomes (related to $P\left(F_k\right)$) that don´t require $F_k$ to be part of a SIC-POVM, but of a CSI-POVM. This new Bayes law of total probability also requires a different rescaling.

## 7.3 New route in the Bayesian ocean

In Chapter 6 we introduced the generalization of SIC-POVMs known as CSI-POVMs. In this section we use CSI-POVMs for showing that a new variety of Bayes law of total probability can be constructed for quantum probability assignments, giving the general formulation for some prior known coefficients of the density matrix and then showing the law for some CSI-POVMs whose existence has already been prove. This is a result we will publish soon in an article that is under development.

We now proceed to define the superoperator $\mathcal{G} = \sum_k |\tilde{\Pi}_k)(\tilde{\Pi}_k|$ where the $\tilde{\Pi}_k$ are the proyectors proportional to the $\tilde{F}_k$ elements of the CSI-POVM [1] and satisfy Eq. (6.22). Then,

$$\mathcal{G}|\tilde{\Pi}_j) = (1-\mu)|\tilde{\Pi}_j) + \mu\tilde{D}|I), \tag{7.13}$$

with $\tilde{D} = D/d$ and $\mu = (D-d)/d(D-1) = (\tilde{D}-1)/(D-1)$ is the Hilbert-Schmidt product between different elements of the CSI-POVM. Defining the auxiliary constant $\alpha = (1-\mu d)$ we can write the superoperator $\mathcal{G}$ as

$$\mathcal{G} = \tilde{D}\left(\alpha\mathbf{I} + (1-\alpha)\left(\frac{\mathcal{I}}{d}\right)\right). \tag{7.14}$$

The inverse of this superoperator is then given by

$$\mathcal{G}^{-1} = \frac{1}{\alpha\tilde{D}}\left(\mathbf{I} - (1-\alpha)\left(\frac{\mathcal{I}}{d}\right)\right). \tag{7.15}$$

We have then that the identity superoperator can be cast as

$$\mathbf{I} = \frac{1}{\alpha\tilde{D}}\sum_k\left\{|\tilde{\Pi}_k)(\tilde{\Pi}_k| - \frac{(1-\alpha)}{d}|I)(\tilde{\Pi}_k|\right\}. \tag{7.16}$$

Using this identity we can write a reduced state $\tilde{\rho}$ as $\tilde{\rho} = \mathbf{I}|\tilde{\rho})$ or equivalently

$$\tilde{\rho} = \frac{1}{\alpha}\sum_k\left\{\tilde{\Pi}_k - \frac{(1-\alpha)}{d}I\right\}Tr(\tilde{F}_k\tilde{\rho}). \tag{7.17}$$

From this latter equation we obtain finally

$$\tilde{\rho} = \frac{1}{\alpha}\sum_k\left\{\tilde{D}P(\tilde{F}_k) - \frac{(1-\alpha)}{d}\right\}\tilde{F}_k, \tag{7.18}$$

Which is the CSI-POVM tomographic reconstruction formula analogue to Eq. (7.11). This latter equation can be recovered from Eq. (7.18) considering the case $D = d^2$.

Upon Eq. (7.18) we build the analogy to the Bayes law. In order to do this we need first to specify the known part of the state $\Lambda_k\rho$ which defines the CSI-POVM. In our case we take into account the diagonal traceless part of $\rho$ as in equation 6.23 where $p_k = Tr(\rho H_k) = P(H_k)$ with $\{H_k = |k\rangle\langle k|\}$ a POVM formed by normalized projectors

[1] To avoid any misunderstanding, in this chapter we will use the $\tilde{(\cdot)}$ hat to identify CSI-POVM probabilities and operators, since they are used to expand the reduced state $\tilde{\rho}$.

onto the canonical base $\{|k\rangle\}$ (with $k = 1, \ldots, d$). This choice of a priori information is motivated by the fact that $\Lambda_k \rho$ clearly implies a Bayes law plus a constant. Let us now examine the CSI-POVM when the a priori information are the diagonal terms of the traceless part of the density matrix. In this case we have $m = d-1$, $D = d^2 - d + 1$ and $\mu = (d-1)/d^2$. As we proved in Chapter 6, in this case CSI-POVM exists whenever the dimension of the Hilbert space is given by a positive integer power $n$ of an arbitrary prime number $p$ plus one, that is $d = p^n + 1$. For this values the tomographic formula for $\tilde{\rho}$ becomes

$$\tilde{\rho} = \sum_k \left\{ (d^2 - d + 1)\tilde{p}_k - \frac{d-1}{d} \right\} \tilde{F}_k. \tag{7.19}$$

