UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

Adaptive Quantum State Tomography with Fisher Symmetric Measurements.

Tesis presentada a la Facultad de Ciencias Físicas y Matemáticas
de la Universidad de Concepción para optar al grado de
Magíster en Ciencias con Mención en Física
por:

Constanza Patricia Vargas Rosales

Marzo de 2025
Concepción, Chile

Profesor Guía: Dr. Aldo Delgado Hidalgo

A mis padres

Agradecimientos

Primero, me gustaría expresar mi más profunda gratitud a mis padres por apoyar siempre mi decisión de seguir una carrera científica y por ser mi mayor respaldo. A mi padre, por inspirarme la alegría de la lectura y el conocimiento, y a mi madre, por su permanente cuidado y motivación. A mi hermano, por ser una presencia constante en mi vida. Y a toda mi familia, por su apoyo incondicional y por sentirse orgullosos de mi trayectoria.

También estoy profundamente agradecido con las increíbles personas que he conocido a lo largo de estos años en la universidad, quienes han hecho de este viaje una experiencia enriquecedora y placentera. A mis amigos Constanza, Adheris, Lilianne, Tomás, Luis y Alejandro: gracias por mantenerse unidos en todo momento, por las risas y la camaradería en el camino. A mis compañeros de oficina, Jorge y Katherine, por hacer que los largos días de trabajo fueran más amenos. A todos mis profesores, por sentar las bases de mi comprensión de la física. Y un agradecimiento especial a Francisco, por estar siempre ahí para mí, por creer en mí, apoyarme y por las interminables discusiones sobre física que han moldeado mi perspectiva.

También quisiera expresar mi más sincera gratitud al Dr. Aldo Delgado y al Dr. Luciano Pereira por su invaluable orientación durante la investigación y redacción de esta tesis. Mi agradecimiento también se extiende al grupo Concepción MIRO por las enriquecedoras discusiones que han contribuido significativamente a mi trabajo. Sin su apoyo y mentoría, este proyecto habría sido considerablemente más desafiante. Finalmente, agradezco al Instituto Milenio de Investigación en Óptica (MIRO) por su apoyo financiero.

Resumen

La estimación de estados cuánticos desempeña un papel crucial en el procesamiento de información cuántica, incluyendo la comunicación, la computación y la metrología cuánticas, donde la caracterización precisa del estado es esencial para su evaluación y optimización. En la presente Tesis proponemos un método adaptativo de tres etapas para estimar estados cuánticos puros de dimensión d utilizando mediciones simétricas de Fisher (FSM) y una de medición de un solo shot en una base. El resultado de la primera medición se emplea para construir dos FSM que conjuntamente estiman el estado desconocido hasta un conjunto de medida nula. Esta estimación luego se utiliza para adaptar una tercera FSM, que refina la reconstrucción final del estado.

Nuestro enfoque logra una infidelidad de estimación promedio que se acerca al límite inferior de Gill-Massar (GMB) sin necesidad de información previa más allá de la pureza del estado, lo que extiende la aplicabilidad de las FSM a cualquier estado puro desconocido. Además, el número total de resultados de mediciones escala linealmente como $7d-3$, lo que reduce significativamente la complejidad en comparación con los métodos que dependen de mediciones colectivas en múltiples copias. Este trabajo resalta el potencial de las técnicas de estimación adaptativa para una caracterización eficiente y precisa de estados cuánticos.

Keywords – Tomografía de Estados Cuánticos, Información Cuántica, Medidas simétricas de Fisher

Abstract

Quantum state estimation plays a crucial role in quantum information processing, including quantum communication, computation, and metrology, where accurate state characterization is essential for evaluation and optimization. We propose a three-stage adaptive method for estimating arbitrary d-dimensional pure quantum states using Fisher-symmetric measurements (FSM) and a single-shot measurement on a basis. The initial measurement outcome is used to construct two FSMs that jointly estimate the unknown state up to a null measure set. This estimate then guides the adaptation of a third FSM, which refines the final state reconstruction.

Our approach achieves an average estimation infidelity that closely approaches the Gill-Massar lower bound (GMB) without requiring prior information beyond the state's purity, thereby extending the applicability of FSM to any unknown pure state. Moreover, the total number of measurement outcomes scales linearly as $7d - 3$, significantly reducing the complexity compared to methods relying on collective measurements across multiple copies. This work highlights the potential of adaptive estimation techniques for efficient and precise quantum state characterization.

Keywords – Quantum State Estimation, Quantum Information, Fisher Symmetric Measurement

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Quantum state tomography (QST) is a fundamental aspect of quantum information science, enabling the reconstruction of a quantum system's state based on measurement outcomes. It is essential for the validation and verification of quantum technologies, such as quantum computing, quantum communication, and quantum metrology. As we approach the realization of scalable quantum technologies, the importance of accurately measuring and characterizing quantum systems cannot be overstated. Quantum computing, for instance, has the potential to revolutionize industries such as cryptography, optimization, and materials science by solving problems that are intractable for classical computers. Quantum communication, with its promise of unbreakable encryption via quantum key distribution, is poised to redefine secure communication. Additionally, quantum metrology leverages quantum principles to achieve measurement precision beyond classical limits, impacting fields ranging from timekeeping and navigation to sensing gravitational waves and detecting dark matter.

The successful deployment of these technologies relies heavily on our ability to perform precise, reliable, and efficient quantum state tomography. However, the high dimensionality and complexity of quantum systems, particularly in the presence of noise and imperfections, present significant challenges for conventional QST methods. Traditional techniques, such as linear inversion and maximum likelihood estimation, while robust, often struggle with computational inefficiencies and are sensitive to the presence of noise, which can lead to inaccurate state reconstructions. These challenges are especially prominent in systems with large

numbers of qubits or high-dimensional quantum systems, where the sheer amount of data required for tomography becomes computationally prohibitive.

Thus, there is a pressing need for novel approaches that can improve the efficiency, accuracy, and scalability of QST. One promising avenue is the use of Fisher Symmetric Measurements (FSM), which optimize the information extracted from each measurement by minimizing statistical uncertainty through the Fisher Information Matrix. FSMs have been shown to offer significant improvements in state reconstruction, but their reliance on prior knowledge of the system's state limits their applicability in situations where the state is completely unknown or poorly understood.

This thesis proposes a novel framework that enhances the applicability of FSM in quantum tomography by integrating adaptive measurement schemes. These adaptive methods allow the quantum state to be inferred progressively, adjusting measurement strategies based on the data collected. We present an adaptive method that uses three FSMs and a single-shot measurement to estimate any pure state in arbitrary dimension near-optimally. Notably, this method eliminates the requirement for prior information or collective measurements. This method also shows that any pure state can be estimated near-optimally using only seven measurement bases, two more than the minimum required. This work opens new avenues for FSMs and addresses a fundamental pure state estimation problem. By extending FSM to work effectively with completely unknown quantum states, this work addresses a major gap in the current state-of-the-art in QST.

Through both theoretical developments and numerical simulations, this thesis demonstrates the feasibility and effectiveness of the proposed adaptive FSM approach, paving the way for more accurate, efficient, and scalable quantum state tomography. The proposed framework not only advances the field of quantum tomography but also provides critical tools for the future of quantum technologies, especially in the context of high-dimensional quantum systems where traditional methods are increasingly inadequate.

Chapter 2 introduces the mathematical framework necessary for this work, covering fundamental concepts such as vector spaces, Dirac's ket notation commonly used in Quantum Mechanics. And essential notions of probability and statistics crucial for understanding estimation theory and Fisher Information, which play a key

role throughout this thesis. Additionally, the Maximum Likelihood Estimator (MLE), one of the most widely used estimation methods, is presented.

Chapter 3 outlines the postulates of Quantum Mechanics, along with key definitions such as quantum states, distance measures between quantum states, and quantum measurements. These concepts are fundamental for introducing the problem of quantum state estimation and quantum tomography, which is the focus of Chapter 5. Moreover, this chapter provides the groundwork for generalizing Fisher Information and extending the Cramér-Rao bound to quantum experiments in Chapter 4.

Chapter 6 explores the construction of measurement strategies that achieve the Gill-Massar bound, focusing on Fisher-symmetric measurements. Chapter 7 presents the key results of this thesis, presenting the adaptive protocol that utilizes Fisher-symmetric measurements, along with numerical simulations and respective plots that demonstrate its performance. Finally, Chapter 8 discusses potential future research directions and open problems related to this work.

1.1 Contribution and Activities

The work carried out in this thesis resulted in the following paper, with a preprint available on ArXiv:

- *Near-optimal pure state estimation with adaptive Fisher-symmetric measurements*, C. Vargas, L. Pereira, A. Delgado. https://doi.org/10.48550/arXiv.2412.04555 [1].

The Python code for the simulations, along with the generated data, is publicly accessible in the following GitHub repository:

- https://github.com/CoVargasRo/FSM

Additionally, I had the opportunity to present this work during poster sessions at various workshops:

- *Workshop MIRO*, Faculty of Physical and Mathematical Sciences, Universidad de Chile, November 11-15, 2024.
- *XXIV Chilean Physics Symposium 2024*, Universidad de La Frontera, Temuco, November 20-22, 2024.

- *Quantum Optics X*, Puerto Varas, December 9-13, 2024.

Throughout my Master's studies, I actively participated in various science outreach initiatives:

- Led the workshop *"Introduction to Qiskit"* at the *Second MIRO Quantum Computing School*, Pontificia Universidad Católica de Chile, Santiago (January 8-12, 2024).
- Participated in *"The Quantum Box: A Journey Through Light"* Project at Teatro BioBio, MIRO (March - July 2024):
  - Promoted the project at the Citizen Dome during the REC Festival, Concepción (March 23-24, 2024).
  - Supervised the exhibition and conducted guided visits for schools and the general public, explaining the physics behind each module.
- Explained light-related physics phenomena to prospective science students at the MIRO stand during the *Universidad de Concepción Open House Fair* (October 9-10, 2024).
- Taught quantum computing to high school students at the *"Quantum Builder" Workshop*, MIRO (October 16-17, 2024).
- Led the workshop *"Introduction to Qiskit 2"* at the *Third MIRO Quantum Computing School*, Universidad de Chile, Santiago (January 6-10, 2025).

Chapter 2

Mathematical Framework

Linear algebra studies vector spaces and linear transformations, forming the mathematical foundation of Quantum Mechanics. Quantum states are modeled as vectors in Hilbert spaces , with linear operators O representing physical observables. The standard quantum mechanical notation for a vector in a vector space is the Dirac notation $|\psi\rangle$. The vector space of most interest to us is $\mathbb{C}_n$.

Since Quantum Mechanics is inherently probabilistic we also need to review some important concepts about statistics such as sample space and random variables. When performing an experiment in which a quantum system is measured, the possible values are random and are taken from a probability distribution given by the quantum state. Having this, we can perform statistical inference to extract information about the system.

An important quantity in the area of estimators is how much information you can extract from the system when performing an experiment, this is given by the Fisher Information. This is an important concept in the length of this research.

This chapter reviews key principles and notation essential for understanding Quantum Mechanics and quantum state estimation.

2.1 Linear Algebra

The natural language of Quantum Mechanics is linear algebra [2]. In this section, we review Hilbert spaces, vectors, and linear operators. Most of this information is taken from [3].

2.1.1 Hilbert Space

The space $\mathcal{H}$ is a linear vector space on the body of complex numbers $\mathbb{C}$ with operations addition and multiplication by a scalar, $(\mathcal{H}, +, \cdot)$

$$+ : \mathcal{H} \times \mathcal{H} \to \mathcal{H}, \tag{2.1.1}$$

$$\cdot : \mathbb{C} \times \mathcal{H} \to \mathcal{H}. \tag{2.1.2}$$

The elements of this space $\mathcal{H}$ are called *vectors*. Having $|\alpha\rangle, |\beta\rangle, |\gamma\rangle \in \mathcal{H}$ and $a, b \in \mathbb{C}$, the sum operation satisfy,

$$\text{Commutation} : |\alpha\rangle + |\beta\rangle = |\beta\rangle + |\alpha\rangle, \tag{2.1.3}$$

$$\text{Associativity} : |\alpha\rangle + (|\beta\rangle + |\gamma\rangle) = (|\alpha\rangle + |\beta\rangle) + |\gamma\rangle, \tag{2.1.4}$$

$$\text{Null vector} : |\alpha\rangle + |0\rangle = |\alpha\rangle, \tag{2.1.5}$$

$$\text{Inverse vector} : |\alpha\rangle + |-\alpha\rangle = |0\rangle, \tag{2.1.6}$$

and the operation of scalar multiplication satisfy,

$$\text{Distribution of vector sum} : a(|\alpha\rangle + |\beta\rangle) = a\,|\alpha\rangle + a\,|\beta\rangle, \tag{2.1.7}$$

$$\text{Distribution of scalar product} : (a + b)\,|\alpha\rangle = a\,|\alpha\rangle + b\,|\alpha\rangle, \tag{2.1.8}$$

$$\text{Associativity} : a(b\,|\alpha\rangle) = (ab)\,|\alpha\rangle, \tag{2.1.9}$$

$$\text{Multiplication by 0 and 1} : 0\,|\alpha\rangle = \emptyset \quad , \quad 1\,|\alpha\rangle = |\alpha\rangle. \tag{2.1.10}$$

The *inner product* is a function which takes as input two vectors $|\alpha\rangle$ and $|\beta\rangle$ from a vector space and produces a complex number as output,

$$(\cdot, \cdot) : \mathcal{H} \times \mathcal{H} \to \mathbb{C}. \tag{2.1.11}$$

The standard quantum mechanical notation for the inner product $(|\alpha\rangle, |\beta\rangle)$ is $\langle\alpha|\beta\rangle$, where $|\alpha\rangle$ and $|\beta\rangle$ are vectors in the inner product space, and the notation $\langle\alpha|$ is used for the dual vector to the vector $|\alpha\rangle$. This operation has the following

properties:

$$\text{Linearity} : (|\alpha\rangle, a\,|\beta\rangle + b\,|\gamma\rangle) = a\,\langle\alpha|\beta\rangle + b\,\langle\alpha|\gamma\rangle\,, \tag{2.1.12}$$

$$\text{Hermitic} : \langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle^*\,, \tag{2.1.13}$$

$$\text{Positivity} : \langle\alpha|\alpha\rangle \geq 0 \quad \text{with equality iff} \quad |\alpha\rangle = 0. \tag{2.1.14}$$

A vector space $\mathcal{H}$ with such an inner product is then called a *Hilbert space*.

Vectors $|\alpha\rangle, |\beta\rangle$ are *orthogonal* if their inner product is zero, $\langle\alpha|\beta\rangle = 0$. We define the *norm* of a vector $|\alpha\rangle$ by

$$||\,|\alpha\rangle\,|| = \sqrt{\langle\alpha|\alpha\rangle}. \tag{2.1.15}$$

A *unit vector* or normalized vector is a vector $|\alpha\rangle$ such that $||\,|\alpha\rangle\,|| = 1$. We can normalize a vector by dividing it by its norm, $|\alpha\rangle\,/||\,|\alpha\rangle\,||$. The norm of a vector satisfy the Cauchy-Schwarz inequality,

$$|\,\langle\alpha|\beta\rangle\,| \leq ||\,|\alpha\rangle\,||\,||\,|\beta\rangle\,||. \tag{2.1.16}$$

A set of $|\alpha_i\rangle$ vectors with index i is *orthonormal* if each vector is a unit vector and different vectors in the set are orthogonal, that is $\langle\alpha_i|\alpha_j\rangle = \delta_{ij}$.

A spanning set for a vector space is a set of vectors $|v_1\rangle, ..., |v_n\rangle$ such that any vector $|\alpha\rangle$ in the vector space can be written as a linear combination of vectors in that set

$$|\alpha\rangle = \sum_{i=1}^{n} a_i\,|v_i\rangle\,, \tag{2.1.17}$$

where $a_i = \langle v_i|\alpha\rangle$. A set of non-zero vectors $|v_1\rangle, \cdots, |v_n\rangle$ are *linearly dependent* if there exists a set of complex numbers $a_1, ..., a_n$ with $a_i \neq 0$ for at least one value of i, such that

$$a_1\,|v_1\rangle + a_2\,|v_2\rangle + \cdots + a_n\,|v_n\rangle = 0. \tag{2.1.18}$$

A set of vectors is *linearly independent* if it is not linearly dependent. A set of linearly independent vectors that span the space is called a *basis*. The number of vectors on any basis is called the *dimension* of the space. If the elements of the basis are normalized and orthogonal, this basis is called *orthonormal*. We will only be interested in finite-dimensional vector spaces, thus a finite-dimensional Hilbert space.

It is possible to construct an orthonormal basis from a linearly dependent set of vectors using the Gram-Schmidt procedure. Given a set of linearly dependent vectors $\{|v_1\rangle, |v_2\rangle, \ldots, |v_n\rangle\}$ in an inner product space, this method generates an orthonormal basis $\{|u_1\rangle, |u_2\rangle, \ldots, |u_n\rangle\}$ through an iterative process. The procedure is as follows:

1. Define the first orthonormal vector as

$$|u_1\rangle = \frac{|v_1\rangle}{\| |v_1\rangle \|}$$

   if $|v_1\rangle \neq 0$.

2. For each subsequent vector $|v_k\rangle$, remove the components that are parallel to the previously constructed orthonormal vectors:

$$|w_k\rangle = |v_k\rangle - \sum_{j=1}^{k-1} \langle u_j|v_k\rangle \, |u_j\rangle \, .$$

3. Normalize the resulting vector to obtain the next orthonormal basis element:

$$|u_k\rangle = \frac{|w_k\rangle}{\| |w_k\rangle \|}$$

   if $|w_k\rangle \neq 0$.

By applying this procedure iteratively, we obtain a set of mutually orthonormal vectors that span the same space as the original set.

With respect to a prescribed basis any given vector is uniquely represented by the (ordered) n-tuple of its components:

$$|a\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \langle v_1|a\rangle \\ \langle v_2|a\rangle \\ \vdots \\ \langle v_n|a\rangle \end{pmatrix}, \tag{2.1.19}$$

the dual-vector $\langle a|$, known as the bra, is defined as

$$\langle a| = (a_1^*, a_2^*, \cdots, a_n^*) \, . \tag{2.1.20}$$

Furthermore, the inner product in the matrix representation is,

$$\langle a|b\rangle = \left(\sum_{i=0}^{d-1} \alpha_i \left|v_i\right\rangle, \sum_{j=0}^{d-1} \beta_j \left|v_j\right\rangle\right), \tag{2.1.21}$$

$$= \sum_{i=0}^{d-1}\sum_{j=0}^{d-1} \alpha_i^* \beta_j \left\langle v_i|v_j\right\rangle, \tag{2.1.22}$$

$$= \sum_{i=0}^{d-1}\sum_{j=0}^{d-1} \alpha_i^* \beta_j \delta_{i,j}, \tag{2.1.23}$$

$$= \sum_{i=0}^{d-1} \alpha_i^* \beta_i, \tag{2.1.24}$$

$$= (\alpha_1^*, \alpha_2^*, \ldots, \alpha_n^*) \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} \tag{2.1.25}$$

2.1.2 Linear Operators

A linear operator between vector spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ is defined to be any function $A : \mathcal{H}_1 \to \mathcal{H}_2$ which is linear in its inputs,

$$A(a\left|\alpha\right\rangle + b\left|\gamma\right\rangle) = aA\left|\alpha\right\rangle + bA\left|\beta\right\rangle, \tag{2.1.26}$$

with $\left|\alpha\right\rangle, \left|\beta\right\rangle \in \mathcal{H}_1$ and $A\left|\alpha\right\rangle, A\left|\beta\right\rangle \in \mathcal{H}_2$. When $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$ the set of operators is called $\mathcal{L}(\mathcal{H})$. We will be working with this kind of operators.

The product of two linear operators A and B in $\mathcal{L}(\mathcal{H})$, AB, is defined as the following

$$(AB)\left|\alpha\right\rangle = A(B\left|\alpha\right\rangle). \tag{2.1.27}$$

In general, $AB \neq BA$. The *commutator* of A and B is defined as

$$[A, B] = AB - BA. \tag{2.1.28}$$

If A commutates with B then $[A, B] = 0$ and $AB = BA$.

An important linear operator is the *identity operator* $\mathbb{I}_{\mathcal{H}}$, which is defined as the

unique operator that given any linear operator A

$$\mathbb{I}_{\mathcal{H}} A = A \mathbb{I}_{\mathcal{H}} = A, \tag{2.1.29}$$

meaning that it leaves the operator A unchanged and it commutates with A.

We can represent an operator using the *outer product*. The outer product between the vectors $|\alpha\rangle$ and $|\beta\rangle$ is the operator $|\alpha\rangle\langle\beta|$ whose action is defined by

$$(|\alpha\rangle\langle\beta|)|\gamma\rangle = \langle\beta|\gamma\rangle|\alpha\rangle, \quad \forall |\gamma\rangle \in \mathcal{H}. \tag{2.1.30}$$

Let $\{|i\rangle\}_{i=0,\ldots,d-1}$ be an orthonormal basis for the vector space $\mathcal{H}$. So an arbitrary vector $|\psi\rangle$ can be written on this basis as $|\psi\rangle = \sum_i \lambda_i |i\rangle$ with $\lambda_i = \langle i|\psi\rangle \in \mathbb{C}$. Therefore,

$$\left(\sum_i |i\rangle\langle i|\right)|\psi\rangle = \sum_i |i\rangle\langle i|\psi\rangle = \sum_i \lambda_i |i\rangle = |\psi\rangle. \tag{2.1.31}$$

Since the last equation is true for all $|\psi\rangle$, it follows that

$$\sum_{i=0}^{d-1} |i\rangle\langle i| = \mathbb{I}_{\mathcal{H}}. \tag{2.1.32}$$

This equation is known as the *completeness relation*, with $\mathbb{I}_{\mathcal{H}}$ the identity in the $\mathcal{H}$ space. This can be used to represent any operator $A \in \mathcal{L}(\mathcal{H})$ in the outer product notation. Having two basis $\{|i\rangle\}_{i=0,\cdots,d-1}$ and $\{|j\rangle\}_{j=0,\cdots,d-1}$,

$$A = \mathbb{I}_{\mathcal{H}} A \mathbb{I}_{\mathcal{H}}, \tag{2.1.33}$$

$$= \left(\sum_{i=0}^{d-1} |i\rangle\langle i|\right) A \left(\sum_{j=0}^{d-1} |j\rangle\langle j|\right), \tag{2.1.34}$$

$$= \sum_{i,j=0}^{d-1} \langle i|A|j\rangle |i\rangle\langle j|. \tag{2.1.35}$$

An *eigenvector* of a linear operator A on a vector space is a non-zero vector $|\alpha\rangle$ such that

$$A|\alpha\rangle = \alpha|\alpha\rangle, \tag{2.1.36}$$

where α is a complex number known as the *eigenvalue* of A corresponding to the eigenvector $|\alpha\rangle$. We can find the eigenvectors and eigenvalues of any operator via

the *characteristic equation*,

$$det(A - \alpha \mathbb{I}_{\mathcal{H}}) = 0, \tag{2.1.37}$$

where *det* is the determinant function for matrices.

The *adjoint* of a linear operator A is the operator $A^\dagger$ such that

$$(|\alpha\rangle, A|\beta\rangle) = (A^\dagger|\alpha\rangle, |\beta\rangle). \tag{2.1.38}$$

In the matrix representation, this is to take the matrix of A to the conjugate-transpose matrix, $A^\dagger \equiv (A^*)^\top$, where the $*$ indicates complex conjugation and $\top$ indicates the transpose operation.