From this we see that CSI-POVM tomographic scheme is possible when the a priori information is the diagonal elements of the density matrix, which are clearly experimentally measurable from the projectors $H_k$, and the dimension of the state is $d = p^n + 1$. The complete tomographic formula for an arbitrary state $\rho$ is in this case

$$\rho = \sum_{k=0}^{d-1} P\left(H_k\right) H_k + \sum_{k=d}^{d^2} \left( DP(\tilde{F}_k) - \frac{d-1}{d} \right) \tilde{F}_k, \tag{7.20}$$

were we have replaced Eqs. (6.23) and (7.19) into Eq. (6.15) and the auxiliary variable $\alpha$ takes the value $\alpha = 1/d$ in this case. Eq. (7.20) allow us to write the quantum analogue of Bayes law for the assignment of probabilities, that is

$$\begin{aligned} Q\left(E_m\right) &= \sum_{k=0}^{d-1} P\left(H_k\right) P\left(F_m | H_k\right) \\ &\quad + \sum_{k=d}^{d^2} \left\{ DP(\tilde{F}_k) - \frac{d-1}{d} \right\} P(E_m | \tilde{F}_k). \end{aligned} \tag{7.21}$$

We remark that the optimality of this choice of conditional parameters is symmetric, since as we show in Chapter 6 is also proved that the CSI-POVM when all non-diagonal elements of the density matrix are known, is the set of diagonal projectors $H_k$.

Our Eq. (7.21) is of great relevance since it shows how the procedure introduced in this article allows to get new results and insights into the new rescalings of the total probability law proposed for probability assignments in quantum mechanics. For example, in Eq. (7.21) we have separated the probability law into two parts. The first part related to the projectors $H_k$ is exactly the usual probability law used in the classical case, which is consistent with the fact that $H_k$ form a basis for the diagonal states which are those with the most classical behavior. The second term in Eq. (7.21) includes the contributions obtained from the measurement of the members $\tilde{F}_k$ of the CSI-POVM that in this case represent the contributions from the non-diagonal terms of $\rho$. We can interpret Eq. (7.21), which provides the quantum probability assignment $Q\left(E_m\right)$, as a combination of a *classical* term, represented by the usual law of probabilities, and a *quantum* term, represented by a rescaling of the total law of probabilities for the interference terms.

Our formula (7.18) is very general and open new possibilities for the Quantum bayesianism program. This equation proves that the existence of SIC-POVM is a sufficient but not necessary condition for completing the program and indicates a procedure to find an alternative quantum bayesian rule. The tomographic scheme due to this formula can be performed experimentally with the help of programable spatial light modulators (SLM) based on liquid crystal displays [91]. These are capable of amplitude and phase-modulation of light beams and can be easily controlled by software. A particular application of these devices has been the tomographic reconstruction of a single spatial qudit in dimensions 7 and 8 [92, 93].Due to the high degree of control of the SLM it is possible to generate a large class of initial states, pure or mixed, and to project them onto states whose projectors are the normalization of the CSI-POVM like the $U^k|\phi\rangle$ of Chapter 6. Thereby, an experimental implementation of a tomographic reconstruction based on Eq. (7.18) is well within reach of current experimental techniques.

The error propagation of this scheme is linear and thus it gives better quality of results than linear inversion, also in Chapter 6 we proved that maximizes the likelihood function for a linear estimator, thus on any sense the CSI-POVM tomography here proposed improves the data analysis of experimental results.

Since we can directly rearrange the terms of Eq. (5.21) to take the same form as Eq. (7.18), is that our formula suggest me that the addition of QBism to classical Bayesianism lies in the rescaling constant $\alpha = (1 - \mu d)$ and the minus sign in the second term, rather than the particular value of dimension $d$ as expected by C. Fuchs [80]. This encourages me to lead my research towards a critical review of Bayesian desiderata (Chapter 2), from which I hope to reveal the QBism rules true origin. In analogy to the work of E. T. Jaynes for classical probabilities, this research program should lead to suitable desiderata and clear principles which result in a probability theory with the same statistics as quantum mechanics but without the Hilbert space formalism.