An operator A such that

$$A^\dagger = A, \tag{2.1.39}$$

is known as an *Hermitian* or *self-adjoint* operator. It follows that, an Hermitian operator A has real eigenvalues α,

$$\alpha = \langle\alpha| A|\alpha\rangle = \langle\alpha| A^\dagger|\alpha\rangle^* = \langle\alpha| A|\alpha\rangle^*. \tag{2.1.40}$$

An important class of Hermitian operators is the *projectors*. Suppose W is a k-dimensional vector subspace of the d-dimensional vector space V, with $|0\rangle, |1\rangle, ..., |k\rangle$ an orthonormal basis for W. By definition,

$$P_k = \sum_{i=0}^{k} |i\rangle\langle i| \tag{2.1.41}$$

is the projector onto the subspace W. All projectors satisfy $P^2 = P$.

An operator A is said to be *normal* if $AA^\dagger = A^\dagger A$. Clearly, an operator which is Hermitian is also normal. The *spectral decomposition* is an extremely useful representation theorem for normal operators.

Theorem 1 (of Spectral Decomposition)**.** *Any normal operator N on a vector space V is diagonal with respect to some orthonormal basis $|i\rangle$ for V,*

$$N = \sum_i \lambda_i |i\rangle\langle i|. \tag{2.1.42}$$

Where $|i\rangle$ are the eigenvectors of N with corresponding eigenvalues λ_i. Conversely, any diagonalizable operator is normal.

We can write this in terms of the projectors $P_i = |i\rangle\langle i|$,

$$N = \sum_i \lambda_i P_i. \tag{2.1.43}$$

This projectors satisfy the completeness relation $\sum_i P_i = \mathbb{I}_V$, and the orthonormality relation $P_i P_j = P_i \delta_{ij}$.

An operator U is *unitary* if

$$U^\dagger U = U U^\dagger = \mathbb{I}, \tag{2.1.44}$$

therefore U is normal and has a spectral decomposition. Unitary operators preserve inner products between vectors,

$$(U|\alpha\rangle, U|\beta\rangle) = \langle\alpha| U^\dagger U |\beta\rangle = \langle\alpha| \mathbb{I} |\beta\rangle = \langle\alpha|\beta\rangle. \tag{2.1.45}$$

Let $|i\rangle$ be any orthonormal basis set. Taking $|j\rangle \equiv U|i\rangle$, we see that $|j\rangle$ is also an orthonormal basis set of the same Hilbert space, since unitary operators preserve inner products.

A *positive semi-definite operator* A is defined to be an operator such that for any vector $|\alpha\rangle \neq 0$,

$$(|\alpha\rangle, A|\alpha\rangle) \geq 0. \tag{2.1.46}$$

It follows that, it has non-negative eigenvalues α. All positive semi-definite operators are Hermitian.

Let A be a normal operator with spectral decomposition,

$$A = \sum_i \alpha_i |i\rangle\langle i|, \tag{2.1.47}$$

given an analytic function $f : \mathbb{C} \to \mathbb{C}$, the operation $f(A)$ is defined as,

$$f(A) = \sum_i f(\alpha_i) |i\rangle\langle i|. \tag{2.1.48}$$

The *trace* of a matrix A is defined to be the sum of its diagonal elements,

$$tr(A) = \sum_{i=0}^{d-1} \langle i| A |i\rangle = \sum_{i=0}^{d-1} A_{ii}. \tag{2.1.49}$$

The trace has the properties of being

$$\text{Cyclic:} \quad tr(AB) = tr(BA) \tag{2.1.50}$$

$$\text{Linear:} \quad tr(\alpha A + \beta B) = \alpha tr(A) + \beta tr(B), \tag{2.1.51}$$

where A and B are arbitrary matrices, and α, β are complex numbers. Following from the spectral decomposition, the trace is equal to the sum of the eigenvalues of the operator,

$$tr(A) = \sum_i \alpha_i. \tag{2.1.52}$$

The *Hilbert-Schmidt inner product* is an inner product used in the context of operators. It is defined as

$$\langle A, B\rangle = tr(A^\dagger B). \tag{2.1.53}$$

The *norm* arising from this inner product is called the *Hilbert-Schmidt norm*,

$$||A|| = \left(\mathrm{tr}\left(A^\dagger A\right)\right)^{1/2}. \tag{2.1.54}$$

This norm is unitarily invariant on matrices, it satisfy the equality

$$||UAV|| = ||A||, \tag{2.1.55}$$

for all unitary U, V. Now, let A and B be any two matrices, then

$$||A^\dagger B||^2 \leq ||AA^\dagger||\ ||BB^\dagger||, \tag{2.1.56}$$

for every unitarily invariant norm. This is called the *Cauchy-Schwarz inequality* for matrices. For the *Hilbert-Schmidt norm*, this inequality is

$$|\,\mathrm{tr}\left(A^\dagger B\right)|^2 \leq \mathrm{tr}\left(AA^\dagger\right)\mathrm{tr}\left(BB^\dagger\right). \tag{2.1.57}$$

The *Kronecker product* provides a systematic way to construct linear operators

that act on the tensor product of two Hilbert spaces. Given two linear operators A and B, acting on Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, their Kronecker product $A \otimes B$ defines a new operator on the product space $\mathcal{H}_1 \otimes \mathcal{H}_2$. This operation ensures that the action of each operator is preserved on its respective space, while combining them into a single operator that acts on the composite space. For example, given matrices

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix},$$

their Kronecker product is

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{bmatrix}.$$

2.2 Probability and Statistics

For this work, we approach probability theory from the *Frenquentist interpretation*. The probability that some specific outcome of a process will be obtained can be interpreted to mean the relative frequency with which that outcome would be obtained if the process were repeated a large number of times under similar conditions. Most of this recompilation is taken from [4] and [5].

2.2.1 Experiments and Probabilities

An **experiment** is any process, real or hypothetical, in which the possible outcomes can be identified ahead of time. An **event** is a well-defined set of possible outcomes of the experiment. The collection of all possible outcomes of an experiment is called the **sample space** of the experiment. A common type of hypothetical experiment is repeating a well-defined task infinitely often under similar conditions.

Let $\boldsymbol{S}$ denote the sample space of some experiment, then each possible outcome s of the experiment is said to belong to the space $\boldsymbol{S}$. This is symbolically denoted by the relation $s \in \boldsymbol{S}$. For a sample space $\boldsymbol{S}$ with only finitely many outcomes, the collection of all subsets of $\boldsymbol{S}$ satisfies the following conditions:

1. The sample space $\boldsymbol{S}$ must be an event.
2. If A is an event, then A^c is also an event.
3. If $A_1, A_2, \cdots$ is a countable collection of events, then $\cup_{i=1}^{\infty} A_i$ is also an event.

In a given experiment, it is necessary to assign to each event A in the sample space $\boldsymbol{S}$ a number $Pr(A)$ that indicates the probability that A will occur. In order to satisfy the mathematical definition of probability, the number $Pr(A)$ that is assigned must satisfy three specific axioms.

1. For every event A, $Pr(A) \geq 0$. The probability of every event must be nonnegative.
2. $Pr(S) = 1$, if an event is certain to occur, then the probability of that event is 1.
3. For every infinite sequence of disjoint events $A_1, A_2, \cdots$, $Pr(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} Pr(A_i)$.

Probability: A probability measure, or simply a probability, on a sample space $\boldsymbol{S}$ is a specification of numbers $Pr(A)$ for all events A that satisfy Axioms 1, 2, and 3.

Properties of probability, having $A \in \boldsymbol{S}$:

- For every event A, $Pr(A^c) = 1 - Pr(A)$.
- For every event A, $0 \leq Pr(A) \leq 1$.
- If $A \subset B$, then $Pr(A) \leq Pr(B)$.
- For every two events A and B, $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$.

From now on, we will consider experiments for which the sample space $\boldsymbol{S}$ contains only a finite number of points $s_1, ..., s_n$. The number p_i is the probability that the outcome of the experiment will be s_i with $i = 1, ..., n$. In order to satisfy the axioms of probability, the numbers $p_1, ..., p_n$ must satisfy the following two conditions:

$$p_i \geq 0 \quad \text{for} \quad i = 1, ..., n \quad , \quad \sum_{i=1}^{n} p_i = 1. \tag{2.2.1}$$

Suppose that we learn that an event B has occurred and that we wish to compute the probability of another event A taking into account that we know that B has

occurred. The new probability of A is called the **conditional probability** of the event A given that the event B has occurred and is denoted $Pr(A|B)$. If $Pr(B) > 0$, we compute this probability as

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)}. \tag{2.2.2}$$

The conditional probability $Pr(A|B)$ is not defined if $Pr(B) = 0$. If learning that B has occurred does not change the probability of A, then we say that A and B are *independent*. Two events A and B are independent if $Pr(A \cap B) = Pr(A)Pr(B)$.

Theorem 2 (Multiplication rule for conditional probabilities)**.** *Let A and B be events. If $Pr(B) > 0$, then*

$$Pr(A \cap B) = Pr(B)Pr(A|B). \tag{2.2.3}$$

If $Pr(A) > 0$, then

$$Pr(A \cap B) = Pr(A)Pr(B|A). \tag{2.2.4}$$

From this, if $Pr(A) > 0$ and $Pr(B) > 0$ we can say that

$$Pr(B|A) = \frac{Pr(A|B)Pr(B)}{Pr(A)}. \tag{2.2.5}$$

So, the conditional probability of an event B given that A happened can be calculated if the conditional probability of A given that B happened is known. This is known as Bayes theorem.

The sample space S can be divided into as many events B_i as we want, this is called a partition. It's definition defines an alternative version of Bayes theorem.

Partition Let S denote the sample space of some experiment, and consider k events $B_1, ..., B_k$ in S such that those events are disjoint and $\cup_{i=1}^{k} B_i = S$. It is said that these events form a partition of S.

Theorem 3 (Law of total probability)**.** *Suppose that the events $B_1, ..., B_k$ form a partition of the space S and $Pr(B_j) > 0$ for $j = 1, ..., k$. Then, for every event A in S,*

$$Pr(A) = \sum_{j=1}^{k} Pr(A \cap B_j) = \sum_{j=1}^{k} Pr(B_j)Pr(A|B_j). \tag{2.2.6}$$

From the multiplication rule for conditional probabilities and the law of total probabilities we get an important theorem of conditional probabilities.

Theorem 4 (Bayes)**.** *Let the events $B_1, ..., B_k$ form a partition of the space $\boldsymbol{S}$ such that $Pr(B_j) > 0$ for $j = 1, ..., k$,and let A be an event such that $Pr(A) > 0$. Then, for $i = 1, ..., k$,*

$$Pr(B_i|A) = \frac{Pr(B_i)Pr(A|B_i)}{\sum_{j=1}^{k} Pr(B_j)Pr(A|B_j)}. \tag{2.2.7}$$

2.2.2 Random Variables

A *random variable* is a real-valued function defined in a sample space. Random variables are the main tools used to model unknown quantities in statistical analysis. When a probability measure has been specified in the sample space of an experiment, we can determine probabilities associated with the possible values of each random variable X. The *distribution* of X is the collection of all probabilities of the form $Pr(X \in C)$ for all sets C of real numbers such that $\{X \in C\}$ is an event.

We say that a random variable X has a **discrete distribution** or that X is a **discrete random variable** if X can take only a finite number k of different values $x_1, ..., x_k$ or, at most, an infinite sequence of different values $x_1, x_2, ...$

If a random variable X has a discrete distribution, the **probability function** of X is defined as the function f such that for every real number x, $f(x) = Pr(X = x)$.

Let X be a discrete random variable with probability function f. If x is not one of the possible values of X, then $f(x) = 0$. Also, if the sequence $x_1, x_2, ...$ includes all the possible values of X, then $\sum_{i=1}^{\infty} f(x_i) = 1$.

Two random variables X and Y are *independent* if, for every two sets A and B of real numbers such that $\{X \in A\}$ and $\{Y \in B\}$ are events, $Pr(X \in A \text{ and } Y \in B) = Pr(X \in A)Pr(Y \in B)$.

2.2.3 Expectation Values and Covariance

Let X be a bounded discrete random variable whose probability function is f. The **expectation** of X, denoted by $\mathbb{E}(X)$, is defined as follows:

$$\mathbb{E}(x) = \sum_{x \in X} x f(x). \tag{2.2.8}$$

The expectation of X is also referred to as the *mean* of X or the *expected value* of X. Now, let r be a real valued function of a real variable, then the expected value for r in X is given by

$$\mathbb{E}[r(X)] = \sum_{x \in X} r(x) f(x). \tag{2.2.9}$$

if the mean exists.

Properties of the expectation:

- Linear Function. If $Y = aX + b$, where a and b are finite constants, then $\mathbb{E}(Y) = a\mathbb{E}(X) + b$.
- If $X_1, ..., X_n$ are n random variables such that each expectation $\mathbb{E}(X_i)$ is finite $(i = 1, ..., n)$, then $\mathbb{E}(X_1 + ... + X_n) = \mathbb{E}(X_1) + ... + \mathbb{E}(X_n)$.
- If $X_1, ..., X_n$ are n independent random variables such that each expectation $\mathbb{E}(X_i)$ is finite $(i = 1, ..., n)$, then

$$\mathbb{E}\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} \mathbb{E}(X_i). \tag{2.2.10}$$

Let X be a random variable with finite mean $\mu = \mathbb{E}(X)$. The **variance** of X, denoted by $Var(X)$, is defined as follows:

$$Var(X) = \mathbb{E}[(X - \mu)^2]. \tag{2.2.11}$$

it gives some measure of how spread out is the distribution of X. If X has infinite mean or if the mean of X does not exist, we say that $Var(X)$ does not exist. We

can find an alternative expression for the variance,

$$Var(X) = \mathbb{E}(X^2 - 2X\mu + \mu^2), \tag{2.2.12}$$
$$= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mu + \mu^2, \tag{2.2.13}$$
$$= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2. \tag{2.2.14}$$

The **standard deviation** of X, denoted as σ_x, is the square root of $Var(X)$ if the variance exists.

Properties of the variance:

- For each X, $Var(X) \geq 0$. If X is a bounded random variable, then $Var(X)$ must exist and be finite.
- $Var(X) = 0$ if and only if there exists a constant c such that $Pr(X = c) = 1$.
- For constants a and b, let $Y = aX + b$. Then $Var(Y) = a^2 Var(X)$, and $\sigma_Y = |a|\sigma_X$.
- If $X_1, ..., X_n$ are independent random variables with finite means, then $Var(X_1 + ... + X_n) = Var(X_1) + ... + Var(X_n)$.

Every number m that divides the total probability into two equal parts, then m will be a *median* of the distribution of X. That is, $Pr(X \leq m) = 1/2$ and $Pr(X \geq m) = 1/2$.

Let X and Y be random variables having finite means. Let $\mathbb{E}(X) = \mu_X$ and $\mathbb{E}(Y) = \mu_Y$ be their expected values. The **covariance** of X and Y, which is denoted by $Cov(X, Y)$, is defined as,

$$Cov(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)], \tag{2.2.15}$$

if the expectation exists. For all random variables X and Y such that $\sigma_X < \infty$ and $\sigma_Y < \infty$,

$$Cov(X, Y) = \mathbb{E}\left[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right], \tag{2.2.16}$$
$$= \mathbb{E}\left[XY - \mathbb{E}(X)Y + X\mathbb{E}(Y) - \mathbb{E}(X)\mathbb{E}(Y)\right], \tag{2.2.17}$$
$$= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y). \tag{2.2.18}$$

The covariance between X and Y measure the degree to which X and Y tend to

be large at the same time or the degree to which one tends to be large while the other is small.

If X and Y are independent random variables with $0 < Var(X) < \infty$ and $0 < Var(Y) < \infty$,then

$$Cov(X, Y) = 0. \tag{2.2.19}$$

Schwarz Inequality: For all random variables X and Y such that $\mathbb{E}(XY)$ exists. Then

$$[\mathbb{E}(XY)]^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2). \tag{2.2.20}$$

Cauchy-Schwarz Inequality: Let X and Y be random variables with finite variance. Then

$$[Cov(X, Y)]^2 \leq Var(X)Var(Y). \tag{2.2.21}$$

2.2.4 Large Random Samples

Law of Large Numbers: Suppose that $X_1, ..., X_n$ form a random sample from a distribution for which the mean is μ and for which the variance is finite. Let $\overline{X}_n$ denote the sample mean. Then, $\overline{X}_n$ converges to μ as $n \to \infty$.

It follows that there is high probability that $\overline{X}_n$ will be close to μ if the sample size n is large. Hence, if a large random sample is taken from a distribution for which the mean is unknown, then the arithmetic average of the values in the sample will usually be a close estimate of the unknown mean.

2.2.5 Statistical Inference

In general, statistical inference consists of making probabilistic statements about unknown quantities. Our goal will be to say what we have learned about the unknown quantities after observing some data that we believe contain relevant information.

A **statistical model** consists of an identification of random variables of interest, a specification of a joint distribution for the observable random variables and the identification of any parameters of those distributions that are assumed unknown and hypothetically observable. A **statistical inference** is a procedure that

produces a probabilistic statement about some or all parts of a statistical model.

The distribution that one assigns to θ before observing the other random variables of interest is called its **prior distribution**, $\xi(\theta)$, and it describes our uncertainty about the parameter before observing any data. The conditional distribution of θ given that $X_1, ..., X_n$ is observed is called the **posterior distribution** of θ, $\xi(\theta|x_1, ..., x_n)$.

Suppose that the n random variables $X_1, ..., X_n$ form a random sample from a distribution for which the probability function is $f(\boldsymbol{x}|\theta)$ with $\boldsymbol{x} = (x_1, ..., x_n)$ the observed variable. Suppose also that the value of the parameter θ is unknown and the prior p.f. of θ is $\xi(\theta)$. Then the posterior probability function of θ is obtained using Bayes' theorem for random variables

$$\xi(\theta|\boldsymbol{x}) = \frac{f(x_1|\theta) \cdots f(x_n|\theta)\xi(\theta)}{g_n(\boldsymbol{x})}, \text{ for } \theta \in \Omega, \tag{2.2.22}$$

where Ω is the space of all possible values of the parameter θ and g_n is the marginal joint probability function of $X_1, ..., X_n$

$$g_n(\boldsymbol{x}) = \int_\Omega f(x_1|\theta) \dots f(x_n|\theta)\xi(\theta)d\theta \tag{2.2.23}$$

When the joint probability function of the observations in a random sample is regarded as a function of θ for given values of $x_1, ..., x_n$,

$$f_n(\boldsymbol{x}|\theta) = f(x_1|\theta) \dots f(x_n|\theta) \tag{2.2.24}$$

it is called the **likelihood function**. The likelihood tells us how much the data will alter our uncertainty.

Let $X_1, ..., X_n$ be observable data whose joint distribution is indexed by a parameter θ taking values in a subset Ω of the real line. An *estimator* of the parameter θ is a real-valued function $\delta(X_1, ..., X_n)$ that aims to find the real value of the parameter. If $X_1 = x_1, ..., X_n = x_n$ are observed, then $\delta(x_1, ..., x_n)$ is called the *estimate* of θ. An estimator $\delta(X)$ is an **unbiased estimator** of a function $g(\theta)$ of the parameter θ if

$$\mathbb{E}_\theta[\delta(X)] = g(\theta) \tag{2.2.25}$$

for every possible value of θ. The bias of the estimator is the difference between

the expectation of the estimator and $g(\theta)$.

2.2.6 Maximum Likelihood Estimators

Maximum likelihood estimation is a method that chooses as the estimate of θ the value of θ that provides the largest value of the likelihood function. Therefore, the maximum likelihood estimate is the value of θ that assigned the highest probability to seeing the observed data. It was introduced by R. A. Fisher in 1912.

Let the random variables $X_1, ..., X_n$ form a random sample from a discrete distribution for which the probability function is $f(x|\theta)$. Where the parameter θ can be either a real-valued parameter or a vector. For every observed vector $\boldsymbol{x} = (x_1, ..., x_n)$ in the sample, the value of the joint probability function 2.2.24 will be denoted by $f_n(\boldsymbol{x}|\theta)$. For each possible observed vector $\boldsymbol{x}$, let $\delta(\boldsymbol{x})$ denote a value of θ for which the likelihood function $f_n(\boldsymbol{x}|\theta)$ is a maximum, and let $\hat{\theta} = \delta(\boldsymbol{x})$ be the estimator of θ. This estimator is called a **maximum likelihood estimator** of θ. After $X = x$ is observed, the value $\delta(\boldsymbol{x})$ is called a **maximum likelihood estimate** of θ.

Properties:

1. Let $\hat{\theta}$ be an M.L.E. of θ, and let $g(\theta)$ be a function of θ. Then an M.L.E. of $g(\theta)$ is $g(\hat{\theta})$.
2. The sequence of M.L.E.'s converges to the unknown value of θ as the sample size $n \to \infty$.

2.3 Fisher Information

This section introduces a method for measuring the amount of information that a sample of data contains about an unknown parameter. The information measure can be used to find bounds on the variances of estimators, and it can be used to approximate the variances of estimators obtained from large samples.

The Fisher information is one property of a distribution that can be used to measure how much information one is likely to obtain from a random variable or a random sample.

2.3.1 Single Parameter

Let's consider a random experiment where we extract n outcomes $\boldsymbol{x}$, and we aim to estimate an unknown parameter ϕ, which is encoded in a physical system. We assume that the value of the parameter does not change with each measurement but the result of the experiments do change. So we have a set of possible outcomes $\{x_i\}$ with probability distribution $p(x|\phi)$, this is the conditional probability of the outcome x_i given that the unknown parameter is ϕ. This distribution is known.

We assume an estimator $\tilde{\phi}(x_i)$ of ϕ. The precision of the estimation is given by the variance of the distribution of the estimator. Now, we will find a bound for the variance of an unbiased estimator.

From the normalization property of the probability distribution, we have

$$\sum_{i=1}^{n} p(x_i|\phi) = 1. \tag{2.3.1}$$

Now, deriving this quantity with respect to the parameter ϕ,

$$\begin{aligned}
\frac{\partial}{\partial\phi}\sum_{i=1}^{n} p(x_i|\phi) &= \sum_{i=1}^{n} \frac{\partial}{\partial\phi} p(x_i|\phi)\,, && (2.3.2)\\
&= \sum_{i=1}^{n} \left(\frac{\partial}{\partial\phi}\log p(x_i|\phi)\right) p(x_i|\phi)\,, && (2.3.3)\\
&= \mathbb{E}\left(\frac{\partial}{\partial\phi}\log p(x|\phi)\right)\,, && (2.3.4)\\
&= \mathbb{E}(Z) = 0. && (2.3.5)
\end{aligned}$$

That is, the random variable $Z = \frac{\partial}{\partial\phi}\log p(x|\phi)$ has expected value zero. Since, we are working with unbiased estimators $E(\tilde{\phi}) = \phi$ and assuming that the x_i are independently and identically distributed,

$$\phi = \sum_{i=1}^{n} \tilde{\phi}(x_i) p(x_i|\phi). \tag{2.3.6}$$

Deriving this quantity with respect of ϕ, considering that the estimator $\tilde{\phi}$ does

not depend on the parameter ϕ,

$$\sum_{i=1}^{n} \tilde{\phi}(x_i) \frac{\partial}{\partial \phi} p(x_i|\phi) = \sum_{i=1}^{n} \left(\tilde{\phi}(x_i) \frac{\partial}{\partial \phi} \ln p(x_i|\phi) \right) p(x_i|\phi) \,, \tag{2.3.7}$$

$$= \mathbb{E}(\tilde{\phi} Z) \,, \tag{2.3.8}$$

$$= 1. \tag{2.3.9}$$

We have that $\mathbb{E}(\tilde{\phi} Z) = 1$ and $\mathbb{E}(Z) = 0$. Now, from the Cauchy-Schwarz inequality for the variances we know that

$$Cov(\tilde{\phi}, Z) \leq Var(\tilde{\phi}) Var(Z). \tag{2.3.10}$$

We can calculate the covariance as,

$$Cov(\tilde{\phi}, Z) = \mathbb{E}(\tilde{\phi} Z) - \mathbb{E}(\tilde{\phi}) \mathbb{E}(Z) \,, \tag{2.3.11}$$

$$= 1. \tag{2.3.12}$$

So, the variances of these variables satisfy the inequality

$$Var(\tilde{\phi}) \geq \frac{1}{Var(Z)}. \tag{2.3.13}$$

So, the variance of the estimation is bounded by the inverse of the variance of the variable Z. Now, the variance of this new variable Z, can be calculated as

$$Var(Z) = \mathbb{E}(Z^2) - \mathbb{E}(Z)^2, \tag{2.3.14}$$

$$= \mathbb{E}\left(\left[\frac{\partial}{\partial \phi} \log p(x|\phi) \right]^2 \right). \tag{2.3.15}$$

We get the Crámer-Rao bound for estimators,

$$Var(\tilde{\phi}) \geq \frac{1}{I(\phi)}. \tag{2.3.16}$$

Where

$$I(\phi) = \mathbb{E}\left(\left[\frac{\partial}{\partial\phi}\log p(x|\phi)\right]^2\right), \tag{2.3.17}$$

$$= \sum_{i=1}^{n}\left[\frac{\partial}{\partial\phi}\log p(x_i|\phi)\right]^2 p(x_i|\phi), \tag{2.3.18}$$

$$= \sum_{i=1}^{n}\frac{1}{p(x_i|\phi)}\left[\frac{\partial}{\partial\phi}\, p(x_i|\phi)\right]^2. \tag{2.3.19}$$

is the *Fisher Information*. This quantity only depends on the conditional probability distribution. In estimation theory, the best estimator that you can build is the one with variance $1/I(\phi)$. And the bigger the Fisher Information the better the estimator. But, this bound does not tell us how to construct this optimal estimator $\tilde{\phi}$.