## 7.4 Final remarks

While still writing this thesis many advances on our research and quantum bayesianism have come to light.

There is actually a very advanced draft about the CSI-POVM rule for quantum assignment and a preliminary draft for the MUEBs introduced in Chapter 6. New ideas and research projects based on the tomography scheme proposed in Chapter 4 are under discussion.

Also through personal communication M. Appleby, C. Fuchs and H. Zhu have introduced me to their actual research on the restrictions induced by the QBism rule in the probability space [89, 90] They proposed a new approach to that problem by focusing in the symmetries ( Like Klein in his Erlangen program [94] about geometry) of the allowed probabilities in the probability space.

Also G. Gour has show recently that SIC-POVMs of arbitrary rank (only rank one SIC-POVMs are optimal for reconstruction) exist in all dimensions [95]. This general

SIC-POVM lose some of the properties of rank one SIC-POVMs, but still allow for a QBism rule (dependent of the rank of the SIC-POVM) that also has the same structure of (7.18). This generalization and its consequences over the restrictions of probability space still require more research.

# Bibliography

[1] M. Keyl , *Fundamentals of quantum information theory*, *Physics Reports* **369**, 431 to 548 (2002). 1, 1.1, 1.2.1, 1.2.3, 1.4

[2] Xavier Zubiri, *Naturaleza, Historia, Dios*, *Alianza Editorial* (1999) (primera edicion, Madrid 1944). English translation: *University press of America*, (1981). 1.1, 2.1

[3] Desiderio Papp, *Historia de la física : desde Galileo hasta los umbrales del siglo XX*, Buenos Aires, *Editorial: Espasa Calpe* (1945).

[4] I. Kant, *Kritik der reinen Vernunft*, Riga 1787, *Critica de la razon pura* edicion en castellano de Pedro Ribas *Santillana ediciones*, (2006). 1.1

[5] A. Peres, *Quantum Theory: Concepts and Methods*, *Kluwer, Dordrecht* (1993). 1.1, 2.1

[6] E. H. Hutten, *The Ideas of Physics*, *Oliver and Boyd* , (1967). 1.1

[7] E. T. Jaynes, *Probability Theory, The Logic of Science*, *Cambridge University Press* , (2003). 1.1

[8] P. Gregory, *Bayesian Logical Data Analysis For The Physical Sciences - A Comparative Approach With Mathematica* , *Cambridge Press*, (2010). 1.1, 2.2, 7.1.1

[9] W. F. Stinespring, *Positive functions on C*-algebras*, *Proc. Anner. Math. Soc.* pages 211-216, (1955). 1.1, 2.2, 2.3, 1, 2, 3, 7.1.1

[10] Vergilious Ferm, *History of Philosophical Systems*, *The Philosophical library, New York*, (1950). 1.3.2

[11] I. M. Copi, *Introduction to Logic*, *Macmillan Press* , (1953). 2.1

2.1

[12] Jorge Eduardo Ribera and Maria Teresa Stuven, *Comentario a Ser y Tiempo de Martin Heidegger*, *Ediciones UC*, (2008). 2.1

[13] Aristotle, *Organon*, 4th Century B.C. 2.2

[14] Henri Bergson, *Introduccion a la Metafisica*, *UNAM, Mexico*, (1960). 2.1

[15] S. F. Barker, *Induccion e Hipotesis*, *Editorial Universitaria de Buenos Aires*, (1963). 2.2

[16] D. Hume, *Tratado de la naturaleza humana*, *Editorial Tecnos. Madrid* , (2008). 2.2

[17] C. G. Hempel , *Filosofia de la ciencia natural*, *Editorial Alianza. Madrid* , (2006). 2.2

[18] K. R. Popper, *La logica de la investigacion cientifica*, *Editorial Tecnos. Madrid* , (1994). 2.2

[19] T. Bayes, *An essay toward solving a problem in the doctrine of chances*, *Philosophical Transactions of the Royal Society* pp. 370 to 418, (1763). 2.2

[20] J. Bernoulli, *Ars conjectandi*, *Thurnisiorum* (1713). Reprinted in Die Werke von Jakob Bernoulli *Birkhaeuser* **3** pp. 107 to 286, (1975). 2.2

[21] Laplace, *Memoire sur la probabilite des causes par les evenemens*, *Memoires de l'Academie royale del sciences* **6** 621-656, (1774). Reprinted in Laplace (1878 to 1912), vol. 8, pp. 27 65, Gauthier Villars, Paris. English translation by S. M. Stigler (1986). 2.2

[22] G. Boole, *An Investigation of the laws of Thought*, (1854) *Dover Publications* , (1958). 2.2

[23] G. Polya, *How to Solve It*, *Princeton University Press.* , (1945). 2.2

[24] G. Polya, *Mathematics and Plausible Reasoning*, *Princeton University Press.* 2 Vols., (1954). 2.2

[25] R.T. Cox, *Probability, Frequency, and Reasonable Expectation*, *Am. Jour. Phys.* **14** 1-13, (1946). 2.2

[26] S.M. Barnett and S. Croke, *Quantum State Discrimination*, *Advances in Optics and Photonics* **1** 238 to 278, (2009). 1

[27] J. A. Bergou, *Quantum state discrimination and selected applications*, *Journal of Physics: Conference Series* **84** 012001, (2007). 3

[28] A. S. Holevo, *Statistical decision theory for quantum systems*, *Journal of Multivariate Analysis* **3** 337 to 394, (1973). 3.1.1

[29] H. P. Yuen and R.S. Kennedy and M. Lax, *Optimum testing of multiple hypothesis in quantum detection theory*, *Transactions on information theory IT21* 125134, (1975). 3.1.1

[30] C. W. Helstrom, *Quantum detection and estimation theory*, *Academic Press* , (1976). 3.1, 3.1.1

[31] Y. C. Eldar and A. Mergretski and G. C. Verghese, *Designing optimal quantum detectors via semidefinite programming*, *Transactions on information theory* **49** 1007 to 1012, (2003). 3.1.1

[32] I. D. Ivanovic, *How to differentiate between non orthogonal states*, *Physics Letters A* **123** 257 to 259, (1987). 3.2, 4.3

[33] D. Dieks, *Overlap and distinguishability of quantum states*, *Physics Letters A* **126** 303 to 306, (1988). 3.2, 4.3

[34] A. Peres, *How to differentiate between non orthogonal states*, *Physics Letters A* **128** 19 to 19, (1988). 3.2, 4.3

[35] G. Jaeger and A. Shimony, *Optimal distinction between two non orthogonal quantum states*, *Physics Letters A* **197** 83 to 87, (1995). 3.2

[36] A. Chefles, *Unambigous discrimination between linearly independent quantum states*, *Physics Letters A* **239** 339 to 347, (1998). 3.3

[37] A. Peres and D. R. Terno, *Optimal distinction between two non orthogonal quantum states*, *Journal of Physics A* **31** 7105 to 7111, (1998). 3.3

[38] L. Roa and R. Salazar and C. Hermann-Avigliano and A. B. Klimov, *Conclusive Discrimination among N equidistant pure states*, *Physical Review A* **84** 014302, (2011). 3.3, 3.3, 3.3, 6.4.2

[39] O. Jimenez and L. Roa and A. Delgado, *Probabilistic cloning of equidistant states*, *Physical Review A* **82** 022328, (2010). 3.3, 3.3, 6.4.2

[40] C. Paiva and E. Burgos and O. Jimenez and A. Delgado, *Quantum tomography via equidistant states*, *Physical Review A* **82** 032115, (2010). 3.3

[41] J. B. Altepeter and D. F. V. James and P. G. Kwiat, *Qubit Quantum State Tomography*, *published in the book Quantum State Estimation, Springer* , (2004). 4.2, 4.2

[42] Y. S. Teo and H. Zhu and B. G. Englert and J. Rehacek and Z. Hradil, *Quantum State Reconstruction by Maximizing Likelihood and Entropy*, *PhysRevLett* **107** 020404, (2011). 4.2, 4.2

[43] V. Buzek, *Quantum Tomography from Incomplete Data via MaxEnt Principle*, *published in the book Quantum State Estimation, Springer* , (2004). 4.2

[44] R. Salazar and A. Delgado, *Quantum tomography via unambiguous state discrimination*, *Physical Review A* **86** 012118, (2012). 4.3