2.3.2 Multiparameter

Let's consider a random experiment in which we aim to estimate a set of n parameters $\boldsymbol{s} = (s_1, \cdots, s_n)^\top$. Suppose that the probability distribution of this experiment is $\{p(x|\boldsymbol{s})\}_{x\in\mathcal{I}}$, where $\mathcal{I}$ is the set of all the possible outcomes.

From the experiment we get estimates $\tilde{\boldsymbol{s}}(x) = (\tilde{s}_1(x), \cdots, \tilde{s}_n(x))^\top$ of the original parameters, which are a function of the experimental results. From now we will consider $\tilde{\boldsymbol{s}}(x)$ as the random variable. The uncertainty of the estimation of the parameters $\boldsymbol{s}$ is given by the covariance matrix,

$$C_{ij} = Cov(\tilde{s}_i, \tilde{s}_j) = \mathbb{E}\left[\left(\tilde{s}_i - \mathbb{E}(\tilde{s}_i)\right)\left(\tilde{s}_j - \mathbb{E}(\tilde{s}_j)\right)\right]. \tag{2.3.20}$$

The diagonal elements of this matrix coincide with the variance of the parameters $\boldsymbol{s}$, and taking the trace of this matrix we obtain the mean squared error of the estimation.

From the definition of probability,

$$\sum_{x\in\mathcal{I}} p(x|\boldsymbol{s}) = 1, \qquad \forall\boldsymbol{s}. \tag{2.3.21}$$

We can see that, this quantity does not depend of the parameters $\boldsymbol{s}$,

$$\frac{\partial}{\partial s_i}\sum_{x\in\mathcal{I}} p(x|\boldsymbol{s}) = \sum_{x\in\mathcal{I}} \frac{\partial p(x|\boldsymbol{s})}{\partial s_i}, \tag{2.3.22}$$

$$= \sum_{x\in\mathcal{I}} p(x|\boldsymbol{s}) \frac{\partial \ln p(x|\boldsymbol{s})}{\partial s_i}, \tag{2.3.23}$$

$$= 0. \tag{2.3.24}$$

The expected value of the parameters $\tilde{\boldsymbol{s}}$ is

$$\mathbb{E}(\tilde{s}_i) = \sum_{x\in\mathcal{I}} p(x|\boldsymbol{s})\tilde{s}_i(x). \tag{2.3.25}$$

In first place, since the estimator $\tilde{\boldsymbol{s}}$ does not depend explicitly of the original parameters, we have

$$\frac{\partial \tilde{s}_j(x)}{\partial s_i} = 0. \tag{2.3.26}$$

From this, we can define the following matrix,

$$G_{ij} = \frac{\partial}{\partial s_i}\mathbb{E}\left(\tilde{s}_j(x)\right) \tag{2.3.27}$$

$$= \frac{\partial}{\partial s_i}\left[\sum_{x\in\mathcal{I}} p(x|\boldsymbol{s})\tilde{s}_j(x)\right] \tag{2.3.28}$$

$$= \sum_{x\in\mathcal{I}} \tilde{s}_j(x)\frac{\partial p(x|\boldsymbol{s})}{\partial s_i} \tag{2.3.29}$$

$$= \sum_{x\in\mathcal{I}} p(x|\boldsymbol{s})\tilde{s}_j(x)\frac{\partial \ln p(x|\boldsymbol{s})}{\partial s_i} \tag{2.3.30}$$

$$= \mathbb{E}\left(\tilde{s}_j(x)\frac{\partial}{\partial s_i}\ln p(x|\boldsymbol{s})\right). \tag{2.3.31}$$

Then, given that (2.3.24) is equal to zero, we can subtract it from (2.3.31) without changing its value,

$$G_{ij} = \mathbb{E}\left(\tilde{s}_j(x)\frac{\partial}{\partial s_i}\ln p(x|\boldsymbol{s})\right) - \mathbb{E}(\tilde{s}_j(x))\mathbb{E}\left(\frac{\partial}{\partial s_i}\ln p(x|\boldsymbol{s})\right) \tag{2.3.32}$$

$$= \mathbb{E}\left(\frac{\partial}{\partial s_i}\ln p(x|\boldsymbol{s})\left[\tilde{s}_j(x) - \mathbb{E}\left(\tilde{s}_j(x)\right)\right]\right). \tag{2.3.33}$$

For any pair of vectors $\boldsymbol{u}$, $\boldsymbol{w} \in \mathbb{R}^n$, we have that

$$(\boldsymbol{u}^\top G \boldsymbol{w})^2 = \left(\sum_{i,j=1}^{n} u_i G_{ij} w_j \right)^2 \tag{2.3.34}$$

$$= \left[\sum_{i,j=1}^{n} u_i \mathbb{E}\left(\frac{\partial}{\partial s_i} \ln p(x|\boldsymbol{s}) \left[\tilde{s}_j(x) - \mathbb{E}\left(\tilde{s}_j(x) \right) \right] \right) w_j \right]^2 \tag{2.3.35}$$

$$= \left[\mathbb{E}\left(\sum_{i,j=1}^{n} u_i \frac{\partial}{\partial s_i} \ln p(x|\boldsymbol{s}) \left[\tilde{s}_j(x) - \mathbb{E}\left(\tilde{s}_j(x) \right) \right] w_j \right) \right]^2, \tag{2.3.36}$$

$$\leq \mathbb{E}\left(\left[\sum_{i=1}^{n} u_i \frac{\partial}{\partial s_i} \ln p(x|\boldsymbol{s}) \right]^2 \right) \mathbb{E}\left(\left[\sum_{j=1}^{n} \left[\tilde{s}_j(x) - \mathbb{E}\left(\tilde{s}_j(x) \right) \right] w_j \right]^2 \right). \tag{2.3.37}$$

$$\tag{2.3.38}$$

Using Schwarz inequality for expected values, $[\mathbb{E}(ab)]^2 \leq \mathbb{E}(a^2)\mathbb{E}(b^2)$. Explicitly solving each term, it follows

$$(\boldsymbol{u}^\top G \boldsymbol{w})^2 \leq \mathbb{E}\left(\sum_{i,j=1}^{n} u_i \frac{\partial}{\partial s_i} \ln p(x|\boldsymbol{s}) u_j \frac{\partial}{\partial s_j} \ln p(x|\boldsymbol{s}) \right) \tag{2.3.39}$$

$$\mathbb{E}\left(\sum_{k,l=1}^{n} \left[\tilde{s}_k(x) - \mathbb{E}\left(\tilde{s}_k(x) \right) \right] w_k \left[\tilde{s}_l(x) - \mathbb{E}\left(\tilde{s}_l(x) \right) \right] w_l \right)$$

$$= \left[\sum_{i,j=1}^{n} u_i \mathbb{E}\left(\frac{\partial}{\partial s_i} \ln p(x|\boldsymbol{s}) \frac{\partial}{\partial s_j} \ln p(x|\boldsymbol{s}) \right) u_j \right] \tag{2.3.40}$$

$$\left[\sum_{k,l=1}^{n} w_k \mathbb{E}\left(\left[\tilde{s}_k(x) - \mathbb{E}\left(\tilde{s}_k(x) \right) \right] \left[\tilde{s}_l(x) - \mathbb{E}\left(\tilde{s}_l(x) \right) \right] \right) w_l \right]$$

$$= \left[\sum_{i,j=1}^{n} u_i I_{ij} u_j \right] \left[\sum_{k,l=1}^{n} w_k C_{kl} w_l \right] \tag{2.3.41}$$

$$= \left(\boldsymbol{u}^\top I \boldsymbol{u} \right) \left(\boldsymbol{w}^\top C \boldsymbol{w} \right), \tag{2.3.42}$$

where C can be recognized as the covariance matrix of the estimator $\tilde{\boldsymbol{s}}$, and I is defined as the Fisher Information matrix

$$I_{ij} = \mathbb{E}\left(\frac{\partial}{\partial s_i} \ln p(x|\boldsymbol{s}) \frac{\partial}{\partial s_j} \ln p(x|\boldsymbol{s}) \right). \tag{2.3.43}$$

Then, it follows that the covariance matrix is bounded by

$$\boldsymbol{w}^\top C \boldsymbol{w} \geq \frac{\boldsymbol{u}^\top G \boldsymbol{w} \boldsymbol{w}^\top G^\top \boldsymbol{u}}{\boldsymbol{u}^\top I \boldsymbol{u}}. \tag{2.3.44}$$

The last equation defines infinite bounds due to $\boldsymbol{u}$ in the right side is arbitrary. Calculating the supreme for the right side, defining $\boldsymbol{v} = I^{1/2}\boldsymbol{u}$ and using the Cauchy-Schwarz inequality,

$$\begin{aligned}
\frac{\boldsymbol{v}^\top I^{-1/2} G \boldsymbol{w} \boldsymbol{w}^\top G^\top I^{-1/2} \boldsymbol{v}}{\boldsymbol{v}^\top \boldsymbol{v}} =& \frac{||\boldsymbol{v}^\top I^{-1/2} G \boldsymbol{w}||^2}{\boldsymbol{v}^\top \boldsymbol{v}} && (2.3.45)\\
\leq& \frac{||\boldsymbol{v}||^2 ||I^{-1/2} G \boldsymbol{w}||^2}{||\boldsymbol{v}||^2} && (2.3.46)\\
=& ||I^{-1/2} G \boldsymbol{w}||^2 && (2.3.47)\\
=& \boldsymbol{w}^\top G^\top I^{-1} G \boldsymbol{w}. && (2.3.48)
\end{aligned}$$

Then, the bigger bound for C is given by

$$\boldsymbol{w}^\top C \boldsymbol{w} \geq \boldsymbol{w}^\top G^\top I^{-1} G \boldsymbol{w}. \tag{2.3.49}$$

Then we get,

$$C \geq G^\top I^{-1} G, \tag{2.3.50}$$

which is known as the Crámer-Rao inequality. This inequality implies that the matrix $C - GI^{-1}G$ is semidefinite positive. A specific case of this inequality arises when the estimator is unbiased,

$$\mathbb{E}[\tilde{\boldsymbol{s}}(x)] = \boldsymbol{s}, \tag{2.3.51}$$

so, the distribution mean of the estimator is equal to the real value of the parameter. Here, the covariance matrix is the quadratic error matrix and the matrix G is the

identity,

$$C_{ij} = \mathbb{E}[(\tilde{s}_i - \mathbb{E}(\tilde{s}_i))(\tilde{s}_j - \mathbb{E}(\tilde{s}_j)], \tag{2.3.52}$$

$$= \mathbb{E}[(\tilde{s}_i - s_i)(\tilde{s}_j - s_j)]. \tag{2.3.53}$$

$$G_{ij} = \frac{\partial}{\partial s_i}\mathbb{E}(\tilde{s}_j(x)), \tag{2.3.54}$$

$$= \frac{\partial s_j}{\partial s_i} = \delta_{ij}. \tag{2.3.55}$$

Then, the Cramer-Rao bound for unbiased estimator is

$$C \geq I^{-1}. \tag{2.3.56}$$

This bound establishes a fundamental limit on the precision of estimators, and the Fisher Information Matrix characterizes the maximum amount of information that a dataset can provide about the parameters of a statistical model. The optimal estimator is the one that achieves $C = I^{-1}$. The Maximum Likelihood estimator is an asymptotically optimal estimator, meaning that as the sample size n tends to infinity the estimation tends to saturate the CRB.

Chapter 3

Quantum Mechanics

Quantum Mechanics provides a mathematical framework essential for describing physical systems at microscopic scale. Its core postulates link observable behavior to abstract mathematical objects, enabling the prediction of quantum phenomena.

A fundamental aspect of this framework is the representation of quantum states, either pure states as vectors in Hilbert space or mixed states as operators. Additionally, distance measures between quantum states play a key role in quantum computing, quantum tomography, and quantum error correction.

This chapter introduces the postulates governing quantum states, their evolution, and methods for comparing them.

3.1 Postulates of Quantum Mechanics

Quantum Mechanics is based on a set of fundamental postulates. These postulates introduce the concept of quantum states in Hilbert spaces, observables as Hermitian operators, unitary evolution governed by the evolution operator, and the probabilistic nature of measurements. In this section, we present these postulates, which are essential for understanding quantum state estimation and tomography.

3.1.1 State Space

Postulate 1 (Quantum State)**.** *Any isolated physical system is associated with a complex vector space equipped with an inner product, known as a Hilbert*

space. The state of the system is completely described by a unit vector in this space, referred to as the system's state vector.

The simplest quantum system is the qubit. A qubit has a two-dimensional state space, which can be described using an orthonormal basis $\{|0\rangle, |1\rangle\}$. In this basis, an arbitrary quantum state can be expressed as

$$|\Psi\rangle = \alpha |0\rangle + \beta |1\rangle , \tag{3.1.1}$$

where $\alpha, \beta \in \mathbb{C}$ are complex coefficients. The requirement that $|\Psi\rangle$ be a unit vector, given by $\langle\Psi|\Psi\rangle = 1$, imposes the normalization condition

$$|\alpha|^2 + |\beta|^2 = 1. \tag{3.1.2}$$

As we will see later, this condition ensures that the total probability of measuring the system in any basis state sums to one, a fundamental principle of Quantum Mechanics. It can be seen that, multiplying the state by a complex phase $e^{i\phi} |\Psi\rangle$, the probabilities remain unchanged and the physical state is the same. We say that the state $e^{i\phi} |\Psi\rangle$ is equivalent to $|\Psi\rangle$, up to the global phase factor $e^{i\phi}$.

The **density operator** formalism serves as a fundamental tool for describing quantum systems when their exact state is a statistical mixture of pure states. Specifically, consider a quantum system that may be found in one of several pure states $|\psi_i\rangle$, each with an associated probability p_i. The set $\{p_i, |\psi_i\rangle\}$ constitutes an ensemble of pure states. The mixture of these states is captured by the density operator, defined as

$$\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i| . \tag{3.1.3}$$

Also known as the density matrix, ρ extends the description of quantum states beyond pure states, allowing for a more general representation of quantum systems. Furthermore, the density operator formalism provides an alternative yet equivalent framework for expressing the fundamental postulates of Quantum Mechanics.

Theorem 5 (Characterization of Density Operators)**.** *An operator ρ represents a valid density operator associated with some ensemble $\{p_i, |\psi_i\rangle\}$ if and only if it*

satisfies the following conditions

$$\textbf{\textit{Trace condition}}\ \mathrm{tr}(\rho) = 1. \tag{3.1.4}$$

$$\textbf{\textit{Positivity condition}}\ \rho \geq 0. \tag{3.1.5}$$

3.1.1.1 Pauli Matrices

The Pauli matrices, denoted X, Y, and Z, are a set of 2 by 2 matrices that are extremely useful in Quantum Mechanics. Together with the identity matrix, they form a basis for expanding density operators for quantum states in a 4-dimensional Hilbert space.

The Pauli matrices are given by

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.1.6}$$

They can also be written in the ket representation,

$$I = |0\rangle\langle 0| + |1\rangle\langle 1|, \tag{3.1.7}$$

$$X = |0\rangle\langle 1| + |1\rangle\langle 0|, \tag{3.1.8}$$

$$Y = i\,|0\rangle\langle 1| - i\,|1\rangle\langle 0|, \tag{3.1.9}$$

$$Z = |0\rangle\langle 0| - |1\rangle\langle 1|. \tag{3.1.10}$$

We can see that the matrices X, Y and Z are traceless, and they satisfy the commutation relations,

$$[X, Y] = 2iZ\,;\ [Y, Z] = 2iX\,;\ [Z, X] = 2iY. \tag{3.1.11}$$

Using these matrices, any 2-dimensional density operator can be expanded as

$$\rho = \frac{I + \vec{r}\cdot\vec{\sigma}}{2}, \tag{3.1.12}$$

where $\vec{r} = (r_1, r_2, r_3)^\top$ and $\vec{\sigma} = (X, Y, Z)$. For this operator to be a valid density matrix it needs to satisfy the conditions 3.1.5. It can be seen easily that the trace

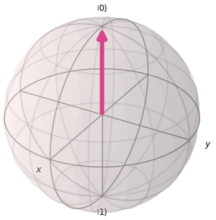

Figure 3.1.1: Bloch sphere representation of the pure state $|0\rangle$ of a qubit.

condition is fulfilled. The positivity condition means that the eigenvalues of ρ must be positive, finding the eigenvalues

$$det(I + \vec{r}\cdot\vec{\sigma} - \lambda I) = det((1-\lambda)I + \vec{r}\cdot\vec{\sigma}), \tag{3.1.13}$$

$$= det\begin{pmatrix} 1-\lambda+r_3 & r_1+ir_2 \\ r_1-ir_2 & 1-\lambda-r_3 \end{pmatrix}, \tag{3.1.14}$$

$$= (1-\lambda+r_3)(1-\lambda-r_3) - (r_1+ir_2)(r_1-ir_2), \tag{3.1.15}$$

$$= (1-\lambda)^2 - r_1^2 - r_2^2 - r_3^2, \tag{3.1.16}$$

$$= (1-\lambda)^2 - |\vec{r}|^2 = 0. \tag{3.1.17}$$

The eigenvalues are $\lambda = 1 - |\vec{r}|$ so, to satisfy the positivity condition $|\vec{r}| \leq 1$. The equality $|\vec{r}| = 1$ holds when the state is pure, and strict inequality $|\vec{r}| < 1$ holds when the state is mixed. This representation is called the Bloch representation and provides a convenient way to visualize qubit states as seen in figure 3.1.1. The vector $\vec{r}$ defines a point on the Bloch sphere, which can be used to graphically represent quantum states of a qubit.

3.1.2 Evolution

Postulate 2 (Unitary Evolution)**.** *The evolution of a closed quantum system is described by a unitary transformation. Specifically, the state $|\Psi\rangle$ of the system at time t_1 is related to the state $|\Psi'\rangle$ at time t_2 by a unitary operator U, which depends only on the times t_1 and t_2*

$$|\Psi'\rangle = U(t_1, t_2)\,|\Psi\rangle\,. \tag{3.1.18}$$

Similarly, the evolution of a quantum state represented by the density matrix ρ is given by

$$\rho' = U(t_1, t_2)\rho U(t_1, t_2)^\dagger, \tag{3.1.19}$$

where $U^\dagger$ is the Hermitian conjugate (or adjoint) of U. This evolution is also unitary and ensures the preservation of the trace and positivity of the density operator.

3.1.3 Quantum Measurement

Postulate 3 (Quantum Measurement)**.** *Quantum measurements are described by a collection $\{M_m\}$ of measurement operators, which act on the state space of the system being measured. The measurement operators satisfy the completeness relation*

$$\sum_m M_m^\dagger M_m = \mathbb{I}. \tag{3.1.20}$$

The index m corresponds to the possible outcomes of the measurement. If the state of the quantum system is $|\Psi\rangle$ immediately before the measurement, the probability of obtaining outcome m is given by

$$p(m) = \langle\Psi| M_m^\dagger M_m |\Psi\rangle . \tag{3.1.21}$$

After the measurement, the state of the system collapses to

$$|\Psi_m\rangle = \frac{M_m |\Psi\rangle}{\sqrt{\langle\Psi| M_m^\dagger M_m |\Psi\rangle}}. \tag{3.1.22}$$

For a quantum system described by the density matrix ρ, immediately before the measurement, the probability of obtaining outcome m is given by

$$p(m) = \mathrm{tr}\big(M_m^\dagger M_m \rho\big), \tag{3.1.23}$$

and the state of the system after the measurement is

$$\rho_m = \frac{M_m^\dagger \rho M_m}{\mathrm{tr}\Big(M_m^\dagger M_m \rho\Big)}. \tag{3.1.24}$$

This postulate tell us that the act of measuring a quantum system changes the state. Thus, the act of observing a quantum system affects the reality of said system. Thing that does not happen classically.

Also, since the system is not the same we loose information about the original

state. Meaning that, in order to obtain all the information about the quantum system, the quantum state, we must perform measures in different copies of the system.

3.1.3.1 Projective Measurements

A projective measurement is characterized by an observable M, which is a Hermitian operator acting on the state space of the system being measured. The observable M can be expressed in terms of its spectral decomposition as

$$M = \sum_m m P_m, \tag{3.1.25}$$

where P_m denotes the projector onto the eigenspace associated with the eigenvalue m. These P_m operators also are Hermitian, $P_m^\dagger = P_m$ and $P_i P_j = \delta_{ij} P_i$. The possible measurement outcomes correspond to the eigenvalues m of M. When a quantum state $|\Psi\rangle$ is measured, the probability of obtaining the outcome m is given by

$$p(m) = \langle\Psi| P_m |\Psi\rangle , \tag{3.1.26}$$

which follows from *Born's rule*. After the measurement, the system collapses into the corresponding post-measurement state

$$|\Psi_m\rangle = \frac{P_m |\Psi\rangle}{\sqrt{\langle\Psi| P_m |\Psi\rangle}}. \tag{3.1.27}$$

If the system is instead described by a density operator ρ, the probability of obtaining the outcome m upon measuring P_m is given by

$$p(m) = \mathrm{tr}(P_m \rho), \tag{3.1.28}$$

and the post-measurement state is

$$\rho_m = \frac{P_m \rho P_m}{\mathrm{tr}(P_m \rho)}. \tag{3.1.29}$$

3.1.3.2 POVM Measurements

When we are only interested in the probability distribution of the outcomes and not on the post-measurement state, it is convenient to use the POVM formalism.

Consider the operator $E_m = M_m^\dagger M_m$. From Postulate 3 and basic linear algebra, it follows that E_m is a positive operator satisfying the condition

$$\sum_m E_m = \mathbb{I}. \tag{3.1.30}$$

The probability of obtaining the outcome m when measuring the state $|\Psi\rangle$ is given by

$$p(m) = \langle\Psi| E_m |\Psi\rangle . \tag{3.1.31}$$

Thus, the collection of operators $\{E_m\}$ fully determines the probabilities of all possible measurement outcomes. These operators, known as POVM elements, define the measurement process. The complete set $\{E_m\}$ is referred to as a POVM.