[45] B. Huttner and A. Muller and J. D. Gautier, H. Zbinden and N. Gisin, *Unambiguous quantum measurement of nonorthogonal states*, *Physical Review A* **54** 3783 to 3789, (1996). 4.5

[46] R. B. M. Clarke and A. Chefles and S. M. Barnett and Erling Riis, *Experimental demonstration of optimal unambiguous state discrimination*, *Physical Review A* **63** 040305, (2001). 4.5

[47] O. Christensen , *An Introduction to Frames and Riesz Bases*, *Birkhauser* , (2003). 5.1

[48] I. Daubechies , *Ten Lectures on Wavelets*, *SIAM* , (1992). 5.1

[49] P. G. Casazza , *The art of frame theory*, *SIAM* , (1992). 5.1

[50] I. Daubechies and A. Grossmann and Y. Meyer , *Painless non orthogonal expansions*, *J. Math. Phys.* **27** 1271, (1986). 5.1

[51] J. M. Renes and R. Blume Kohout and A. J. Scott and C. M. Caves , *Symmetric informationally complete quantum measurements*, *J. Math. Phys.* **45** 2171, (2004). 5.3, 5.3

[52] I. D. Ivanovic , *Geometrical description of quantal state determination*, *J. Phys. A* **14**, (1981). 5.3, 5.3

[53] W. K. Wootters and B. D. Fields, *Optimal state determination by mutually unbiased measurements*, *Ann. Phys.* **191** 363, (1989). 5.3, 5.3, 6.1

[54] G. Zauner , *Quantendesigns - Grundzuge einer nichtkommutativen Designtheorie*, *PhD thesis, University of Vienna* , (1999). 5.3, 5.5.1

[55] S. G. Hoggar , *64 lines from a quaternionic polytope*, *Geom. Dedic.* **69** 287, (1998). 5.3

[56] M. Grassl , *On SIC POVMs and MUBs in dimension 6*, *Proceedings of the ERATO Conference on Quantum Information Science, Tokyo* p. 60, (2004). 5.3

[57] D. M. Appleby , *Symmetric informationally complete positive operator valued measures and the extended Clifford group*, *J. Math. Phys.* **46** 052107, (2005). 5.3

[58] M. Grassl , *Tomography of quantum states in small dimensions*, *Electron. Notes Discrete Math.* **20** 151, (2005). 5.3

[59] C. M. Caves, *Quantum error correction and reversible operations*, *J. Supercond.* **12** 707, (1999). 5.1

[60] P. Rungta and W. J. Munro and K. Nemoto and P. Deuar and G. J. Milburn C. M. Caves , *Qudit entanglement*, *In Directions in Quantum Optics: A Collection of Papers Dedicated to the Memory of Dan Walls (Springer Verlag, Berlin)* p. 149, (2000).

[61] P. Rungta and V. Buzek and C. M. Caves and M. Hillery and G. J. Milburn , *Universal state inversion and concurrence in arbitrary dimensions*, *Phys. Rev. A* **64** 042315, (2001).

[62] P. G. Casazza and M. Fickus and J. Kovacdevic and M. T. Leon and J. C. Tremain, *A physical interpretation for finite tight frames*, *Birkhauser, Boston* , (2006). 5.1, 5.1

[63] S. D. Li , *On general frame decompositions*, *Numer. Func. Anal. Optimiz.* **16** 1181, (1995). 5.1

[64] R. T. Horn and A. J. Scott and J. Walgate and R. Cleve and A. I. Lvovsky and B. C. Sanders , *Classical and quantum fingerprinting with shared randomness and one sided error*, *Quantum Inf. Comput.* **5** 258, (2005). 5.1

[65] A. J. Scott, *Tight informationally complete quantum measurements*, *J. Phys. A: Math. Gen.* **39** 13507, (2006). 1, 5.1

[66] R. Salazar and D. Goyeneche and A. Delgadoa and C. Saavedra, *Constructing symmetric informationally complete positive-operator-valued measures in Bloch space*, *Physics Letters A* **376** 325 to 329, (2012). 5.4, 5.4

[67] A. J. Macfarlane A. Sudbery and P. H. Weisz, *On Gell Mann Matrices d and f Tensors, Octects and Parametrizations of SU(3)*, *Commun. math. Phys.* **11**, 77 to 90 (1968). 5.4

[68] D. Goyeneche and R. Salazar and A. Delgado, *Characterization of fiducial states in prime dimensions via mutually unbiased bases*, *Phys. Scr.* 014031, (2013). 5.5

[69] H. Georgi, *Lie Algebras in Particle Physics*, *Westview Press* , (1999). 5.4

[70] A. J. Scott and M. Grassl, *Symmetric informationally complete positive operator valued measures: A new computer study*, *J. Math. Phys.* **51** 042203, (2010). 5.3, 5.5.3

[71] S. Bandyopadhyay and P. Boykin and V. Roychowdhury and F. Vatan, *A new proof for the existence of mutually unbiased bases*, *Algoritmica* **34** 512, (2002). 5.5.1

[72] D. M. Appleby , *SIC POVMS and MUBS: Geometrical Relationships in Prime Dimension*, *AIP Conference Proceedings* **1101** Issue 1, p223, (2009). 5.5.2

[73] K. S. Gibbons and M. J. Hoffman and W. K. Wootters, *Discrete phase space based on finite fields*, *Phys. Rev. A* **70** 062101, (2004). 5.5.2

[74] U. Larsen , *Superspace geometry: the exact uncertainty relationship between complementary aspects*, *J. Phys. A: Math. Gen.* **23** 1041, (1990). 5.5.3

[75] G. M. D Ariano and P. Perinotti and M. F. Sacchi, *Informationally complete measurements and groups representation*, *J. Opt. B* **6** S487, (2004). 5.1

[76] D. Petz and L. Ruppert and A. Szanto, *Conditional SIC POVMs*, *to be published, http://arxiv.org/abs/1202.5741* , (2012). 6.3, 6.4.1

[77] D. Petz and L. Ruppert , *Efficient quantum tomography needs complementary and symmetric measurements*, *Reports on Mathematical Physics* **69** 161 to 177, (2012). 6.4.1

[78] D. Kalra, *Complex equiangular cyclic frames and erasures*, *Linear Algebra and its Applications* **419** 373 to 399, (2006). 6.3

[79] J. Singer, *A theorem in finite projective geometry and some applications to number theory*, *Trans. Amer. Math. Soc.* **43** 377 to 385, (1938). 6.3

[80] C. A. Fuchs , *QBism, the Perimeter of Quantum Bayesianism*, *arXiv:1003.5209v1* , (2010). 7.1, 7.3

[81] C. M. Caves, *Symmetric Informationally Complete POVMs*, *posted at http://info.phys.unm.edu/caves/reports/infopovm.pdf* , (1999). 7.2

[82] R. E. Kass and L. A. Wasserman, *The selection of prior distributions by formal rules*, *Statistical Association* **90** 928 to 934, (1996). 7.1.1, 7.2

[83] M. Christandl and R. Renner, *Reliable Quantum State Tomography*, *Phys. Rev. Lett.* **109** 120403, (2012). 7.2

[84] R. Blume Kohout, *Optimal reliable estimation of quantum states*, *New Journal of Physics* **12** 043034, (2010). 7.2

[85] C. M. Caves and C. A. Fuchs and R. Schack, *Unknown quantum states: The quantum de Finetti representation*, *Jour. of Math Physics* **43** Num 9, (2002). 7.1.1, 7.1.2

[86] P. Monari and D. Cocchi, *Induction and Probability*, *Biblioteca di Statistica, CLUEB, Bologna* , (1993). 7.1.1, 7.1.1, 7.1.1

[87] D. Heath and W. Sudderthi , *De Finetti theorem on exchangeable variables*, *American Statistician.* **30** (4) 188, (1976). 7.1.1

[88] C. M. Caves and C.A. Fuchs and R. Schack, *Quantum Probabilities as Bayesian Probabilities*, *Phys. Rev. A* **65** 022305, (2002). 7.1

[89] D. M. Appleby and A. Ericsson and C. A. Fuchs, *Properties of QBist State Spaces*, *Foundations of Physics* **41** Issue 3, pp 564 to 579, (2011). 7.4

[90] C. A. Fuchs , *Charting the Shape of Quantum State Space*, *AIP Conf. Proc.* **1363** 305, (2011). 7.4

[91] G. Lima, A. Vargas, L. Neves, R. Guzmán, and C. Saavedra, Opt. Express **17**, 10688 (2009). 7.3