3.1.4 Composite Systems

Postulate 4 (State Space of Composite Systems)**.** *The state space of a composite quantum system is given by the Kronecker tensor product of the state spaces of its individual subsystems. If we consider a collection of n systems, the corresponding state space is*

$$\mathcal{H}_{tot} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n \tag{3.1.32}$$

This postulate is fundamental for describing systems composed of multiple physical subsystems, such as particles, spins, or electrons. It also applies in scenarios where a single system has multiple degrees of freedom that must be considered simultaneously.

For a collection of n systems where the i-th system is in the state $|\Psi_i\rangle$, then the overall state of the composite system is represented as

$$|\Psi_1\rangle \otimes |\Psi_2\rangle \otimes \cdots \otimes |\Psi_n\rangle . \tag{3.1.33}$$

Similarly, if each subsystem i is described by a density operator ρ_i, the joint state of the entire system is given by the tensor product of these density matrices

$$\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n. \tag{3.1.34}$$

There exist a special type of composed pure states that cannot be written as

a tensor product of the states of each subsystem, this kind of states are called *entangled states*. From now on, we will omit the tensor product symbol. For example, the Bell state for 2 qubits in the product space $\mathcal{H}_1 \otimes \mathcal{H}_2$ is a widely used entangled state

$$|\Phi\rangle_{\text{Bell}} = \frac{1}{\sqrt{2}}(|0\rangle_1 |0\rangle_2 + |1\rangle_1 |1\rangle_2), \tag{3.1.35}$$

where the subscripts indicate to what subsystem belongs each state. It is clear to see that is not possible to write this state as a product of each subsystem state.

Entanglement, corresponds to a kind of *quantum correlation*, since a measurement in one subsystem also affects the state of the other subsystem that is not being measured. For example, if we perform a measure in subsystem 1 and the $|0\rangle$ state is detected, the post-measurement state of the total system is

$$|\phi\rangle = \frac{|0\rangle_1 \langle 0|_1 |\Phi\rangle_{\text{Bell}}}{|\langle 0|_1 |\Phi\rangle_{\text{Bell}}|} = |0\rangle_1 |0\rangle_2 . \tag{3.1.36}$$

We have certainty that the system 2 is in the state $|0\rangle$. Entanglement enables stronger-than-classical correlations and for this is used as a resource for quantum technologies [3], such as quantum communication and quantum computation.

3.2 Distance Measures for Quantum States

In Quantum Mechanics, distance measures quantify the difference between quantum states, while measures like fidelity assess their similarity [3]. This section reviews the trace distance, a true metric, and the fidelity, a key tool for evaluating quantum state similarity and transformations.

Given two quantum states ρ and σ, the properties of a metric $d(\rho, \sigma)$ on the space of density operators are

1. **Non-negativity** $d(\rho, \sigma) \geq 0$, with equality only if $\rho = \sigma$.
2. **Symmetry** $d(\rho, \sigma) = d(\sigma, \rho)$.
3. **Triangle inequality** $d(\rho, \tau) \leq d(\rho, \sigma) + d(\sigma, \tau)$.

3.2.1 Trace Distance

The *trace distance* between two quantum states ρ and σ is defined as

$$D(\rho,\sigma) = \frac{1}{2}\,\mathrm{tr}\,||\rho-\sigma||, \tag{3.2.1}$$

where $||A|| = \sqrt{A^\dagger A}$ is the matrix norm. Explicitly, this distance is

$$D(\rho,\sigma) = \frac{1}{2}\,\mathrm{tr}\left(\sqrt{(\rho-\sigma)^\dagger(\rho-\sigma)}\right) \tag{3.2.2}$$

For qubits, with density operators given by

$$\rho = \frac{I+\vec{r}\cdot\vec{\sigma}}{2}, \quad \sigma = \frac{I+\vec{s}\cdot\vec{\sigma}}{2}, \tag{3.2.3}$$

the trace distance reduces to

$$\begin{aligned} D(\rho,\sigma) &= \frac{1}{2}\,\mathrm{tr}\,||\rho-\sigma||, && (3.2.4)\\ &= \frac{1}{2}\,\mathrm{tr}\left|\left|\frac{(\vec{r}-\vec{s})\cdot\vec{\sigma}}{2}\right|\right|, && (3.2.5)\\ &= \frac{1}{2}|\vec{r}-\vec{s}|. && (3.2.6) \end{aligned}$$

Thus, the distance between two single-qubit states is half the Euclidean distance between their corresponding points on the Bloch sphere.

A useful property of the trace distance is that it is invariant under unitary transformations

$$\begin{aligned} D(U\rho U^\dagger, U\sigma U^\dagger) &= \frac{1}{2}\,\mathrm{tr}\left(\sqrt{(U(\rho-\sigma)U^\dagger)^\dagger(U(\rho-\sigma)U^\dagger)}\right), && (3.2.7)\\ &= \frac{1}{2}\,\mathrm{tr}\left(\sqrt{U(\rho-\sigma)^\dagger U^\dagger U(\rho-\sigma)U^\dagger}\right), && (3.2.8)\\ &= \frac{1}{2}\,\mathrm{tr}\left(\sqrt{(\rho-\sigma)^\dagger(\rho-\sigma)}\right), && (3.2.9)\\ &= D(\rho,\sigma). && (3.2.10) \end{aligned}$$

The trace distance also satisfies the properties of a metric on the space of density operators. This metric is widely used in the context of *state discrimination* [6]. The trace distance provides an operational meaning in terms of the

maximum probability of correctly distinguishing between ρ and σ in a single-shot measurement. Specifically, if we are given a quantum state that is either ρ or σ with equal probability, and we perform the optimal measurement to distinguish them, the maximum success probability is [6]

$$P_D = \frac{1}{2}[1 + D(\rho, \sigma)]. \tag{3.2.11}$$

Where the maximum probability $P_D = 1$ is reached when the states are orthogonal, and the minimum $P_D = 0$ when the states are very close, almost identical.

3.2.2 Quantum Fidelity

The fidelity is not a metric on density operators, but it is a useful measure for comparing quantum states. The fidelity between two states ρ and σ is defined as

$$F(\rho, \sigma) = \operatorname{tr}\left(\sqrt{\rho^{1/2}\sigma\rho^{1/2}}\right)^2. \tag{3.2.12}$$

This gives a measure that quantifies how alike are two states. When two states are completely different (orthogonal) the fidelity is minimum with value zero, and its maximum is when the states are the same with value 1.

There are three important special cases where explicit formulas for fidelity can be derived. The first is when ρ and σ commute (i.e., they are diagonal in the same basis)

$$\rho = \sum_i r_i \left|i\right\rangle \left\langle i\right|, \quad \sigma = \sum_i s_i \left|i\right\rangle \left\langle i\right|, \tag{3.2.13}$$

for some orthonormal basis $\{|i\rangle\}$. In this case, we can express the fidelity as

$$F(\rho, \sigma) = \left(\operatorname{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}\right)^2, \tag{3.2.14}$$

$$= \left(\operatorname{tr}\sqrt{\sum_i r_i s_i \left|i\right\rangle \left\langle i\right|}\right)^2, \tag{3.2.15}$$

$$= \left(\sum_i \sqrt{r_i s_i}\right)^2. \tag{3.2.16}$$

The fidelity between a pure state $|\psi\rangle$ and an arbitrary state ρ is given by

$$F(|\psi\rangle, \rho) = (\mathrm{tr}\sqrt{|\psi\rangle\langle\psi|\rho|\psi\rangle\langle\psi|})^2, \tag{3.2.17}$$

$$= \langle\psi|\rho|\psi\rangle. \tag{3.2.18}$$

The fidelity between two pure states $|\psi\rangle$ and $|\phi\rangle$ is

$$F(|\psi\rangle, |\phi\rangle) = \mathrm{tr}\left(\sqrt{|\psi\rangle\langle\psi|\phi\rangle\langle\phi|}\right)^2, \tag{3.2.19}$$

$$= |\langle\phi|\psi\rangle|^2. \tag{3.2.20}$$

This quantity is invariant under unitary transformations

$$F(U\rho U^\dagger, U\sigma U^\dagger) = F(\rho, \sigma). \tag{3.2.21}$$

Although fidelity is not a metric, it can be transformed into one. Uhlmann's theorem [3] states that the fidelity between two states is equal to the maximum inner product between their purifications. This suggests we define the angle between two states ρ and σ as

$$A(\rho, \sigma) = \arccos(F(\rho, \sigma)). \tag{3.2.22}$$

This angle is non-negative, symmetric in its inputs, zero if and only if $\rho = \sigma$, and obeys the triangle inequality.

We can also define an useful quantity, named *Infidelity* defined in terms of the quantum fidelity,

$$\mathcal{I}(\rho, \sigma) = 1 - F(\rho, \sigma). \tag{3.2.23}$$

For pure states the infidelity is

$$\mathcal{I}(|\psi\rangle, |\phi\rangle) = 1 - |\langle\psi|\phi\rangle|^2. \tag{3.2.24}$$

Most of the time we are interested in the average behavior of the tomographic methods. The average infidelity is defined as the expectation of the infidelity over

the space of estimators for the same unknown state, that is,

$$\bar{\mathcal{I}}(|\psi\rangle) = \frac{1}{\Omega}\int I(|\psi\rangle, |\hat{\psi}\rangle)\, d\hat{\psi}. \tag{3.2.25}$$

The *Bures distance* [7] is another distance measure that can be written in terms of the quantum fidelity. This is defined as

$$d_B^2(\rho_1, \rho_2) = 2(1 - \sqrt{F(\rho_1, \rho_2)}). \tag{3.2.26}$$

Which has a maximum of $\sqrt{2}$ when the states are orthogonal, and a minimum of zero when the states are equal. So this quantity gives a measure of how different are two states.

The fidelity and the trace distance are related through the following inequality [3]

$$1 - \sqrt{F(\rho, \sigma)} \leq D(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)}. \tag{3.2.27}$$

The implication is that the trace distance and the fidelity are qualitatively equivalent measures of closeness for quantum states.

In the regime of infinitesimally close states ρ and $\rho + d\rho$, or high fidelity $F \approx 1$, we can approximate $\sqrt{F(\rho, \rho + d\rho)} \approx F(\rho, \rho + d\rho)$. With this, the Bures distance for close states can be written in terms of the infidelity $\mathcal{I}$ [8],

$$d_B^2(\rho_1, \rho_2) \approx \mathcal{I}. \tag{3.2.28}$$

Chapter 4

Quantum Fisher Information

4.1 Quantum Fisher Information

In this section, we generalize the Fisher information matrix and the Cramér-Rao bound to the quantum case. Mostly based on the thesis [9] and review [10].

In Quantum Mechanics, probabilities arise from measurements in experiments. Measurements are described by a set of operators $\{E_m\}$ called POVM, such that $E_m > 0$ and $\sum_m E_m = \mathbb{I}$. The probability of obtaining the outcome m when measuring E_m given that the system is in the quantum state ρ is given by Born's rule

$$P(m|\rho) = \mathrm{tr}(\rho E_m), \tag{4.1.1}$$

we can see that the distribution depends on the measurement performed.

4.1.1 Single Parameter Estimation

Consider a quantum system prepared in the state ρ, which depends on an unknown parameter ϕ. By performing a measurement $\{E_m\}$, we extract information about ϕ through the probability distribution of measurement outcomes given by Born's rule, $P(m|\rho(\phi))$, for conciseness, we write this as $P(m|\phi)$. This is analogous to the probability distribution $p(x|\boldsymbol{s})$ in the classical experiment.

The Fisher information, previously introduced in 2.3.19, quantifies the amount of

information that the measurement provides about ϕ and is given by

$$I(\phi) = \sum_m \frac{1}{p(m|\phi)} [\partial_\phi p(m|\phi)]^2. \tag{4.1.2}$$

Substituting Born's rule, we obtain

$$\partial_\phi p(m|\phi) = \partial_\phi \operatorname{tr}(E_m \rho), \tag{4.1.3}$$

$$= \operatorname{tr}(E_m \, \partial_\phi \rho). \tag{4.1.4}$$

To characterize the evolution of $\rho(\phi)$ with respect to ϕ, we implicitly define the Symmetric Logarithmic Derivative (SLD) L_ϕ as the operator that satisfies the equation

$$\partial_\phi \rho = \frac{1}{2}(L_\phi \rho(\phi) + \rho(\phi) L_\phi). \tag{4.1.5}$$

Next, we verify whether the SLD operator is Hermitian. Using the fact that both $\rho(\phi)$ and E_m are Hermitian, we take the Hermitian conjugate of both sides

$$\Big(\frac{\partial \rho(\phi)}{\partial \phi}\Big)^\dagger = \frac{\partial \rho(\phi)}{\partial \phi}, \tag{4.1.6}$$

$$= \frac{1}{2}\Big(\rho(\phi) L_\phi(\phi) + L_\phi(\phi)\rho(\phi)\Big)^\dagger, \tag{4.1.7}$$

$$= \frac{1}{2}\Big(L_\phi(\phi)^\dagger \rho(\phi)^\dagger + \rho(\phi)^\dagger L_\phi(\phi)^\dagger\Big), \tag{4.1.8}$$

$$= \frac{1}{2}\Big(L_\phi(\phi)^\dagger \rho(\phi) + \rho(\phi) L_\phi(\phi)^\dagger\Big). \tag{4.1.9}$$

Since the derivative of $\rho(\phi)$ is Hermitian, it follows that $L_\phi^\dagger = L_\phi$, confirming that the SLD operator is indeed Hermitian. Substituting this into (4.1.4), we obtain

$$\partial_\phi p(m|\phi) = \operatorname{tr}(E_m \, \partial_\phi \rho), \tag{4.1.10}$$

$$= \frac{1}{2} \operatorname{tr}\left(E_m (L_\phi \rho(\phi) + \rho(\phi) L_\phi)\right), \tag{4.1.11}$$

$$= \frac{1}{2} \left[\operatorname{tr}(E_m L_\phi \rho(\phi)) + \operatorname{tr}(E_m \rho(\phi) L_\phi)\right]. \tag{4.1.12}$$

We can use the fact that,

$$\begin{aligned}\mathrm{tr}(E_m\rho(\phi)L_\phi) &= \mathrm{tr}^*[(E_m\rho(\phi)L_\phi)^\dagger], &(4.1.13)\\ &= \mathrm{tr}^*(L_\phi\rho(\phi)E_m), &(4.1.14)\\ &= \mathrm{tr}^*(E_mL_\phi\rho(\phi)). &(4.1.15)\end{aligned}$$

So, the sum of the traces is

$$\mathrm{tr}\left(E_m\rho(\phi)L(\phi)\right) + \mathrm{tr}\left(E_mL(\phi)\rho(\phi)\right) = \mathrm{tr}\left(E_mL(\phi)\rho(\phi)\right) + \mathrm{tr}\left(E_mL(\phi)\rho(\phi)\right)^*, \tag{4.1.16}$$

$$= 2\,\mathrm{Re}\left[\mathrm{tr}\left(E_mL(\phi)\rho(\phi)\right)\right]. \tag{4.1.17}$$

Thus, the Fisher Information is given by

$$I(\phi) = \sum_m \frac{1}{\mathrm{tr}(E_m\rho)}\,\mathrm{Re}^2\left[\mathrm{tr}(E_mL_\phi\rho(\phi))\right]. \tag{4.1.18}$$

This quantity explicitly depends on both the POVM $\{E_m\}$ and the quantum state ρ. To determine its upper bound over all possible measurements, we apply the Cauchy-Schwarz inequality $\left|\mathrm{tr}\big(A^\dagger B\big)\right|^2 \le \mathrm{tr}\big(AA^\dagger\big)\,\mathrm{tr}\big(BB^\dagger\big)$,

$$\begin{aligned}I(\phi) &\le \sum_m \frac{1}{\mathrm{tr}(E_m\rho)}\left|\mathrm{tr}\left(E_mL_\phi\rho(\phi)\right)\right|^2, &(4.1.19)\\ &= \sum_m \frac{1}{\mathrm{tr}(E_m\rho)}\left|\mathrm{tr}\left(\sqrt{\rho(\phi)}\sqrt{E_m}\sqrt{E_m}L_\phi\sqrt{\rho(\phi)}\right)\right|^2, &(4.1.20)\\ &\le \sum_m \frac{1}{\mathrm{tr}(E_m\rho)}\,\mathrm{tr}\left(\sqrt{\rho(\phi)}\sqrt{E_m}\left[\sqrt{\rho(\phi)}\sqrt{E_m}\right]^\dagger\right) &(4.1.21)\\ &\quad\ \mathrm{tr}\left(\sqrt{E_m}L_\phi\sqrt{\rho(\phi)}\left[\sqrt{E_m}L_\phi\sqrt{\rho(\phi)}\right]^\dagger\right),\\ &= \sum_m \frac{1}{\mathrm{tr}(E_m\rho)}\,\mathrm{tr}\left(\rho(\phi)E_m\right)\mathrm{tr}\left(L_\phi E_mL_\phi\rho(\phi)\right), &(4.1.22)\\ &= \sum_m \mathrm{tr}\left(L_\phi E_mL_\phi\rho(\phi)\right), &(4.1.23)\\ &= \mathrm{tr}\left(L_\phi\sum_m E_m\ L_\phi\rho(\phi)\right), &(4.1.24)\\ &= \mathrm{tr}\left(L_\phi^2\rho(\phi)\right). &(4.1.25)\end{aligned}$$

where we used tha fact that the operators E_m and ρ are semidefinite positive, so we can write $E_m = \sqrt{E_m}\sqrt{E_m}$. This final expression defines the *Quantum Fisher Information* (QFI)

$$F(\phi) = \mathrm{tr}\left(L_\phi^2 \rho(\phi)\right). \tag{4.1.26}$$

This quantity depends only on the unknown state that is of interest. Since $I(\phi) \leq F(\phi)$, we obtain a fundamental bound on parameter estimation

$$\mathrm{Var}(\tilde{\phi}) \geq \frac{1}{I(\phi)} \geq \frac{1}{F(\phi)}. \tag{4.1.27}$$

This inequality, known as the *Quantum Cramér-Rao bound*, provides the ultimate limit on the precision of parameter estimation.

From the perspective of estimation theory, an optimal measurement saturates both the classical and quantum bounds. For this to happen, the following inequalities present in the previous calculations,

$$\mathrm{Re}^2\left[\mathrm{tr}(E_m L_\phi \rho(\phi))\right] = \left|\mathrm{tr}\left(E_m L_\phi \rho(\phi)\right)\right|^2. \tag{4.1.28}$$

and

$$\left|\mathrm{tr}\left(\sqrt{\rho(\phi)}\sqrt{E_m}\sqrt{E_m}L_\phi\sqrt{\rho(\phi)}\right)\right|^2 = \mathrm{tr}\left(\sqrt{\rho(\phi)}\sqrt{E_m}\left[\sqrt{\rho(\phi)}\sqrt{E_m}\right]^\dagger\right) \tag{4.1.29}$$
$$\mathrm{tr}\left(\sqrt{E_m}L_\phi\sqrt{\rho(\phi)}\left[\sqrt{E_m}L_\phi\sqrt{\rho(\phi)}\right]^\dagger\right).$$

needs to be equalities. So, the optimal measurements met the following conditions

- $\mathrm{tr}(\rho(\phi)E_m L_\phi) \in \mathbb{R}$, ensuring that the measurement outcomes remain real-valued.
- $\sqrt{E_m}\sqrt{\rho(\phi)} = c\sqrt{E_m}L_\phi\sqrt{\rho(\phi)}$, meaning that $\sqrt{E_m}$ is a projector onto the eigenstates of L_ϕ.

4.1.2 Multiparameter Estimation

Consider a d-dimensional quantum system in a state $\rho(\boldsymbol{s})$ characterized by the parameters $\boldsymbol{s} = \left(s_1 \cdots s_n\right)^\top$, which we aim to estimate. Given a set of measurement operators $\{E_i\}$ satisfying the completeness relation $\sum_i E_i = I$,

the probability of obtaining the outcome i in a measurement, according to the postulates of Quantum Mechanics, is given by

$$p(i|\boldsymbol{s}) = \operatorname{tr}\big(E_i\rho(\boldsymbol{s})\big). \tag{4.1.30}$$

Now we will find the Cramer-Rao bound for unbiased quantum estimators. Substituting this probability expression (4.1.30) into the Classical Fisher Information matrix (2.3.43) for unbiased estimators, we obtain

$$I_{ij} = \sum_{k=1}^{m} p(k|\boldsymbol{s}) \left(\frac{\partial}{\partial s_i} \ln p(k|\boldsymbol{s}) \frac{\partial}{\partial s_j} \ln p(k|\boldsymbol{s}) \right) \tag{4.1.31}$$

$$= \sum_{k=1}^{m} \frac{1}{p(k|\boldsymbol{s})} \frac{\partial p(k|\boldsymbol{s})}{\partial s_i} \frac{\partial p(k|\boldsymbol{s})}{\partial s_j} \tag{4.1.32}$$

$$= \sum_{k=1}^{m} \frac{1}{p(k|\boldsymbol{s})} \operatorname{tr}\left(E_k \frac{\partial \rho(\boldsymbol{s})}{\partial s_i}\right) \operatorname{tr}\left(E_k \frac{\partial \rho(\boldsymbol{s})}{\partial s_j}\right). \tag{4.1.33}$$

Let $\boldsymbol{w} \in \mathbb{R}^n$, then

$$\boldsymbol{w}^\top I \boldsymbol{w} = \sum_{i,j=1}^{n} w_i I_{ij} w_j \tag{4.1.34}$$

$$= \sum_{i,j=1}^{n} w_i \sum_{k=1}^{m} \frac{1}{p(k|\boldsymbol{s})} \operatorname{tr}\left(E_k \frac{\partial \rho(\boldsymbol{s})}{\partial s_i}\right) \operatorname{tr}\left(E_k \frac{\partial \rho(\boldsymbol{s})}{\partial s_j}\right) w_j \tag{4.1.35}$$

$$= \sum_{k=1}^{m} \frac{1}{p(k|\boldsymbol{s})} \operatorname{tr}\left(E_k \sum_{i=1}^{n} w_i \frac{\partial \rho(\boldsymbol{s})}{\partial s_i}\right) \operatorname{tr}\left(E_k \sum_{j=1}^{n} \frac{\partial \rho(\boldsymbol{s})}{\partial s_j} w_j\right) \tag{4.1.36}$$

$$= \sum_{k=1}^{m} \frac{1}{p(k|\boldsymbol{s})} \operatorname{tr}\left(E_k \sum_{i=1}^{n} w_i \frac{\partial \rho(\boldsymbol{s})}{\partial s_i}\right)^2. \tag{4.1.37}$$

Here define the *Symmetric Logarithmic Derivative* (SLD) operators, $L_i(\boldsymbol{s})$, corresponding to the parameters $\boldsymbol{s}$, as the operators satisfying the equation

$$\frac{\partial \rho(\boldsymbol{s})}{\partial s_i} = \frac{1}{2}\Big(\rho(\boldsymbol{s}) L_i(\boldsymbol{s}) + L_i(\boldsymbol{s})\rho(\boldsymbol{s})\Big). \tag{4.1.38}$$

Then, summing over the vector elements w_i

$$\sum_{i=1}^{n} w_i \frac{\partial \rho(\boldsymbol{s})}{\partial s_i} = \frac{1}{2} \sum_{i=1}^{n} w_i \Big(\rho(\boldsymbol{s}) L_i(\boldsymbol{s}) + L_i(\boldsymbol{s}) \rho(\boldsymbol{s}) \Big) \tag{4.1.39}$$

$$= \frac{1}{2} \left(\rho(\boldsymbol{s}) \sum_{i=1}^{n} w_i L_i(\boldsymbol{s}) + \sum_{i=1}^{n} w_i L_i(\boldsymbol{s}) \rho(\boldsymbol{s}) \right) \tag{4.1.40}$$

$$= \frac{1}{2} \left(\rho(\boldsymbol{s}) L(\boldsymbol{s}) + L(\boldsymbol{s}) \rho(\boldsymbol{s}) \right). \tag{4.1.41}$$

Defining

$$L(\boldsymbol{s}) = \sum_{i=1}^{n} w_i L_i(\boldsymbol{s}), \tag{4.1.42}$$

and substituting this into (4.1.37), we obtain

$$\boldsymbol{w}^\top I \boldsymbol{w} = \sum_{k=1}^{m} \frac{1}{p(k|\boldsymbol{s})} \operatorname{tr} \left[E_k \frac{1}{2} \Big(\rho(\boldsymbol{s}) L(\boldsymbol{s}) + L(\boldsymbol{s}) \rho(\boldsymbol{s}) \Big) \right]^2 \tag{4.1.43}$$

$$= \frac{1}{4} \sum_{k=1}^{m} \frac{1}{p(k|\boldsymbol{s})} \Big(\operatorname{tr} \big(E_k \rho(\boldsymbol{s}) L(\boldsymbol{s}) \big) + \operatorname{tr} \big(E_k L(\boldsymbol{s}) \rho(\boldsymbol{s}) \big) \Big)^2. \tag{4.1.44}$$

Since both E_k and $\rho(\boldsymbol{s})$ are Hermitian, we verify whether $L(\boldsymbol{s})$ is also Hermitian. From the definition of the SLD, we have

$$\left(\frac{\partial \rho(\boldsymbol{s})}{\partial s_i} \right)^\dagger = \frac{\partial \rho(\boldsymbol{s})}{\partial s_i} \tag{4.1.45}$$

$$= \frac{1}{2} \Big(\rho(\boldsymbol{s}) L_i(\boldsymbol{s}) + L_i(\boldsymbol{s}) \rho(\boldsymbol{s}) \Big)^\dagger \tag{4.1.46}$$

$$= \frac{1}{2} \Big(L_i(\boldsymbol{s})^\dagger \rho(\boldsymbol{s})^\dagger + \rho(\boldsymbol{s})^\dagger L_i(\boldsymbol{s})^\dagger \Big) \tag{4.1.47}$$

$$= \frac{1}{2} \Big(L_i(\boldsymbol{s})^\dagger \rho(\boldsymbol{s}) + \rho(\boldsymbol{s}) L_i(\boldsymbol{s})^\dagger \Big). \tag{4.1.48}$$

Since $\rho(\boldsymbol{s})$ is Hermitian, for the equality to hold, it must be that $L_i^\dagger = L_i$, meaning that each SLD operator L_i is Hermitian. Consequently, the sum $L(\boldsymbol{s}) =$

$\sum_{i=1}^{n} w_i L_i(\boldsymbol{s})$ is also Hermitian. With this, we observe that

$$\operatorname{tr}\Big(E_k\rho(\boldsymbol{s})L(\boldsymbol{s})\Big) = \operatorname{tr}\Big(E_k^\dagger\rho(\boldsymbol{s})^\dagger L(\boldsymbol{s})^\dagger\Big) \tag{4.1.49}$$

$$= \operatorname{tr}\Big[\Big(L(\boldsymbol{s})\rho(\boldsymbol{s})E_k\Big)^\dagger\Big] \tag{4.1.50}$$

$$= \operatorname{tr}\Big(E_k L(\boldsymbol{s})\rho(\boldsymbol{s})\Big)^*, \tag{4.1.51}$$

where we used the cyclic property of the trace operation and the identity $\operatorname{tr}\big(A^\dagger\big) = \operatorname{tr}(A)^*$. Then, we obtain

$$\operatorname{tr}\Big(E_k\rho(\boldsymbol{s})L(\boldsymbol{s})\Big) + \operatorname{tr}\Big(E_k L(\boldsymbol{s})\rho(\boldsymbol{s})\Big) = \operatorname{tr}\Big(E_k L(\boldsymbol{s})\rho(\boldsymbol{s})\Big) + \operatorname{tr}\Big(E_k L(\boldsymbol{s})\rho(\boldsymbol{s})\Big)^*, \tag{4.1.52}$$

$$= 2\operatorname{Re}\Big[\operatorname{tr}\Big(E_k L(\boldsymbol{s})\rho(\boldsymbol{s})\Big)\Big]. \tag{4.1.53}$$

Substituting this into (4.1.44), we find

$$\boldsymbol{w}^\top I\boldsymbol{w} = \frac{1}{4}\sum_{k=1}^{m}\frac{1}{p(k|\boldsymbol{s})}\Big(2\operatorname{Re}\Big[\operatorname{tr}\Big(E_k L(\boldsymbol{s})\rho(\boldsymbol{s})\Big)\Big]\Big)^2 \tag{4.1.54}$$

$$= \sum_{k=1}^{m}\frac{1}{p(k|\boldsymbol{s})}\Big[\operatorname{Re}\operatorname{tr}\Big(E_k L(\boldsymbol{s})\rho(\boldsymbol{s})\Big)\Big]^2 \tag{4.1.55}$$

$$\leq \sum_{k=1}^{m}\frac{1}{p(k|\boldsymbol{s})}\Big|\operatorname{tr}\Big(E_k L(\boldsymbol{s})\rho(\boldsymbol{s})\Big)\Big|^2. \tag{4.1.56}$$

Since the operators $E_k, \rho(s)$ are positives, we can decompose them as the product of their square roots $E_k = \sqrt{E_k}\sqrt{E_k}$, using this in (4.1.56)

$$\boldsymbol{w}^\top I\boldsymbol{w} = \sum_{k=1}^{m}\frac{1}{p(k|\boldsymbol{s})}\left|\operatorname{tr}\left(\sqrt{E_k}\sqrt{E_k}L(\boldsymbol{s})\sqrt{\rho(\boldsymbol{s})}\sqrt{\rho(\boldsymbol{s})}\right)\right|^2 \tag{4.1.57}$$

$$= \sum_{k=1}^{m}\frac{1}{p(k|\boldsymbol{s})}\left|\operatorname{tr}\left(\sqrt{\rho(\boldsymbol{s})}^\dagger\sqrt{E_k}^\dagger\sqrt{E_k}L(\boldsymbol{s})\sqrt{\rho(\boldsymbol{s})}\right)\right|^2 \tag{4.1.58}$$

$$= \sum_{k=1}^{m}\frac{1}{p(k|\boldsymbol{s})}\left|\operatorname{tr}\left(\left[\sqrt{E_k}\sqrt{\rho(\boldsymbol{s})}\right]^\dagger\left[\sqrt{E_k}L(\boldsymbol{s})\sqrt{\rho(\boldsymbol{s})}\right]\right)\right|^2. \tag{4.1.59}$$

$$\tag{4.1.60}$$

Using the Cauchy-Schwarz inequality $|\operatorname{tr}(A^\dagger B)|^2 \leq \operatorname{tr}(AA^\dagger)\operatorname{tr}(BB^\dagger)$ it follows

$$\begin{aligned}
\boldsymbol{w}^\top I \boldsymbol{w} \leq & \sum_{k=1}^{m} \frac{1}{p(k|\boldsymbol{s})} \operatorname{tr}\Big(E_k \rho(\boldsymbol{s})\Big) \operatorname{tr}\Big(E_k L(\boldsymbol{s}) \rho(\boldsymbol{s}) L(\boldsymbol{s})\Big), && (4.1.61)\\
= & \sum_{k=1}^{m} \operatorname{tr}\Big(E_k L(\boldsymbol{s}) \rho(\boldsymbol{s}) L(\boldsymbol{s})\Big) && (4.1.62)\\
= & \operatorname{tr}\Big(\sum_{k=1}^{m} E_k L(\boldsymbol{s}) \rho(\boldsymbol{s}) L(\boldsymbol{s})\Big) && (4.1.63)\\
= & \operatorname{tr}\Big(\rho(\boldsymbol{s}) L(\boldsymbol{s})^2\Big) && (4.1.64)\\
= & \operatorname{tr}\left(\rho(\boldsymbol{s}) \sum_{i=1}^{n} w_i L_i(\boldsymbol{s}) \sum_{j=1}^{n} w_j L_j(\boldsymbol{s})\right) && (4.1.65)\\
= & \sum_{i=1}^{n}\sum_{j=1}^{n} w_i \operatorname{tr}\left(\rho(\boldsymbol{s}) L_i(\boldsymbol{s}) L_j(\boldsymbol{s})\right) w_j && (4.1.66)\\
= & \sum_{i=1}^{n}\sum_{j=1}^{n} w_i \frac{1}{2} \operatorname{tr}\left(\rho(\boldsymbol{s})\Big[L_i(\boldsymbol{s}) L_j(\boldsymbol{s}) + L_j(\boldsymbol{s}) L_i(\boldsymbol{s})\Big]\right) w_j && (4.1.67)\\
= & \sum_{i=1}^{n}\sum_{j=1}^{n} w_i \big[J(\boldsymbol{s})\big]_{ij} w_j && (4.1.68)\\
= & \boldsymbol{w}^\top J(\boldsymbol{s}) \boldsymbol{w}. && (4.1.69)
\end{aligned}$$

Finally, we obtain a superior limit for the classical Fisher Information I,

$$I \leq J, \tag{4.1.70}$$

where J is defined as the Quantum Fisher Information matrix,

$$J_{ij}(\boldsymbol{s}) = \frac{1}{2} \operatorname{tr}\left(\rho(\boldsymbol{s})\Big[L_i(\boldsymbol{s}) L_j(\boldsymbol{s}) + L_j(\boldsymbol{s}) L_i(\boldsymbol{s})\Big]\right). \tag{4.1.71}$$

The Quantum Crámer-Rao bound establishes a fundamental limit on the precision of parameter estimation in quantum systems. This bound is expressed as an inequality for the covariance matrix C of an unbiased estimator:

$$C \geq I^{-1} \geq J^{-1}, \tag{4.1.72}$$

where I is the classical Fisher information matrix, and J is the quantum Fisher

information matrix (QFIM). The fact that J^{-1} provides a tighter lower bound than I^{-1} suggests that Quantum Mechanics imposes a fundamental limit on estimation precision, which is stricter than its classical counterpart.

Unlike the classical Fisher information, the quantum Fisher information matrix $J(\boldsymbol{s})$ does not explicitly depend on the choice of measurement operators $\{E_i\}$. Instead, it is determined solely by the quantum state $\rho(\boldsymbol{s})$ and the parameters being estimated. Importantly, the inequality (4.1.72) holds only for unbiased estimators.

From this relation, we observe that Quantum Mechanics allows for an enhancement in estimation precision, surpassing what is achievable using only classical statistical methods. This realization has led to the development of *Quantum Metrology* [11–15], a field dedicated to optimizing measurement processes by exploiting uniquely quantum properties, such as entanglement and coherence. Quantum metrology has significant applications in fields like atomic clocks, gravitational wave detection, and quantum-enhanced imaging.

A fundamental question in quantum estimation theory is whether the quantum Crámer-Rao bound can actually be attained, and if so, what quantum measurement strategy enables achieving this limit. Identifying such a strategy is crucial, as it defines the optimal estimation technique that maximizes precision within the constraints imposed by Quantum Mechanics. However, the derivation of the bound itself does not indicate which measurement process, if any, saturates it. In fact, it has been established that attaining the quantum Crámer-Rao bound generally requires the use of collective measurements on multiple copies of the quantum state [16, 17].

A deeper connection between quantum estimation theory and quantum state distinguishability was established by Braunstein and Caves [7]. They demonstrated that the Symmetric Logarithmic Derivative (SLD) Fisher Information naturally induces a statistical distance between quantum states that are infinitesimally close, which corresponds to the *Bures distance* previosly introduced in 3.2.26 and 3.2.28 (D_B). This relationship is given by:

$$D_B^2(\rho(\boldsymbol{\theta}), \rho(\boldsymbol{\theta} + d\boldsymbol{\theta})) = \frac{1}{4} \sum_{ij} J(\boldsymbol{\theta})_{ij} \, d\theta_i d\theta_j, \tag{4.1.73}$$

where $J(\theta)$ represents the QFIM. Now, if the $d\boldsymbol{\theta}$ arises from an estimation, $d\theta_i d\theta_j = C(\boldsymbol{\theta})_{ij}$ is the covariance and

$$D_B^2(\rho(\boldsymbol{\theta}), \rho(\boldsymbol{\theta} + d\boldsymbol{\theta})) = \frac{1}{4}\sum_{ij} J(\boldsymbol{\theta})_{ij}\, C(\boldsymbol{\theta})_{ij} = \frac{1}{4}\,\mathrm{tr}(J(\boldsymbol{\theta})C(\boldsymbol{\theta})). \tag{4.1.74}$$

From this, we can get an expression for a weighted mean squared error,

Since the SLD Fisher Information dictates the precision limits of quantum estimation, this result establishes a crucial link between state distinguishability, statistical inference, and optimal measurement strategies.

4.1.3 N copies ensemble

To estimate the parameters $\boldsymbol{s}$ you must have N equally prepared copies of the state $\rho(\boldsymbol{s})$, that is,

$$\rho^{\otimes N}(\boldsymbol{s}) = \rho(\boldsymbol{s}) \otimes \rho(\boldsymbol{s}) \otimes \cdots \otimes \rho(\boldsymbol{s}). \tag{4.1.75}$$

From the N copies ensemble, information about the $\boldsymbol{s}$ parameters can be obtained through individual measures, the one that can be carried out sequentially on separate particles , or by globally measuring the collective system formed by all the copies.

The logarithmic symmetric derivatives $L_i^{(N)}(\boldsymbol{s})$ of $\rho^{\otimes N}(\boldsymbol{s})$ in this case are defined by

$$\frac{\partial \rho^{\otimes N}(\boldsymbol{s})}{\partial s_i} = \frac{1}{2}\Big(\rho^{\otimes N}(\boldsymbol{s}) L_i^{(N)}(\boldsymbol{s}) + L_i^{(N)}(\boldsymbol{s})\rho^{\otimes N}(\boldsymbol{s})\Big). \tag{4.1.76}$$

Using the product rule we can find an explicit form of this SLD's,

$$\frac{\partial \rho}{\partial s_i} \otimes \rho \otimes \cdots \otimes \rho + \cdots + \rho \otimes \cdots \otimes \rho \otimes \frac{\partial \rho}{\partial s_i} = \frac{1}{2}\Big(\rho^{\otimes N} L_i^{(N)} + L_i^{(N)} \rho^{\otimes N}\Big). \tag{4.1.77}$$

The last equation corresponds to the equation of L_i per subsystem,

$$L_i^{(N)} = L_i \otimes \mathbb{I} \otimes \cdots \otimes \mathbb{I} + \cdots + \mathbb{I} \otimes \cdots \otimes \mathbb{I} \otimes L_i. \tag{4.1.78}$$

Then, the Fisher quantum matrix has the following property

$$J_{ij}^{(N)} = \frac{1}{2}\operatorname{tr}\left(\rho^{\otimes N}\Big[L_i^{(N)}L_j^{(N)} + L_j^{(N)}L_i^{(N)}\Big]\right) \tag{4.1.79}$$

$$= \frac{1}{2}\operatorname{tr}\Big[\rho(L_iL_j + L_jL_i)\otimes\rho\otimes\cdots\otimes\rho + \cdots + \rho\otimes\cdots\otimes\rho\otimes\rho(L_iL_j + L_jL_i)\Big] \tag{4.1.80}$$

$$+ \operatorname{tr}\Big[\rho L_i\otimes\rho L_j\otimes\cdots\otimes\rho + \cdots + \rho\otimes\cdots\otimes\rho L_j\otimes\rho L_i\Big],$$

where there are N terms in the sum (4.1.81) in which there is a $\rho(L_iL_j + L_jL_i)$ per subsystem, and $N(N-1)$ terms in the sum (4.1.81) in which there is a ρL_i and a ρL_j in different subsystems. Now, using the composite system trace property, $\operatorname{tr}(A\otimes B) = \operatorname{tr}(A)\operatorname{tr}(B)$

$$J_{ij}^{(N)} = N\frac{1}{2}\operatorname{tr}\Big(\rho(L_iL_j + L_jL_i)\Big)\operatorname{tr}(\rho)^{N-1} + N(N-1)\operatorname{tr}\Big(\rho L_i\Big)\operatorname{tr}\Big(\rho L_j\Big)\operatorname{tr}(\rho)^{N(N-1)-2} \tag{4.1.81}$$

$$= N\frac{1}{2}\operatorname{tr}\Big(\rho(L_iL_j + L_jL_i)\Big) + N(N-1)\operatorname{tr}\Big(\rho L_i\Big)\operatorname{tr}\Big(\rho L_j\Big), \tag{4.1.82}$$

from the definition of logarithmic symmetric derivative,

$$\operatorname{tr}\Big(\frac{\partial\rho(\boldsymbol{s})}{\partial s_i}\Big) = \frac{1}{2}\operatorname{tr}\Big(\rho(\boldsymbol{s})L_i(\boldsymbol{s}) + L_i(\boldsymbol{s})\rho(\boldsymbol{s})\Big), \tag{4.1.83}$$

$$= \operatorname{tr}\Big(\rho(\boldsymbol{s})L_i(\boldsymbol{s})\Big), \tag{4.1.84}$$

$$= \frac{\partial\operatorname{tr}(\rho(s))}{\partial s_i}, \tag{4.1.85}$$

$$= 0. \tag{4.1.86}$$

With this result the quantity (4.1.82) is reduced to

$$J_{ij}^{(N)} = \frac{N}{2}\operatorname{tr}\left(\rho\Big[L_iL_j + L_jL_i\Big]\right), \tag{4.1.87}$$

$$= NJ_{ij}. \tag{4.1.88}$$

Let $C^{(N)}(\tilde{\boldsymbol{s}})$ be the covariance matrix and $I^{(N)}(\boldsymbol{s})$ the classical Fisher matrix for

an estimate with sample size N. The quantum Cramér-Rao bound becomes

$$C^{(N)} \geq \left[I^{(N)}\right]^{-1} \geq \frac{1}{N} J^{-1}, \tag{4.1.89}$$

where J corresponds to the Quantum Fisher Information matrix for a single copy.

4.1.4 Gill-Massar bound

Measurements on the ensemble members can be performed either individually or collectively. In the latter case, the measurements involve several ensemble members at once, which are often characterized by the use of quantum entanglement and because of this they are more difficult to perform experimentally than individual measurements. For this reason, individual measurements are generally preferred over collective ones.

Individual measurements, however, have a drawback. It is possible to show that with individual measurements, it is not possible to saturate the quantum Cramér-Rao bound, meaning that it is not possible to find a POVM whose classical Fisher information matrix equals the quantum Fisher information matrix. In this scenario, the known limit is the *Gill-Massar bound for infidelity.*

From the inequality 4.1.89 it follows that

$$I^{(N)} J^{-1} \leq N \mathbb{I}_{n\times n}, \tag{4.1.90}$$

where $I^{(N)}$ is the Classical Fisher Information matrix for a sample of size N , J^{-1} is the Quantum Fisher Information matrix and $\mathbb{I}_{n\times n}$ is the identity matrix of size $n \times n$, with n the number of parameters. Taking the trace of this we get,

$$\mathrm{tr}\left(I^{(N)} J^{-1}\right) \leq nN. \tag{4.1.91}$$

This means that the quantity $\mathrm{tr}\left(I^{(N)} J^{-1}\right)$ is bounded by at least N times the number of parameters. This inequality is saturated when the Fisher matrices are $I^{(N)} = NJ$, the same to saturate the Crámer-Rao bound. This inequality expresses how, in the quantum case, information about one parameter can be traded for information about another [18, 19].

We can obtain a different bound, more restrictive in the case when ρ is in a pure state or a mixed state. The following theorems [18] show this:

Theorem 6. *When measurements are performed on N identically prepared copies of a system in a pure state ρ, the classical Fisher matrix I and the quantum Fisher matrix J must satisfy the Gill-Massar bound*

$$Tr(I^{(N)}J^{-1}) \leq N(d-1). \tag{4.1.92}$$

Where, d is the dimension of the Hilbert space to which ρ belongs.

This inequality turns into an equality when the POVM being performed is *exhaustive*, i.e., whose elements are proportional to one dimensional projections. The next theorem is the Gill-Massar inequality for mixed states.

Theorem 7. *When separable measurements are performed on N identically prepared copies of a system in a mixed state ρ, the classical Fisher matrix I and the quantum Fisher matrix J must satisfy the Gill-Massar bound*

$$Tr(I^{(N)}J^{-1}) \leq N(d-1). \tag{4.1.93}$$

Where, d is the dimension of the Hilbert space to which ρ belongs.

In the mixed state inequality, the bound also is achieved by performing exhaustive measurements. From theorems 6 and 7 we can conclude that in general is impossible to construct an optimal measurement for all the parameters via separable measurements. The next theorem provides an optimal measurement protocol for mixed states.

Theorem 8. *In the case of mixed states, it is in general possible to devise a collective measurement for which the Fisher information does not satisfy the inequality 4.1.92.*

4.1.5 GM Bound for Weighted MSE

From 4.1.89 we can also derive a bound for weighted mean squared errors WC, W being a positive weight matrix, and C the covariance matrix. Then, the

Crámer-Rao bound gives us the following inequality

$$\mathrm{tr}\big(WC^{(N)}\big) \geq \mathrm{tr}\big(WI^{(N)-1}\big) \geq N\,\mathrm{tr}\big(WJ^{-1}\big). \tag{4.1.94}$$

The first bound is achieved performing MLE, the second is saturated when the SLD operator commutates. Aditionally, from the Gill-Massar bound 4.1.92 and the classical Crámer-Rao 4.1.89 bound we get the next inequality

$$\mathrm{tr}\big(J^{-1}C^{-1}\big) \leq \mathrm{tr}\big(J^{-1}I\big) \leq N(d-1), \tag{4.1.95}$$

this sets a lower bound for the scaled WMSE. Assuming that J,C and W are positive $n \times n$ matrices, now we can obtain the minimum value of $\mathrm{tr}(WC)$ under the constrain $\mathrm{tr}(J^{-1}C^{-1}) = N(d-1)$,

$$\mathrm{tr}(WC) = \mathrm{tr}\left(J^{-1/2}WJ^{-1/2}J^{1/2}CJ^{1/2}\right), \tag{4.1.96}$$

$$= \mathrm{tr}\left(J^{-1/2}WJ^{-1/2}J^{1/2}CJ^{1/2}\right)\frac{\mathrm{tr}\big(J^{-1/2}C^{-1}J^{-1/2}\big)}{N(d-1)}, \tag{4.1.97}$$

$$\geq (\mathrm{tr}\sqrt{J^{-1/2}WJ^{-1/2}})^2\frac{1}{N(d-1)}. \tag{4.1.98}$$

Where we used the Cauchy-Schwarz inequality for traces. Now, the minimum is given by

$$\mathrm{tr}(WC) = (\mathrm{tr}\sqrt{J^{-1/2}WJ^{-1/2}})^2\frac{1}{N(d-1)}. \tag{4.1.99}$$

And this is achieved when $C = I^{-1}$ and the covariance matrix is

$$C = N(d-1)J^{-1/2}\frac{\mathrm{tr}\sqrt{J^{-1/2}WJ^{-1/2}}}{\sqrt{J^{-1/2}WJ^{-1/2}}}J^{-1/2}. \tag{4.1.100}$$

The Gill-Massar bound for the Scaled Mean Squared Bures distance 4.1.73 derives from 4.1.99 with $W(\rho) = J(\rho)/4$,

$$\mathrm{tr}(WC) \geq (\mathrm{tr}\sqrt{J^{-1/2}WJ^{-1/2}})^2\frac{1}{N(d-1)}, \tag{4.1.101}$$

$$= (\mathrm{tr}\sqrt{J^{-1/2}\frac{J}{4}J^{-1/2}})^2\frac{1}{N(d-1)}, \tag{4.1.102}$$

$$= \frac{n^2}{4N(d-1)}. \tag{4.1.103}$$

with n the number of parameters being estimated. Finally, the ultimate bound depends on the kind of state being estimated. In the case of estimating a pure state which has $n = 2(d-1)$ parameters, the bound is given by

$$\mathrm{tr}(WC) \geq \frac{d-1}{N}. \tag{4.1.104}$$

Remembering that the Bures distance for close states can be written in terms of the infidelity, as shown in 3.2.28, we can find a lower bound for this quantity. The Gill-Massar bound establishes that the optimal estimation process of unknown pure quantum states through individual (or separable) measurements on the ensemble has a bounded average infidelity [18, 20] such that

$$\bar{\mathcal{I}}(|\psi\rangle) \geq \frac{d-1}{N} \quad \forall\ |\psi\rangle, \tag{4.1.105}$$

where d is the dimension of the Hilbert space of the states to be estimated, and N is the size of the ensemble. We can then use the average infidelity as a metric for the error in the estimation generated by a measurement.