[92] G. Lima, L. Neves, R. Guzmán, E. S. Gómez, W. A. T. Nogueira, A. Delgado, A. Vargas, and C. Saavedra, Opt. Exp. **19**, 3542 (2011). 7.3

[93] G. Lima , F. A. Torres-Ruiz , L. Neves , A. Delgado , C. Saavedra and S. Pádua, J. Phys. B: At. Mol. Opt. Phys. **41**,185501 (2008). 7.3

[94] F. Klein , *A comparative review of recent researches in geometry*, Complete English Translation is here http://arxiv.org/abs/0807.3161, (1872). 7.4

[95] G. Gour, *General symmetric informationally complete measurements exist in all dimensions.*, *arXiv:1305.6545* [qu ant-ph], (2013). 7.4

# Agradecimientos

Antes que nada quiero agradecer a mis padres y familia. Si no fuera por su dedicación, paciencia y cariño incondicional no sería la persona que soy hoy en día. La formación integral y amor al conocimiento que me lega mi padre, así como la ayuda práctica y apoyo constante de mi madre son regalos y muestras de amor invaluables con los que viviré siempre.

Agradezco también a mi tutor Aldo Delgado, por su eterna comprensión y empatía. Esta tesis y sus resultados no serian posibles sin su colaboración y la sabiduría con la que supo darme la libertad que necesitaba para abordar los problemas a los que mí espíritu se inclinaba.

Así mismo a todos los profesores que han sido fuente de inspiración y apoyo en este camino de la física, entre ellos C. Saavedra, G. Rubilar, P. Salgado, J. Araneda, J. Días de Valdés, F. Borotto, J. Aguirre , P. Manidurai y L. Braga.

En forma especial agradezco en este sentido al Michael Kurgansky con quién nuestras conversaciones sobre la vida y la vida en la ciencia han significado para mí un hito en mi formación, que espero poder legar a las siguientes generaciones.

También quiero aquí recordar a mis compañeros de generación con quienes compartí las batallas que deben ser libradas en este estudiar y búsqueda de la verdad: Tomás Ramos, Patricio Muñoz, Esteban Sepúlveda, Miguel Solís, Mauricio Santibañes, Alejandra Muñoz , Carla Hermann y muchos otros.

Gracias también a toda la gente del CEFOP, programa Milenium y beca CONICYT que con su apoyo económico me abrieron muchas puertas y pude salir a expandir mis límites así como tener una seguridad y comodidad necesarias para la investigación.

Pero esto ha durado ya poco más de ocho años y en el camino nuevas generaciones se me han presentado y con las cuales, en nuestra interacción me he sentido revitalizado para seguir adelante en esta búsqueda de la verdad, así es el caso de: Pablo Solano, Cristian Jara, Gabriel Araneda, Nataly Cisternas y más recientemente Jorge Johnny Moya, Felipe Matus, Karina Ferrada, Nicolás Pérez y Aníbal Utreras

Hay tantos y tantas de esta facultad a quienes debería agradecer por las conversaciones y discusiones que no pueden ser separadas de quién soy ahora. Por eso pido disculpas a aquellos quienes mereciendo ser mencionados, mi memoria no les haga tal justicia.

Sin embargo, ¡por injusta que sea mi memoria! Demasiado sería el pecado si dejo de lado a las solidas amistades que labre escribiendo para la revista Anifreak. Ellos han sido ese soporte tan grande y tan escaso que llamamos amistad. Con ellos no puedo dejar de sentirme en deuda por ofrecérmela con entrega desinteresada y sincera. Me refiero a Guillermo Pereira, Manuel Durán, Pablo Pizarro, Mauricio Peñailillo, Nicolás-Martinez Conde, Chingt-ta Lai, Cristina Silva y Camila Navarrete.

Si Anifreak fue una veta mineral de amistad, en mi camino también he encontrado otras piedras preciosas no de menor calidad e importancia, como: Daniela Lagos, Juan Clavería y Alex Brumm.

Y así con esta memoria pecadora, del noveno círculo de Dante, no podría escaparme si no agradezco los afectos y cariños de Daniela Grandón, que fueron mí gran sustento en esta tesis.

A todos ¡muchas gracias!

Realizado por TBOTEX