Chapter 5

Quantum State Tomography

Quantum state tomography refers to a collection of statistical methods used to reconstruct the density matrix of an unknown quantum state based on experimental data [5, 21]. These methods are essential for validating and characterizing quantum systems [22, 23], particularly as quantum technologies advance toward more complex architectures.

The demand for efficient and precise quantum tomography techniques is driven by the continuous development of increasingly intricate quantum devices. Ensuring the correct functionality of quantum operations requires reliable statistical methods to process the inherently probabilistic outcomes of quantum measurements. Consequently, quantum tomography plays a crucial role in the verification and benchmarking of quantum devices like quantum sensors [24, 25].

At its core, quantum tomography consists of two fundamental components:

1. **A set of quantum measurements**, which extract information about the unknown state.
2. **An estimator**, which maps the measurement results to an estimated density matrix.

A fundamental question in quantum tomography is: *What is the optimal measurement strategy for a system of a given dimensionality?* The answer lies in the concept of *informationally complete* (IC) measurements. To uniquely determine an unknown quantum state, the set of measurements must be *tomographically complete*. This requires at least d^2 independent measurement outcomes in a system

of Hilbert space dimension d, accounting for the $d^2 - 1$ real parameters needed to specify a Hermitian density matrix with unit trace, while satisfying the constraint $\sum_i p_i = 1$.

Different tomography protocols—estimation strategies, based on a particular choice of an estimator and measurements, should be compared to each other and optimized for precision. The notion of optimality may, however, depend on the chosen figure of merit; i.e., strictly speaking, it may be different if the estimation precision is measured in terms of (in)fidelity or, for example, Hilbert–Schmidt distance. We will therefore focus on infidelity [7, 26] as a figure of merit for statistical reasons discussed above. As usual in estimation tasks, we are mostly concerned with the behavior of infidelity as a function of the dimension d and sample size N [8, 27–32], which is the overall number of measurement outcomes registered. From a more statistically rigorous point of view, it is reasonable to consider the average performance of tomographic protocols over all states.

Algorithms have been engineered to improve the statistical errors on a given sample of experimental data—the so-called adaptive tomography and then maximum likelihood strategies have been used that improved dramatically statistical errors.

5.1 Linear Inversion Tomography

Let us consider a quantum system of dimension d and a basis $\{B_j\}_{j=1,\ldots,d^2}$ of $d \times d$ Hermitian matrices. Analyzing the parameters needed to characterize a density matrix, we note that:

- Complex matrices have $2d^2$ independent real parameters,

$$\begin{pmatrix} a_{00} + ib_{00} & \ldots & \ldots \\ \ldots & \ldots & \ldots \\ \ldots & \ldots. & a_{dd} + ib_{dd} \end{pmatrix}.$$

- Hermitian matrices have d^2 independent real parameters,

$$\begin{pmatrix} a_{00} & a_{01} + ib_{01} & \ldots \\ a_{01} - ib_{01} & \ldots & \ldots \\ \ldots & \ldots. & a_{dd} \end{pmatrix}.$$

- The density matrix of a quantum system in dimension d has $n = d^2 - 1$ independent real parameters, as it must not only be Hermitian but also normalized, meaning its trace must equal one,

$$\sum_i a_{ii} = 1.$$

  In other words, $d^2 - 1$ measurements outcomes are needed to unequivocally reconstruct all quantum states. However, experimentally, d^2 measurements outcomes are required to normalize the results and obtain probabilities.

Let's consider a set of M_m measurements $\{\Pi_j\}$ that are projectors from a basis of POVM elements. We can write the density operator in this basis as

$$\rho = \sum_j S_j B_j, \quad S_j \in \mathbb{R}, \tag{5.1.1}$$

To perform state tomography of ρ with respect to this set of observables, we must determine the parameters S_j. By performing a measurement on $\{\Pi_j\}$, the expectation value of the measurement is given by

$$\langle \Pi_j \rangle = \mathrm{tr}(\rho \Pi_j), \tag{5.1.2}$$

$$= \mathrm{tr}\left[\left(\sum_i S_i B_i \right) \Pi_j \right], \tag{5.1.3}$$

$$= \sum_i S_i \, \mathrm{tr}(B_i \Pi_j). \tag{5.1.4}$$

Defining a column vector of the expectation values of each observable,

$$\vec{P} = \left(\langle \Pi_0 \rangle, ..., \langle \Pi_m \rangle \right)^\top, \tag{5.1.5}$$

the vector of the parameters S_i of the system's density matrix,

$$\vec{S} = (S_0, ..., S_n)^\top, \tag{5.1.6}$$

and a matrix of the traces of products between the observables Π_j and the basis

elements B_j,

$$M = \begin{pmatrix} \mathrm{tr}(\Pi_0 B_0) & ... & \mathrm{tr}(\Pi_0 B_n) \\ ... & ... & ... \\ \mathrm{tr}(\Pi_m B_0) & ... & \mathrm{tr}(\Pi_m B_n) \end{pmatrix}. \tag{5.1.7}$$

This matrix contains the projections of the basis elements B_j onto the basis of observables Π_i. We can rewrite the equation 5.1.4 for the expectation values as

$$\vec{P} = M\vec{S}. \tag{5.1.8}$$

Thus, to perform state tomography, i.e., to find the S_j, we must solve for the vector $\vec{S}$. If $n = m$, the number of observables is the same as the number of parameters and the matrix M is square, and the solution to the system of equations is

$$\vec{S} = M^{-1}\vec{P}. \tag{5.1.9}$$

If the number of observables is greater than the number of basis elements of the density matrix, $m > n$, there is redundant information, and we must solve the least-squares problem

$$\vec{S} = \arg\min_{\vec{r}} ||M\vec{r} - \vec{P}||^2. \tag{5.1.10}$$

To do this, we minimize $||M\vec{r} - \vec{P}||^2$ with respect to $\vec{r}$. First, we rewrite the expression to minimize,

$$||M\vec{r} - \vec{P}||^2 = \left| \sum_i \sum_j M_{ij} r_j \hat{i} - \sum_i P_i \hat{i} \right|^2, \tag{5.1.11}$$

$$= \left| \sum_i \left(\sum_j M_{ij} r_j - P_i \right) \hat{i} \right|^2, \tag{5.1.12}$$

$$= \sum_i \left(\sum_j M_{ij} r_j - P_i \right)^2. \tag{5.1.13}$$

Calculating the minimization,

$$\frac{\partial}{\partial r_k}||M\vec{r}-\vec{P}||^2 = \frac{\partial}{\partial r_k}\sum_i\left(\sum_j M_{ij}r_j - P_i\right)^2, \tag{5.1.14}$$

$$= 2\sum_i\left(\sum_j M_{ij}r_j - P_i\right)\left(\sum_j M_{ij}\delta_{kj}\right), \tag{5.1.15}$$

$$= 2\sum_i\left(\sum_j M_{ij}r_j - P_i\right)M_{ik}. \tag{5.1.16}$$

To find the $\vec{r}$ that minimizes the expression, we set the result above to zero, obtaining

$$\sum_i\sum_j M_{ij}M_{ik}r_j = \sum_i M_{ik}P_i, \tag{5.1.17}$$

$$M^\dagger M\vec{r} = M^\dagger\vec{P}. \tag{5.1.18}$$

Solving for the vector $\vec{r}$, the tomography solution is given by

$$\boxed{\vec{S} = (M^\dagger M)^{-1}M^\dagger\vec{P}.} \tag{5.1.19}$$

Here, $M^\dagger M$ is positive semi-definite, so it always has an inverse. Finally, the reconstruction of the density matrix is given by equation 5.1.1 with the parameters $\vec{S}$ obtained through minimization. If M is not invertible, or $m < n$, we say that the $\{\Pi_j\}$ are informationally incomplete for ρ.

5.1.1 Qubit Tomography

As shown in chapter 3.1.1.1, we can use the Pauli matrices σ_i to write quantum states of dimension 2, the qubit. In this representation, the density operator of a qubit is

$$\rho = \frac{1}{2}(\mathbb{I} + s_x\sigma_x + s_y\sigma_x + s_z\sigma_z)\,,\ s_i \in \mathbb{R} \tag{5.1.20}$$

where s_i are the parameters that determine the quantum state. For simplicity, we can rewrite this operator in the ket representation,

$$\rho = \frac{1}{2}\big[(1+s_z)\,|0\rangle\langle 0| + (1-s_z)\,|1\rangle\langle 1| + (s_x+is_y)\,|0\rangle\langle 1| + (s_x-is_y)\,|1\rangle\langle 0|\,\big]. \tag{5.1.21}$$

So, in order to perform the quantum tomography we need to estimate the s_x, s_y and s_z parameters.

In the case of a single qubit it suffices to perform three measurements with two outcomes each for the tomography [8], described by the following projectors:

$$M_{1\pm} = \frac{1}{2}(|0\rangle \pm |1\rangle)(\langle 0| \pm \langle 1|), \tag{5.1.22}$$

$$M_{2\pm} = \frac{1}{2}(|0\rangle \pm i\,|1\rangle)(\langle 0| \mp i\,\langle 1|), \tag{5.1.23}$$

$$M_{3+} = |0\rangle\langle 0| \ , \ M_{3-} = |1\rangle\langle 1|\,. \tag{5.1.24}$$

By measuring these projectors we obtain the probabilities of observing them. Following Born's rule,

$$p_{i\pm} = \mathrm{tr}(M_{i\pm}\rho). \tag{5.1.25}$$

For each projector this is,

$$p_{1\pm} = \mathrm{tr}(M_{1\pm}\rho), \tag{5.1.26}$$

$$= \frac{1}{4}(\langle 0| \pm \langle 1|)\big((1+s_z)\,|0\rangle\langle 0| + (1-s_z)\,|1\rangle\langle 1| \tag{5.1.27}$$

$$+ (s_x + is_y)\,|0\rangle\langle 1| + (s_x - is_y)\,|1\rangle\langle 0|\,\big)(|0\rangle \pm |1\rangle),$$

$$= \frac{1}{2}(1 \pm s_x). \tag{5.1.28}$$

$$p_{2\pm} = \mathrm{tr}(M_{2\pm}\rho), \tag{5.1.29}$$

$$= \frac{1}{4}(\langle 0| \pm i\,\langle 1|)\big((1+s_z)\,|0\rangle\langle 0| + (1-s_z)\,|1\rangle\langle 1| \tag{5.1.30}$$

$$+ (s_x + is_y)\,|0\rangle\langle 1| + (s_x - is_y)\,|1\rangle\langle 0|\,\big)(|0\rangle \mp i\,|1\rangle),$$

$$= \frac{1}{2}(1 \pm s_y). \tag{5.1.31}$$

$$p_{3+} = \mathrm{tr}(M_{3+}\rho) = \frac{1}{2}(1+s_z). \tag{5.1.32}$$

$$p_{3-} = \mathrm{tr}(M_{3-}\rho) = \frac{1}{2}(1-s_z). \tag{5.1.33}$$

With this, we can obtain the unknown parameters as,

$$s_i = p_{i+} - p_{i-}. \tag{5.1.34}$$

Experimentally, the corresponding frequencies of detection $n_{i\pm}$ estimate the real probabilities $p_{i\pm} = n_{i\pm}/(n_{i+} + n_{i-})$, and consequently the unknown parameters.

The linear estimation model is:

$$p_i = \mathrm{tr}(\rho\Pi_i) \quad \rightarrow \quad \text{estimator} \quad \hat{p_i} = \frac{n_i}{N}, \tag{5.1.35}$$

where n_i denotes the number of detections associated with the measurement operator Π_i, and N is the total number of measurements. Despite its conceptual simplicity, linear inversion tomography has a significant drawback: the estimated density matrix may not always be a valid quantum state.

The associated estimation error ε, scales as (citar algun libro):

$$\varepsilon = O\left(\frac{1}{\sqrt{N}}\right). \tag{5.1.36}$$

This implies that reducing the error requires increasing the sample size. However, in practical scenarios, the number of measurements is always finite, meaning that some level of error is unavoidable. Including the statistical fluctuations due to noise, deviations from the expected probabilities is expected. As a result, the reconstructed density matrix may violate key physical constraints, such as Hermiticity, unit trace, or positive semi-definiteness. Furthermore, as the Hilbert space dimension d increases, the likelihood of obtaining an estimated density matrix that is not positive semi-definite also increases.

To address this issue, *Maximum Likelihood Estimation* (MLE) is employed. This statistical inference method refines the reconstruction by ensuring that the estimated density matrix remains a physically valid quantum state, best fitting the experimental data within the constraints imposed by Quantum Mechanics.

5.2 Maximum Likelihood Estimation

The *Maximum Likelihood Estimator* (MLE), previously introduced in subsection 2.2.6 is a widely used statistical inference method that provides an efficient estimation strategy, as it asymptotically saturates the *Cramér-Rao Bound* [33, 34]. Originally introduced by Fisher in 1920, MLE is based on the principle of selecting the parameter values that maximize the *likelihood function*, which represents the

probability of obtaining the observed data given a specific model.

The likelihood function quantifies how well a particular set of parameters explains the measured data. It depends on both the observed dataset and the parameters that define the underlying probability distribution or model. Formally, if the model parameters are represented by a vector $\boldsymbol{\theta}$ and the observed data by a vector $\boldsymbol{x}$, the likelihood function is given by the joint probability:

$$L(\boldsymbol{x}|\boldsymbol{\theta}) = f(x_1|\boldsymbol{\theta}) \dots f(x_n|\boldsymbol{\theta}), \tag{5.2.1}$$

and it describes the probability distribution governing the observations.

The MLE estimate, $\hat{\theta}_{MLE}$, corresponds to the parameter values that maximize the likelihood function. In practice, it is often more convenient to maximize the *log-likelihood function*, $\log L(\theta)$, rather than the likelihood function itself. This is because the likelihood function can be highly peaked or numerically difficult to handle, whereas the logarithm transforms it into a more manageable form while preserving the location of the maximum:

$$\hat{\theta}_{MLE} = \arg\max_{\theta}(\log L(\theta)). \tag{5.2.2}$$

In the case of quantum tomography, the quantum state we wish to estimate plays the role of the model parameters. Data are obtained through an experiment depending on the POVM $\{E_k\}$ implemented, which is assumed to be informationally complete. Then, the probability p_k of obtaining result k is given by $p_k = \mathrm{tr}(E_k\rho)$. Since the experiment is conducted with an ensemble of N identically and independently prepared systems, a random sequence of results will be generated, characterized by frequencies n_k and in one-to-one relation with the elements of the POVM. Clearly

$$N = \sum_k n_k \quad \text{and} \quad \frac{n_k}{N} \approx p_k. \tag{5.2.3}$$

In this case, the likelihood function is given by [33, 35, 36]

$$L(\rho) = \prod_k \mathrm{tr}(\rho E_k)^{n_k}, \tag{5.2.4}$$

Since, its preferable to work with the log-likelihood, we have

$$\log(L(\rho)) = \sum_k n_k tr(\rho E_k). \tag{5.2.5}$$

Finally, the state estimate ρ is obtained by maximizing $L(\rho)$ (or $\log L(\rho)$) over the space of semi-positive operators with unit trace, that is,

$$\hat{\rho}_{MLE} = \arg\left(\max_{\rho\geq 0, \mathrm{tr}(\rho)=1} \log L(\rho)\right). \tag{5.2.6}$$

To ensure these conditions we consider the Cholesky decomposition [37] of the density matrix:

$$\rho(\vec{t}) = \frac{T^\dagger(\vec{t})T(\vec{t})}{\mathrm{tr}\big(T^\dagger(\vec{t})T(\vec{t})\big)}, \tag{5.2.7}$$

where T is a complex lower triangular matrix. The numerator in the decomposition ensures positivity, and the denominator ensures normalization. The form of the matrix T is:

$$T(\vec{t}) = \begin{pmatrix} t_1 & 0 & 0 & ... & 0 \\ t_d + it_{d+2} & t_2 & 0 & ... & 0 \\ ... & ... & ... & ... & 0 \\ t_{d^2-1} + it_{d^2} & t_{d^2-3} + it_{d^2-2} & ... & ... & t_d \end{pmatrix}. \tag{5.2.8}$$

In this way, the optimization 5.2.6 is performed over the vector $\vec{t} = (t_1, ..., t_{d^2})$ with $t_i \in \mathbb{R}$. Some algorithms used to perform the optimizations utilize specific parametrization to incorporate the quantum constraints [38, 39] but the convergence of the likelihood is slow. To solve this, *superfast maximum likelihood* [34] utilize an accelerated projected-gradient method that ensures the quantum nature of the estimation and a fast convergence.

5.3 Pure Quantum State Estimation

The tomography methods discussed earlier aim to estimate the density matrix, initially unknown, of a quantum system. These methods are general; however, they require a large number of measurements. Standard quantum tomography (SQT) [35] requires on the order of d^2 measurements, and for multi-qubit systems, 4^n with n the number of quits. In other cases like Symmetrically Informationally

Complete POVM (SIC-POVM) [20, 40–45] or MLE [38], it is simply difficult to implement them, or the method does not exist for that dimension such as Mutually Unbiased Bases (MUBs) [46–52].

A possible solution to this problem is to employ a priori information about the unknown state such as it's purity [53–55]. In this way, the amount of unknown parameters to estimate and the number of measurement to perform reduces from the order d^2 to d. This is pure state quantum tomography.

Pure states take the form:

$$|\Psi\rangle = \sum_{k=0}^{d-1} (a_k + ib_k) |k\rangle \,, \; a_k, b_k \in \mathbb{C}. \tag{5.3.1}$$

Considering the constraints of normalization and global phase, only $2(d-1)$ parameters are needed to characterize this state. This means the number of measurements grows linearly with the dimension, thereby reducing the number of measurements required as well as the amount of data post-processing.

When performing measurements on the basis, we obtain the probabilities:

$$P_k = \mathrm{tr}(|k\rangle \langle k| \, \rho) = |\langle k|\Psi\rangle|^2 = |a_k + ib_k|^2. \tag{5.3.2}$$

Thus, to find the parameters that describe the unknown pure state we must solve a system of quadratic equations. There exists different method for performing measures and the estimation for pure states as shown in [56, 57].

However, it is difficult to generate pure quantum states. Since they are not completely isolated and decoherence occurs. Despite this, it is possible to experimentally create high-purity states, such as those used in quantum technologies.

5.4 5 Bases Based Quantum Tomography

A longstanding problem in Quantum Mechanics is determining the minimum number of observables required for the complete characterization of an unknown pure quantum state. In [54], it was demonstrated that the full reconstruction of an unknown pure quantum state can be achieved using only 5 measurement bases.

For a pure d-dimensional state given by (5.3.1), we consider an arbitrary basis $\mathcal{B}_0 = \{|i\rangle\}$ of the Hilbert space, where $i = 0, \ldots, d-1$. Relative to this basis, the additional four measurement bases are defined as:

$$\mathcal{B}_1 = \left\{ \left|\varphi_{\pm}^{\nu}\right\rangle_1 = \frac{1}{\sqrt{2}}(|2\nu\rangle \pm |2\nu+1\rangle) \right\}, \tag{5.4.1}$$

$$\mathcal{B}_2 = \left\{ \left|\tilde{\varphi}_{\pm}^{\nu}\right\rangle_2 = \frac{1}{\sqrt{2}}(|2\nu\rangle \pm i\,|2\nu+1\rangle) \right\}, \tag{5.4.2}$$

$$\mathcal{B}_3 = \left\{ \left|\varphi_{\pm}^{\nu}\right\rangle_3 = \frac{1}{\sqrt{2}}(|2\nu+1\rangle \pm |2\nu+2\rangle) \right\}, \tag{5.4.3}$$

$$\mathcal{B}_4 = \left\{ \left|\tilde{\varphi}_{\pm}^{\nu}\right\rangle_4 = \frac{1}{\sqrt{2}}(|2\nu+1\rangle \pm i\,|2\nu+2\rangle) \right\}, \tag{5.4.4}$$

where $\nu \in [0, (d-2)/2]$. For odd-dimensional systems, the integer part of $(d-2)/2$ is considered, and each basis is completed with the state $|d\rangle$. The probabilities associated with measuring each element of these bases define a system of equations that uniquely determines the unknown quantum state, as demonstrated in [54]. Notably, the measurement basis $\mathcal{B}_0$ is only used to define the structure of the other measurement bases $\mathcal{B}_i$ for $i = 1, \ldots, 4$, ensuring the feasibility of state reconstruction.

In [55], the 5BB-QT method was refined to enhance estimation accuracy, measured in terms of the infidelity $\mathcal{I}$. The modified approach utilizes measurements from basis $\mathcal{B}_0$ to estimate the amplitudes $|a_k + ib_k|$ and to optimize the adaptation of the remaining bases. Additionally, a permutation operator U_p was introduced to reduce errors in phase estimation. This improved approach was named the *Improved Corrected* 5BB-QT (IC5BB-QT). The average infidelity of the estimation was found to scale as:

$$\bar{\mathcal{I}}(|\psi_i\rangle) \approx 1.76\, \frac{d^{1.87}}{N}, \tag{5.4.5}$$

where the prefactor $\bar{\alpha} = 1.76$ represents the average value obtained from simulations in dimensions $d = 4, 8, 16, 32, 64$.

5.5 Adaptive Quantum Tomography

An adaptive strategy for quantum state estimation refers to a method in which the measurement settings or protocols are dynamically adjusted based on the outcomes of prior measurements [8]. Adaptive strategies aim to improve the

efficiency and accuracy of quantum state estimation by focusing measurements on parts of the state space that are most informative. They have been proven to be particularly valuable when resources are limited or when the quantum system has high dimensionality, as they reduce the number of measurements required to achieve a given level of accuracy [29, 58]. This contrasts with static strategies, where the measurement settings are predetermined and do not change during the process.

The simplest protocol consists in performing a two-stage standard tomography [29, 59, 60]. In the first stage, a low accuracy estimate $\hat{\rho}_0$ of ρ is obtained performing standard tomography on a sample of size $N_0 < N$. The basis of eigenvectors of this estimate is then used to adapt the measurement writing the observables σ_i in the basis of eigenvector of $\hat{\rho}_0$ for a second, higher accuracy tomography on the remaining sample of size $N - N_0$. This process is shown schematically in the figure 5.5.1.

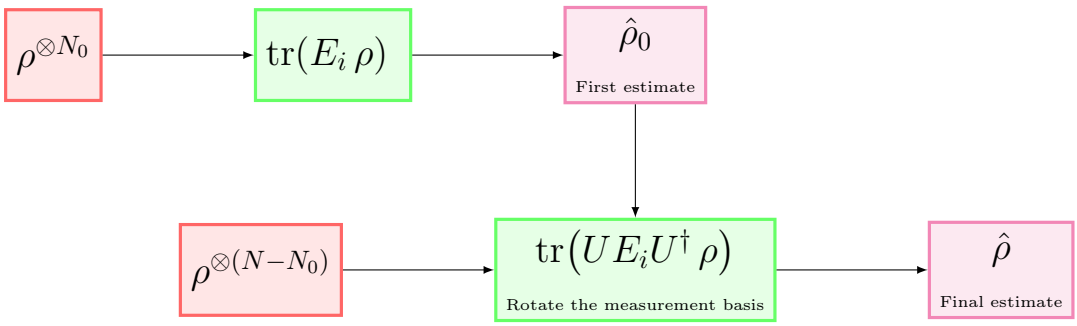

Figure 5.5.1: General diagram for two-stage adaptive quantum state tomography.

Other, more sophisticated approaches include *self-learning* (or *self-guided*) tomographic methods [61], which iteratively optimize a utility function at each step of the reconstruction process as shown in figure 5.5.2. These methods estimate the gradient of the utility function by performing two measurements at each iteration. While they have demonstrated superior performance compared to non-adaptive tomography, they are also characterized by a high computational cost. Experimental implementations on trapped ions [62] and photons [58] have shown that self-learning tomography can outperform traditional, non-adaptive techniques.

A class of self-learning adaptive tomographic methods aims to optimize the *information gain* during the estimation process [63]. However, their

performance in terms of *average infidelity*—a crucial figure of merit in quantum estimation—remains unclear. Consequently, methods that directly optimize the fidelity between the estimated and true quantum state are often preferred [25, 64–66].

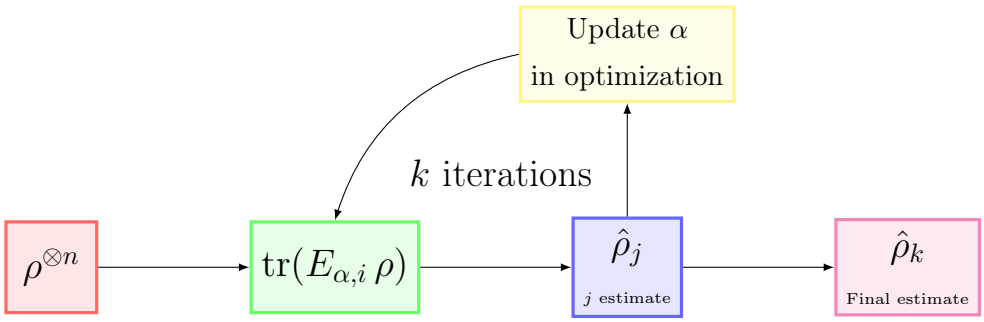

Figure 5.5.2: General diagram for self-learning tomography.

Due to the significant computational cost associated with fidelity-based optimization, some approaches incorporate *Bayesian inference* to enhance efficiency and improve estimation accuracy [67, 68]. Bayesian methods allow for adaptive updates of prior knowledge about the quantum state, leading to more efficient data acquisition and state reconstruction.

5.5.1 Qubit Two-stage Adaptive Tomography

Considering the 3-measurement tomography 5.1.1, we can build a two-step adaptive tomography for qubits. Considering an ensemble of N copies of the unknown state:

1. Perform the standard tomography on a fraction N_0 of the N qubits and obtain a preliminary estimate $\hat{\rho}_1$ for ρ.

2. Rotate the measurement basis, such that it coincides with the eigenbasis of $\hat{\rho}_1$, and perform the standard tomography on the remaining $N - N_0$ copies. Finally, we get the estimate $\hat{\rho}_2$.

This strategy was first mentioned by Gill and Massar in [18] later refined by Řeháček et al. [69]. This scheme has been experimentally proved to perform better than non-adaptive stategies for qubits [29]. Later this adaptive strategy was generalized for qudits in [30].

Determining the optimal choice of $N_0 = \beta N$ is not straightforward. Numerical simulations in [29] suggest that the optimal value is $\beta = 1/2$, which yields the best performance in practical implementations.

Chapter 6

Fisher Symmetric Measurements

A Fisher-symmetric POVM [70] is a POVM that saturates the Gill-Massar bound. So, is a set of rank-one elements constructed such that the Fisher information is proportional to the identity, meaning that all parameters are estimated with the same quality relative to the quantum limit for estimating each separately, thus symmetric. But, this kind of measurement are constructed knowing that the unknown state is close to a *fiducial state*, so there is apriori information. Consequently, this measures can determine a quantum state in a local neighborhood of a fiducial state, meaning that they are *locally informationaly complete*. Local estimates can be obtained from locally optimal measurements [71, 72] or as a result of a global estimation strategy [64, 65, 68, 73, 74]. In [70] they present a Fisher-symmetric POVM for pure states, calling this set a *pure Fisher-symmetric informationally complete* (PFSIC) POVM. Since then, there has been an effort to expand this measurements to other problems, like tomography via collective measurement using Fisher Symmetric measurements [75].

6.1 Fisher Symmetric Measurements

Let us consider an unknown quantum pure state $|\chi(x)\rangle$ near a known fiducial pure state $|0\rangle$. This state can be written as

$$|\chi(x)\rangle = |0\rangle + \sum_{j=1}^{d-1}(x^{j0} + ix^{j1})\,|j\rangle\,, \tag{6.1.1}$$

with $|x^{k\sigma}| \ll 1$ infinitesimal real parameters. Writing the state $|\chi\rangle$ as a density matrix and retaining only the linear order elements in the parameters, we get

$$\rho(x) = |0\rangle\langle 0| + \sum_{j\sigma} x^{j\sigma} X_{j\sigma}, \tag{6.1.2}$$

where

$$X_{j\sigma} = (-i)^{\sigma}(|0\rangle\langle j| + (-1)^{\sigma}|j\rangle\langle 0|). \tag{6.1.3}$$

These Hermitian operators satisfy the orthogonality condition $\mathrm{tr}(X_{j\sigma}X_{k\tau}) = 2\delta_{jk}\delta_{\sigma\tau}$.

A Positive Operator Valued Measurement (POVM) is a Fisher Symmetric Measurement (FSM) if it saturates the Gill-Massar Inequality

$$\mathrm{tr}\big(J^{-1}I\big) \leq d - 1, \tag{6.1.4}$$

and its Classical Fisher Information Matrix I is proportional to the Quantum Fisher Information Matrix J. Both conditions are fulfilled when the Classical FIM C is related with the Quantum FIM Q by

$$I = \frac{1}{2}J, \tag{6.1.5}$$

in the case of pure states. In the following sections, we calculate both Quantum and Classical Fisher Information Matrices for the state (6.1.1) and find the conditions of a POVM to be a FSM.

6.1.1 Quantum Fisher Information

The Quantum FIM of the state ρ is a symmetric matrix J with elements

$$J_{j\sigma,k\tau} = \frac{1}{2}\,\mathrm{tr}[\rho(L_{j\sigma}L_{k\tau} + L_{k\tau}L_{j\sigma})], \tag{6.1.6}$$

where the $\{L_{j\sigma}\}$ are the symmetric logarithmic derivative operators, implicitly determined by the equation

$$\frac{\partial\rho}{\partial x^{j\sigma}} = \frac{1}{2}(L_{j\sigma}\rho + \rho L_{j\sigma}). \tag{6.1.7}$$

When the state is pure, that is $\rho(x) = |\psi\rangle\langle\psi|$, the Quantum Fisher Information Matrix is simply given by [10]

$$J_{j\sigma,k\tau} = 4\mathrm{Re}\left[\langle\partial_{j\sigma}\psi|\partial_{k\tau}\psi\rangle - \langle\psi|\partial_{j\sigma}\rangle\langle\partial_{k\tau}\psi|\psi\rangle\right], \tag{6.1.8}$$

where

$$|\partial_{j\sigma}\psi\rangle = \frac{\partial}{\partial x^{j\sigma}}|\psi\rangle. \tag{6.1.9}$$

Thereby, the Quantum Fisher Information of the state $|\chi(x)\rangle$ is

$$J_{j\sigma,k\tau} = 2\,\mathrm{tr}(X_{j\sigma}X_{k\tau}) = 4\delta_{j\sigma,k\tau}, \tag{6.1.10}$$

or equivalently $J(\rho) = 4\mathbb{I}_{2d-2}$.

6.1.2 Classical Fisher Information

Let us consider an arbitrary POVM with n rank-1 elements, $\mathcal{E} = \{E^\alpha = |\varphi^\alpha\rangle\langle\varphi^\alpha|\}$, where

$$|\varphi^\alpha\rangle = \sum_{k=0}^{d-1}(\beta_k^\alpha + i\gamma_k^\alpha)\,|k\rangle\,. \tag{6.1.11}$$

with $\{\beta_k^\alpha\}$ and $\{\gamma_k^\alpha\}$ are real coefficients. Considering the completeness condition, $\sum_\alpha E^\alpha = \mathbb{I}$, we have

$$\sum_\alpha E^\alpha = \sum_\alpha\sum_{k,j}(\beta_k^\alpha + i\gamma_k^\alpha)(\beta_j^\alpha - i\gamma_k^\alpha)\,|k\rangle\langle j| \tag{6.1.12}$$

$$= \sum_\alpha\sum_{k,j}\left(\beta_k^\alpha\beta_j^\alpha + \gamma_k^\alpha\gamma_j^\alpha + i(\beta_j^\alpha\gamma_k^\alpha - \beta_k^\alpha\gamma_j^\alpha)\right)|k\rangle\langle j| \equiv \sum_{k,j}\delta_{k,j}\,|k\rangle\langle j|\,. \tag{6.1.13}$$

Defining the n-dimensional real vector $\vec{\beta}_k = (\beta_k^0, \ldots, \beta_k^n)$ and $\vec{\gamma}_k = (\gamma_k^0, \ldots, \gamma_k^n)$, the completeness condition becomes

$$\begin{aligned}
&\vec{\beta}_k\cdot\vec{\beta}_j + \vec{\gamma}_k\cdot\vec{\gamma}_j = \delta_{k,j}, \quad j,k = 0,\ldots,d-1,\\
&\vec{\gamma}_k\cdot\vec{\beta}_j - \vec{\beta}_k\cdot\vec{\gamma}_j = 0, \quad j,k = 0,\ldots,d-1.
\end{aligned} \tag{6.1.14}$$

Since we can choose the global phase of each state $|\varphi^\alpha\rangle$, we can set the constant accompanying the vector $|0\rangle$ to be real, i.e., $\gamma_0 = 0$,

$$|\varphi^\alpha\rangle = \beta_0^\alpha |0\rangle + \sum_{k=1}^{d-1} (\beta_k^\alpha + i\gamma_k^\alpha) |k\rangle . \tag{6.1.15}$$

With this we also have the following conditions for β_0,

$$\begin{aligned} &\vec{\beta}_0 \cdot \vec{\beta}_0 = 1 \\ &\vec{\beta}_0 \cdot \vec{\beta}_j = \vec{\beta}_0 \cdot \vec{\gamma}_j = 0, \quad j = 1, ..., d-1. \end{aligned} \tag{6.1.16}$$

The classical FIM of the state $|\chi\rangle$ when is measured by $\mathcal{E}$ is

$$I_{j\sigma,k\tau} = \sum_\alpha \frac{1}{p(\alpha|x)} \frac{\partial p(\alpha|x)}{\partial x^{j\sigma}} \frac{\partial p(\alpha|x)}{\partial x^{k\tau}}, \tag{6.1.17}$$

where $p(\alpha|x)$ is the probability distribution. Calculating the probabilities for the POVM $\mathcal{E}$, we obtain

$$\begin{aligned} p(\alpha|x) &= \mathrm{tr}(E^\alpha \rho(x)), \\ &= \mathrm{tr}\left[E^\alpha \left(|0\rangle\langle 0| + \sum_{j\sigma} x^{j\sigma} X_{j\sigma} \right) \right], \qquad (6.1.18) \\ &= \langle 0| E^\alpha |0\rangle + \sum_{j\sigma} x^{j\sigma} (-i)^\sigma [\langle j| E^\alpha |0\rangle + (-1)^\sigma \langle 0| E^\alpha |j\rangle]. \qquad (6.1.19) \end{aligned}$$

Calculating the matrix elements of each operator E^α,

$$\langle 0| E^\alpha |0\rangle = (\beta_0^\alpha + i\gamma_0^\alpha)(\beta_0^\alpha - i\gamma_0^\alpha) = (\beta_0^\alpha)^2, \tag{6.1.20}$$

$$\langle j| E^\alpha |0\rangle = (\beta_j^\alpha + i\gamma_j^\alpha)(\beta_0^\alpha - i\gamma_0^\alpha) = \beta_0^\alpha(\beta_j^\alpha + i\gamma_j^\alpha), \tag{6.1.21}$$

$$\langle 0| E^\alpha |j\rangle = (\beta_0^\alpha + i\gamma_0^\alpha)(\beta_j^\alpha - i\gamma_j^\alpha) = \beta_0^\alpha(\beta_j^\alpha - i\gamma_j^\alpha). \tag{6.1.22}$$

we can rewrite the probabilities as

$$p(\alpha|x) = (\beta_0^\alpha)^2 + \beta_0^\alpha \sum_j \left(x^{j0} \beta_j^\alpha + x^{j1} \gamma_j^\alpha \right) \tag{6.1.23}$$

Thereby, the derivatives of the probabilities are

$$\frac{\partial p(\alpha|x)}{\partial x^{j0}} = 2\beta_0^\alpha \beta_j^\alpha \tag{6.1.24}$$

$$\frac{\partial p(\alpha|x)}{\partial x^{j1}} = 2\beta_0^\alpha \gamma_j^\alpha \tag{6.1.25}$$

$$\tag{6.1.26}$$

Replacing $p(\alpha|x)$ and $\partial p(\alpha|x)/\partial x^{j\sigma}$ into Eq.(6.1.17) and approximating to first order , we obtain the Classical FIM

$$\begin{aligned} I_{j0,k0} &= 4\vec{\beta}_j \cdot \vec{\beta}_k, \\ I_{j1,k1} &= 4\vec{\gamma}_j \cdot \vec{\gamma}_k, \quad j,k = 1,...,d-1. \\ I_{j0,k1} &= 4\vec{\beta}_j \cdot \vec{\gamma}_k, \end{aligned} \tag{6.1.27}$$

Imposing the relation between the FIMs $I = J/2$, we obtain the conditions

$$2\vec{\beta}_j \cdot \vec{\beta}_k = \delta_{jk}, \tag{6.1.28}$$

$$2\vec{\gamma}_j \cdot \vec{\gamma}_k = \delta_{jk}, \tag{6.1.29}$$

$$2\vec{\beta}_j \cdot \vec{\gamma}_k = 0. \tag{6.1.30}$$

6.1.3 FSM conditions

We can summarize the conditions on the coefficients $\{\beta_j^\alpha\}$ and $\{\gamma_j^\alpha\}$ to E^α being a FSM as

$$\begin{aligned} &\vec{\beta}_0 \cdot \vec{\beta}_0 = 1, \\ &\vec{\beta}_0 \cdot \vec{\beta}_j = \vec{\beta}_0 \cdot \vec{\gamma}_j = \vec{\beta}_j \cdot \vec{\gamma}_k = 0, \quad j,k = 1,\ldots,d-1, \\ &\vec{\beta}_j \cdot \vec{\beta}_k = \vec{\gamma}_j \cdot \vec{\gamma}_k = \frac{1}{2}\delta_{jk}. \end{aligned} \tag{6.1.31}$$

Thereby, the set of $2d-1$ real vectors $\{\beta_0, \sqrt{2}\beta_k, \sqrt{2}\gamma_k\}_{k=1}^{d-1}$ form an orthonormal set. Consequently, the number of elements of the POVM has to satisfy $n \geq 2d-1$ since the dimension of the vectors has to be equal or larger to the number of states in order to have an orthonormal set. The conditions in Eq. (6.1.14) and (6.1.31) can be written in a more compact way considering the complex coefficients

$\omega_k^\alpha = \beta_k^\alpha + i\gamma_k^\alpha$ as

$$\sum_{\alpha=1}^{n} \omega_k^\alpha \omega_j^\alpha = 0, \quad j,k = 1,\ldots,d-1, \tag{6.1.32}$$

$$\sum_{\alpha=1}^{n} \omega_k^\alpha \omega_j^{\alpha *} = \delta_{jk}, \quad j,k = 0,\ldots,d-1. \tag{6.1.33}$$

The set of measurements defined by the conditions Eqs. (6.1.32) and (6.1.33) is sufficient to estimate simultaneously all parameters of a quantum state $|\chi\rangle$ close to the fiducial state $|0\rangle$ with the ultimate precision in infidelity given by the Gill-Massar bound

$$\bar{\mathcal{I}} = \frac{d-1}{N}, \tag{6.1.34}$$

with N the sample size used in the measurement of $\mathcal{E}$.

6.2 Explicit Construction of Fisher Symmetric Measurements

Following the explicit construction given by C. Caves et al [70], we are going to find a set of orthonormal vectors $\{u_\xi\}_{\xi=0,\ldots,2d-2} = \{\boldsymbol{b_0}, \sqrt{2}\boldsymbol{b_j}, \sqrt{2}\boldsymbol{c_j}\}_{j=1,\ldots,d-1}$ in $n = 2d-1$ dimensions that characterize a minimal FSM.

Considering the basis $\{\boldsymbol{e_\xi}\}_{\xi=0,\ldots,2d-2}$ where $\boldsymbol{e_\xi}$ has a 1 in the ξ-th position en zeroes elsewhere. Since b_0 has all positive components, and to take advantage of the symmetry of the PFSIC, we can choose

$$\boldsymbol{b_0} = \frac{1}{\sqrt{n}} \sum_{\xi=0}^{2d-2} \boldsymbol{e_\xi} \quad \Longleftrightarrow \quad b_0^\xi = \frac{1}{\sqrt{n}}, \ \xi = 0,\ldots,2d-2. \tag{6.2.1}$$

This choice ensures $\boldsymbol{b_0} \cdot \boldsymbol{b_0} = 1$. Now, the condition for the other vectors is $\boldsymbol{b_0} \cdot \boldsymbol{b_j} = 0$ and $\boldsymbol{b_0} \cdot \boldsymbol{c_j} = 0$, so we need to construct a set of $2d-2$ vectors orthogonal to $\boldsymbol{b_0}$. We can do this from the basis $\{\boldsymbol{e_\xi}\}$, subtracting the projection of each element with $\boldsymbol{b_0}$,

$$\boldsymbol{v_\xi} = \boldsymbol{e_\xi} - \boldsymbol{b_0}(\boldsymbol{b_0} \cdot \boldsymbol{e_\xi}) = \boldsymbol{e_\xi} - \frac{1}{\sqrt{n}}\boldsymbol{b_0}, \quad \xi = 0,\ldots,2d-2. \tag{6.2.2}$$

There is one extra element for it to be an orthonormal set, we can see that from

$$\boldsymbol{v_\xi} \cdot \boldsymbol{v_\nu} = \left(\boldsymbol{e_\xi} - \frac{1}{\sqrt{n}}\boldsymbol{b_0}\right) \cdot \left(\boldsymbol{e_\nu} - \frac{1}{\sqrt{n}}\boldsymbol{b_0}\right), \tag{6.2.3}$$

$$= \boldsymbol{e_\xi} \cdot \boldsymbol{e_\nu} + \frac{1}{n} - 2\frac{1}{n}, \tag{6.2.4}$$

$$= \delta_{\xi\nu} - \frac{1}{n}. \tag{6.2.5}$$

We can choose one of the vectors and subtract a multiple of it from the other elements, we chose $\boldsymbol{v_0} = \boldsymbol{e_0} - \frac{1}{\sqrt{n}}\boldsymbol{b_0}$. Now, the elements are of the form $\boldsymbol{u_\xi} = \boldsymbol{v_\xi} - k\,\boldsymbol{v_0}$ with $k \in \mathcal{R}$.

Then, the product between two elements is

$$\boldsymbol{u_\xi} \cdot \boldsymbol{u_\nu} = (\boldsymbol{v_\xi} - k\,\boldsymbol{v_0}) \cdot (\boldsymbol{v_\nu} - k\,\boldsymbol{v_0}), \tag{6.2.6}$$

$$= \boldsymbol{v_\xi} \cdot \boldsymbol{v_\nu} - k\boldsymbol{v_0} \cdot \boldsymbol{v_\nu} - k\boldsymbol{v_\xi} \cdot \boldsymbol{v_0} + k^2\boldsymbol{v_0} \cdot \boldsymbol{v_0}, \tag{6.2.7}$$

$$= \delta_{\xi\nu} - \frac{1}{n} - k(\delta_{0\nu} - \frac{1}{n}) - k(\delta_{0\xi} - \frac{1}{n}) + k^2(1 - \frac{1}{n}), \tag{6.2.8}$$

$$= \delta_{\xi\nu} - \frac{1}{n} + 2k\frac{1}{n} + k^2(1 - \frac{1}{n}). \tag{6.2.9}$$

We have that, for this elements to be orthogonal

$$k^2(1 - \frac{1}{n}) + 2k\frac{1}{n} - \frac{1}{n} = 0. \tag{6.2.10}$$

So $k = 1/(\sqrt{n}+1)$ in order for $\{u_\xi\}$ to be an orthonormal set of vectors orthogonal to $\boldsymbol{b_0}$. Now, explicitly this set of vectors is

$$\boldsymbol{u_\xi} = \boldsymbol{v_\xi} - \frac{1}{\sqrt{n}+1}\boldsymbol{v_0} = \boldsymbol{e_\xi} - \frac{1}{\sqrt{n}+1}(\boldsymbol{e_0} + \boldsymbol{b_0}) \quad , \xi = 1, \ldots, 2d-2. \tag{6.2.11}$$

Now, for $j = 1, \ldots, d-1$ we chose

$$\boldsymbol{b_j} = \frac{1}{\sqrt{2}}\boldsymbol{u_{2j-1}} = \frac{1}{\sqrt{2}}(\boldsymbol{e_{2j-1}} - \frac{1}{\sqrt{n}+1}(\boldsymbol{e_0} + \boldsymbol{b_0})), \tag{6.2.12}$$

$$\boldsymbol{c_j} = \frac{1}{\sqrt{2}}\boldsymbol{u_{2j}} = \frac{1}{\sqrt{2}}(\boldsymbol{e_{2j}} - \frac{1}{\sqrt{n}+1}(\boldsymbol{e_0} + \boldsymbol{b_0})). \tag{6.2.13}$$

With components,

$$b_j^\xi = \frac{1}{\sqrt{2}}\left(\delta_{2j-1}^\xi - \frac{1}{\sqrt{n}+1}(\delta_0^\xi + \frac{1}{\sqrt{n}})\right), \tag{6.2.14}$$

$$c_j^\xi = \frac{1}{\sqrt{2}}\left(\delta_{2j}^\xi - \frac{1}{\sqrt{n}+1}(\delta_0^\xi + \frac{1}{\sqrt{n}})\right). \tag{6.2.15}$$

Now, the elements of the FSM POVM with fiducial state $|0\rangle$ are of the form,

$$\left|\phi^\xi\right\rangle = b_0^\xi |0\rangle + \sum_{j=1}^{d-1}(b_j^\xi + ic_j^\xi)|j\rangle . \tag{6.2.16}$$

with $\xi = 0, \ldots, 2d-2$. Replacing the found coefficients.

$$\left|\phi^\xi\right\rangle = \frac{1}{\sqrt{n}}|0\rangle + \frac{1}{\sqrt{2}}\sum_{j=1}^{d-1}\left[\delta_{2j-1}^\xi + i\delta_{2j}^\xi - \frac{(1+i)}{\sqrt{n}+1}\left(\delta_0^\xi + \frac{1}{\sqrt{n}}\right)\right]|j\rangle . \tag{6.2.17}$$

We can write $(1+i) = \sqrt{2}\, e^{i\pi/4}$. Then, we have

$$\left|\phi^\xi\right\rangle = \frac{1}{\sqrt{n}}\left[|0\rangle + \sqrt{\frac{n}{2}}(\delta_{2k-1}^\xi + i\delta_{2k}^\xi)|k\rangle - e^{i\pi/4}\frac{\sqrt{n}}{\sqrt{n}+1}\sum_{j=1}^{d-1}\left(\delta_0^\xi + \frac{1}{\sqrt{n}}\right)|j\rangle\right]. \tag{6.2.18}$$

with $\xi = 0, \ldots, 2d-2$ and $k = 1, \ldots, d-1$.

Finally, the elements of this minimal FSM with fiducial state $|0\rangle$ are

$$\left|\phi^0\right\rangle = \frac{1}{\sqrt{n}}\left[|0\rangle - e^{i\pi/4}\sum_{j=1}^{d-1}|j\rangle\right], \tag{6.2.19}$$

$$\left|\phi^{2k-1}\right\rangle = \frac{1}{\sqrt{n}}\left[|0\rangle + \sqrt{\frac{n}{2}}|k\rangle - \frac{e^{i\pi/4}}{\sqrt{n}+1}\sum_{j=1}^{d-1}|j\rangle\right], \quad k = 1, ..., d-1 \tag{6.2.20}$$

$$\left|\phi^{2k}\right\rangle = \frac{1}{\sqrt{n}}\left[|0\rangle + i\sqrt{\frac{n}{2}}|k\rangle - \frac{e^{i\pi/4}}{\sqrt{n}+1}\sum_{j=1}^{d-1}|j\rangle\right], \quad k = 1, ..., d-1. \tag{6.2.21}$$

6.2.1 FSM from two orthonormal bases

At the cost of adding one more element to this set, we can construct the minimal-plus-one FSM using two different orthonormal bases, following the construction given in [70]. This can be implemented by randomly selecting

between measurements in these two bases, for instance, by flipping a coin.

The objective is to construct a set of $2d-1$ orthonormal vectors, denoted as $\{\boldsymbol{b_0}, \sqrt{2}\boldsymbol{b_j}, \sqrt{2}\boldsymbol{c_j}\}_{j=1,\dots,d-1}$, within an extended Hilbert space of dimension $n = 2d$. We begin with an arbitrary orthonormal basis $\{\boldsymbol{u_j}\}_{j=0,\dots,d-1}$ in d dimensions, where each basis vector is expressed as:

$$\boldsymbol{u_j} = \sum_{\xi=0}^{d-1} u_j^{\xi} e_{\xi}. \tag{6.2.22}$$

We choose $\boldsymbol{u_0}$ such that all its components are positive, i.e., $u_0^{\xi} > 0$ for $\xi = 0, \dots, d-1$. Using this basis, we define the $(2d)$-dimensional vectors:

$$\boldsymbol{b_0} = \frac{1}{\sqrt{2}} \begin{pmatrix} \boldsymbol{u_0} \\ \boldsymbol{u_0} \end{pmatrix} = \frac{1}{\sqrt{2}} \sum_{\xi=0}^{d-1} u_0^{\xi} (e_{\xi} + e_{d+\xi}), \tag{6.2.23}$$

$$\boldsymbol{b_j} = \frac{1}{\sqrt{2}} \begin{pmatrix} \boldsymbol{u_j} \\ \boldsymbol{0_d} \end{pmatrix} = \frac{1}{\sqrt{2}} \sum_{\xi=0}^{d-1} u_j^{\xi} e_{\xi}, \quad j = 1, \dots, d-1, \tag{6.2.24}$$

$$\boldsymbol{c_j} = \frac{1}{\sqrt{2}} \begin{pmatrix} \boldsymbol{0_d} \\ \boldsymbol{u_j} \end{pmatrix} = \frac{1}{\sqrt{2}} \sum_{\xi=0}^{d-1} u_j^{\xi} e_{d+\xi}, \quad j = 1, \dots, d-1. \tag{6.2.25}$$

Here, $\boldsymbol{0_d}$ represents the d-dimensional zero vector. These vectors are mutually orthogonal and satisfy the normalization conditions required for $\{\boldsymbol{b_0}, \sqrt{2}\boldsymbol{b_j}, \sqrt{2}\boldsymbol{c_j}\}_{j=1,\dots,d-1}$. Substituting them into the expression for the FSM elements in (6.2.16), we obtain:

$$\left|\phi^{\xi}\right\rangle = \frac{1}{\sqrt{2}} \left(u_0^{\xi} \left|0\right\rangle + \sum_{j=1}^{d-1} u_j^{\xi} \left|j\right\rangle \right) = \frac{1}{\sqrt{2}} \left|\chi^{\xi}\right\rangle, \tag{6.2.26}$$

$$\left|\phi^{d+\xi}\right\rangle = \frac{i}{\sqrt{2}} \left(-i u_0^{\xi} \left|0\right\rangle + \sum_{j=1}^{d-1} u_j^{\xi} \left|j\right\rangle \right) = \frac{i}{\sqrt{2}} \left|\tau^{\xi}\right\rangle, \tag{6.2.27}$$

for $\xi = 0, \dots, d-1$. Here, the sets $\{\left|\chi^{\xi}\right\rangle\}$ and $\{\left|\tau^{\xi}\right\rangle\}$ define two orthonormal bases. It is straightforward to verify that all elements of both bases have a nonzero projection onto the fiducial state $|0\rangle$.

Chapter 7

Adaptive Fisher Symmetric Tomography

As discussed in the previous chapter, Fisher-symmetric measurements provide an estimation protocol that guarantees accuracy close to the Gill-Massar bound. However, this high accuracy is only achievable when prior information about the unknown state is available. In particular, beyond knowing that the state is pure, we must also ensure that it is sufficiently close to a known reference pure state, the *fiducial* state. Due to this limitation, the Fisher-symmetric protocol cannot be directly applied to estimate an arbitrary unknown pure state.

To overcome this constraint, we can introduce an adaptive tomography scheme in which the measurement process is iteratively refined to approach the unknown state. This approach requires adapting the POVM at each step and defining an optimization criterion to ensure the convergence of the protocol. However, this adaptability comes at the cost of increased computational complexity.

In this chapter, we demonstrate that by using only two Fisher-symmetric POVMs, we can achieve a high-fidelity analytical reconstruction of a pure state. Moreover, the reconstructed state is sufficiently close to the unknown state so that a third adapted Fisher-symmetric POVM can estimate it with an accuracy near the Gill-Massar bound. Based on these results, we propose a **two-stage adaptive FSM scheme**. We validate this approach through Monte Carlo simulations, using the infidelity measure to quantify its performance. In this way, we generalize the applicability of Fisher-symmetric measurements for estimating arbitrary unknown

pure states in a d-dimensional Hilbert space.

7.1 Analytical Reconstruction from 2 FSM

Let $|\Psi\rangle$ be a d-dimensional pure state with no-null overlap with the fiducial state $|0\rangle$,

$$|\Psi\rangle = a_0 |0\rangle + \sum_{j=0}^{d-1} e^{i\phi_j} a_j |j\rangle , \quad a_0 > 0, \quad a_j \geq 0, \quad \phi_j \in [0, 2\pi]. \tag{7.1.1}$$

And consider two minimal FSM, with $n = 2d - 1$ elements, $\mathcal{E}_\pm = \{\left|\varphi_\pm^\alpha\right\rangle \left\langle\varphi_\pm^\alpha\right|\}$, with

$$\left|\varphi_\pm^\alpha\right\rangle = \beta_0^\alpha |0\rangle \pm \sum_{k=1}^{d-1} (\beta_k^\alpha + i\gamma_k^\alpha) |k\rangle , \quad \beta_0^\alpha > 0, \tag{7.1.2}$$

with $\beta_k^\alpha, \gamma_k^\alpha \in \mathbb{R}$, $\alpha = 0, \ldots, n-1$ and $k = 1, \ldots, d-1$.

Measuring the previous POVM on $|\Psi\rangle$, we get

$$P_\pm^\alpha = |\left\langle\varphi_\pm^\alpha\right|\Psi\rangle|^2 \tag{7.1.3}$$

$$= \left| \left(\beta_0^\alpha \langle 0| \pm \sum_{k=1}^{d-1} (\beta_k^\alpha - i\gamma_k^\alpha) \langle k| \right) \left(a_0 |0\rangle + \sum_{j=0}^{d-1} e^{i\phi_j} a_j |j\rangle \right) \right|^2 , \tag{7.1.4}$$

$$= \left| \beta_0^\alpha a_0 \pm \sum_{j=1}^{d-1} (\beta_j^\alpha - i\gamma_j^\alpha) e^{i\phi_j} a_j \right|^2 , \tag{7.1.5}$$

$$= (\beta_0^\alpha)^2 a_0^2 + \left| \sum_{j=1}^{d-1} (\beta_j^\alpha - i\gamma_j^\alpha) a_j e^{i\phi_j} \right|^2 \pm \beta_0^\alpha a_0 \sum_{j=1}^{d-1} (\beta_j^\alpha + i\gamma_j^\alpha) e^{-i\phi_j} a_j \tag{7.1.6}$$

$$\pm \beta_0^\alpha a_0 \sum_{j=1}^{d-1} (\beta_j^\alpha - i\gamma_j^\alpha) e^{i\phi_j} a_j,$$

$$P_\pm^\alpha = (\beta_0^\alpha)^2 a_0^2 + \left| \sum_{j=1}^{d-1} (\beta_j^\alpha - i\gamma_j^\alpha) a_j e^{i\phi_j} \right|^2 \pm 2\beta_0^\alpha a_0 \sum_{j=1}^{d-1} a_j (\beta_j^\alpha \cos(\phi_j) + \gamma_j^\alpha \sin(\phi_j)). \tag{7.1.7}$$

Subtracting the probabilities $P_{\pm}^{\alpha}$,

$$P_{+}^{\alpha} - P_{-}^{\alpha} = 4\beta_0^{\alpha} a_0 \sum_{j=1}^{d-1} a_j(\beta_j^{\alpha}\cos(\phi_j) + \gamma_j^{\alpha}\sin(\phi_j)), \tag{7.1.8}$$

and multiplying the previous expression by the coefficients β_k^{α} and γ_k^{α}, we have

$$\beta_k^{\alpha}(P_{+}^{\alpha} - P_{-}^{\alpha}) = 4a_0\beta_0^{\alpha} \sum_{j=1}^{d-1} a_j(\beta_k^{\alpha}\beta_j^{\alpha}\cos(\phi_j) + \beta_k^{\alpha}\gamma_j^{\alpha}\sin(\phi_j)), \tag{7.1.9}$$

$$\gamma_k^{\alpha}(P_{+}^{\alpha} - P_{-}^{\alpha}) = 4a_0\beta_0^{\alpha} \sum_{j=1}^{d-1} a_j(\gamma_k^{\alpha}\beta_j^{\alpha}\cos(\phi_j) + \gamma_k^{\alpha}\gamma_j^{\alpha}\sin(\phi_j)). \tag{7.1.10}$$

Summing over all α and remembering the orthogonality conditions Eq. (6.1.31),

$$\sum_{\alpha}\left(\frac{\beta_k^{\alpha}}{\beta_0^{\alpha}}\right)(P_{+}^{\alpha} - P_{-}^{\alpha}) = 4a_0 \sum_{j=1}^{d-1} a_j \sum_{\alpha}(\beta_k^{\alpha}\beta_j^{\alpha}\cos(\phi_j) + \beta_k^{\alpha}\gamma_j^{\alpha}\sin(\phi_j)), \tag{7.1.11}$$

$$= 4a_0 \sum_{j=1}^{d-1} a_j \frac{1}{2}\delta_{jk}\cos(\phi_j), \tag{7.1.12}$$

$$= 2a_0 a_k \cos(\phi_k). \tag{7.1.13}$$

$$\sum_{\alpha}\left(\frac{\gamma_k^{\alpha}}{\beta_0^{\alpha}}\right)(P_{+}^{\alpha} - P_{-}^{\alpha}) = 4a_0 \sum_{j=1}^{d-1} a_j \sum_{\alpha}(\gamma_k^{\alpha}\beta_j^{\alpha}\cos(\phi_j) + \gamma_k^{\alpha}\gamma_j^{\alpha}\sin(\phi_j)), \tag{7.1.14}$$

$$= 4a_0 \sum_{j=1}^{d-1} a_j \frac{1}{2}\delta_{jk}\sin(\phi_j), \tag{7.1.15}$$

$$= 2a_0 a_k \sin(\phi_k). \tag{7.1.16}$$

From these expressions, we can define the quantity

$$\Delta_k = \sum_{\alpha}\left(\frac{\beta_k^{\alpha} + i\gamma_k^{\alpha}}{\beta_0^{\alpha}}\right)(P_{+}^{\alpha} - P_{-}^{\alpha}), \tag{7.1.17}$$

$$= 2a_0 a_k e^{i\phi_k}. \tag{7.1.18}$$

With this, we can find the phases ϕ_k from the measured probabilities as

$$\tan(\phi_k) = \frac{\operatorname{Im}\Delta_k}{\operatorname{Re}\Delta_k}. \tag{7.1.19}$$

On the other hand, adding the probabilities $P_{\pm}^{\alpha}$, we have

$$P_{+}^{\alpha} + P_{-}^{\alpha} = 2(\beta_0^{\alpha})^2 a_0^2 + 2\left|\sum_{k=1}^{d-1}(\beta_k^{\alpha} - i\gamma_k^{\alpha})a_k e^{i\phi_k}\right|^2. \tag{7.1.20}$$

Recalling that

$$\frac{\Delta_k}{2a_0} = a_k e^{i\phi_k}, \tag{7.1.21}$$

we find a quadratic equation for a_0^2,

$$2a_0^2(P_{+}^{\alpha} + P_{-}^{\alpha}) = 4(\beta_0^{\alpha})^2(a_0^2)^2 + \left|\sum_{k=1}^{d-1}(\beta_k^{\alpha} - i\gamma_k^{\alpha})\Delta_k\right|^2. \tag{7.1.22}$$

Now, from Δ_k we have another quadratic equation for a_0^2,

$$\sum_{k=1}^{d-1}|\Delta_k|^2 = 4a_0^2(1 - a_0^2). \tag{7.1.23}$$

Inserting this equation in the previous one, we have a quadratic equation for a_0,

$$2a_0^2(P_{+}^{\alpha} + P_{-}^{\alpha}) = 4(\beta_0^{\alpha})^2(a_0^2 - \frac{1}{4}\sum_{k=1}^{d-1}|\Delta_k|^2) + \left|\sum_{k=1}^{d-1}(\beta_k^{\alpha} - i\gamma_k^{\alpha})\Delta_k\right|^2. \tag{7.1.24}$$

but since a_0 is defined as positive, there is only one valid solution. Solving this equation,

$$a_0^{\alpha} = \sqrt{\frac{\left|\sum_{k=1}^{d-1}(\beta_k^{\alpha} - i\gamma_k^{\alpha})\Delta_k\right|^2 - (\beta_0^{\alpha})^2\sum_{k=1}^{d-1}|\Delta_k|^2}{2(P_{+}^{\alpha} + P_{-}^{\alpha}) - 4(\beta_0^{\alpha})^2}}. \tag{7.1.25}$$

There are $n = 2d - 1$ equations for the coefficient a_0. Analytically, they all have the same value, but since the probabilities arise from experimental measures the values can vary. For that reason, we recommend taking the mean of all the values as the estimation of a_0,

$$a_0 = \frac{1}{n}\sum_{\alpha} a_0^{\alpha}. \tag{7.1.26}$$

Finally, having ϕ_k and a_0 from (7.1.19) and (7.1.26), we can find the coefficients a_k replacing these values into the definition of Δ_k (7.1.18). With this, we can perform an analytical reconstruction of the unknown state $|\Psi\rangle$ from the experimental

probabilities $P_{\pm}^{\alpha}$.

7.2 Arbitrary Phase FSM

In this section we demonstrate that if $\mathcal{E} = \{|\varphi^{\alpha}\rangle \langle\varphi^{\alpha}|\}$ is a FSM, with $|\varphi^{\alpha}\rangle$ given by Eq. (6.1.15), then the POVM $\mathcal{E}' = \{|\varphi'^{\alpha}\rangle \langle\varphi'^{\alpha}|\}$, with

$$|\varphi'^{\alpha}\rangle = \beta_0^{\alpha} |0\rangle + e^{i\phi} \sum_{k=1}^{d-1} (\beta_k^{\alpha} + i\gamma_k^{\alpha}) |k\rangle , \quad \phi \in [0, 2\pi], \tag{7.2.1}$$

is also a FSM. Let us define the new coefficients

$$\beta_k'^{\alpha} = \mathrm{Re}\left(e^{i\phi}(\beta_k^{\alpha} + i\gamma_k^{\alpha})\right) = \cos(\phi)\beta_k^{\alpha} - \sin(\phi)\gamma_k^{\alpha}. \tag{7.2.2}$$

$$\gamma_k'^{\alpha} = \mathrm{Im}\left(e^{i\phi}(\beta_k^{\alpha} + i\gamma_k^{\alpha})\right) = \cos(\phi)\gamma_k^{\alpha} + \sin(\phi)\beta_k^{\alpha}, \tag{7.2.3}$$

and their respective real vectors $\vec{\beta}'_k = (\beta_k'^0, \ldots, \beta_k'^n)$ and $\vec{\gamma}'_k = (\gamma_k'^0, \ldots, \gamma_k'^n)$. Checking if the vectors $\vec{\beta}'_k$ and $\vec{\gamma}'_k$ satisfy the FSM conditions,

$$\vec{\beta}_0 \cdot \vec{\beta}_0 = 1, \tag{7.2.4}$$

$$\vec{\beta}_0 \cdot \vec{\beta}'_k = (\cos(\phi)\vec{\beta}_0 \cdot \vec{\beta}_k - \sin(\phi)\vec{\beta}_0 \cdot \vec{\gamma}_k) = 0, \tag{7.2.5}$$

$$\vec{\beta}_0 \cdot \vec{\gamma}'_k = (\sin(\phi)\vec{\beta}_0 \cdot \vec{\beta}_k + \cos(\phi)\vec{\beta}_0 \cdot \vec{\gamma}_k) = 0, \tag{7.2.6}$$

$$\vec{\beta}'_j \cdot \vec{\gamma}'_k = 2(\cos(\phi)\vec{\beta}_j - \sin(\phi)\vec{\gamma}_j) \cdot (\sin(\phi)\vec{\beta}_j + \cos(\phi)\vec{\gamma}_j), \tag{7.2.7}$$

$$= 2\cos(\phi)\sin(\phi)\vec{\beta}_j \cdot \vec{\beta}_k + 2\cos^2(\phi)\vec{\beta}_j \cdot \vec{\gamma}_k - 2\sin^2(\phi)\vec{\gamma}_j \cdot \vec{\beta}_k \tag{7.2.8}$$

$$- 2\cos(\phi)\sin(\phi)\vec{\gamma}_j \cdot \vec{\gamma}_k$$

$$= 0, \tag{7.2.9}$$

$$\vec{\beta}'_j \cdot \vec{\beta}'_k = 2(\cos(\phi)\vec{\beta}_j - \sin(\phi)\vec{\gamma}_j) \cdot (\cos(\phi)\vec{\beta}_k - \sin(\phi)\vec{\gamma}_k), \tag{7.2.10}$$

$$= 2\cos^2(\phi)\vec{\beta}_j \cdot \vec{\beta}_k - 2\sin(\phi)\cos(\phi)\vec{\gamma}_j \cdot \vec{\beta}_k - 2\sin(\phi)\cos(\phi)\vec{\beta}_j \cdot \vec{\gamma}_k \tag{7.2.11}$$

$$+ 2\sin^2(\phi)\vec{\gamma}_j \cdot \vec{\gamma}_k,$$

$$= \frac{1}{2}\delta_{jk}. \tag{7.2.12}$$

$$\vec{\gamma}'_j \cdot \vec{\gamma}'_k = 2(\sin(\phi)\vec{\beta}_j + \cos(\phi)\vec{\gamma}_j) \cdot (\sin(\phi)\vec{\beta}_k + \cos(\phi)\vec{\gamma}_k),$$

$$= 2\sin^2(\phi)\vec{\beta}_j \cdot \vec{\beta}_k + 2\sin(\phi)\cos(\phi)\vec{\beta}_j \cdot \vec{\gamma}_k + 2\sin(\phi)\cos(\phi)\vec{\gamma}_j \cdot \vec{\beta}_k \tag{7.2.13}$$

$$+ 2\cos^2(\phi)\vec{\gamma}_j \cdot \vec{\gamma}_k,$$

$$= \frac{1}{2}\delta_{jk}. \tag{7.2.14}$$

with $j,k=1,\ldots,d-1$. We can see that the new coefficients $\{\beta_0'^{\alpha},\beta_j'^{\alpha},\gamma_j'^{\alpha}\}_{j=1}^{d-1}$ of the new POVM satisfy the conditions to be a FSM, Eq. (6.1.31), therefore $\mathcal{E}'$ also is a FSM.

7.3 Collections of FSMs are also FSMs

In this section we demonstrate that a collection of m FSMs $\{\mathcal{E}_i\}_{i=0}^{m-1}$, with $n\geq 2d-1$ elements $\mathcal{E}_i=\{|\varphi_i^{\alpha}\rangle\}_{\alpha=0}^{n-1}$, such that

$$|\varphi_i^{\alpha}\rangle=\beta_{i,0}^{\alpha}|0\rangle+\sum_{k=1}^{d-1}(\beta_{i,k}^{\alpha}+i\gamma_{i,k}^{\alpha})|k\rangle\,,\quad \beta_{i,0}^{l}>0,\quad \beta_{i,k}^{\alpha},\gamma_{i,k}^{\alpha}\in\mathbb{R}, \tag{7.3.1}$$

is also a FSM. The coefficients $\{\beta_{i,0}^{\alpha},\beta_{i,k}^{\alpha},\gamma_{i,k}^{\alpha})\}$ satisfy the FSM conditions, Eq. (6.1.31), for each $i=0,\ldots,m-1$.

Let us consider the POVM $\mathcal{E}'=\{\tau_i|\varphi_i^{\alpha}\rangle\}$, with $\{\tau_i\}$ positive coefficients such as $\sum_i|\tau_i|^2=1$. Defining the coefficient of each POVM element,

$$\beta_{i,0}'^{\alpha}=\tau_i\beta_{i,0}^{\alpha}, \tag{7.3.2}$$

$$\beta_{i,k}'^{\alpha}=\tau_i\beta_{i,k}^{\alpha}, \tag{7.3.3}$$

$$\gamma_{i,k}'^{\alpha}=\tau_i\gamma_{i,k}^{\alpha}, \tag{7.3.4}$$

and their respective $mn-$dimensional vectors $\vec{\beta}_0'=(\beta_{0,0}'^{0},\ldots,\beta_{m,0}'^{n})$, $\vec{\beta}_k'=(\beta_{0,k}'^{0},\ldots,\beta_{m,k}'^{n})$, $\vec{\gamma}_0'=(\gamma_{0,k}'^{0},\ldots,\gamma_{m,k}'^{n})$, we can verify if $\mathcal{E}'$ satisfies the conditions for being a FSM, Eq. (6.1.31),

$$\vec{\beta}_0'\cdot\vec{\beta}_0'=\sum_{i=0}^{m-1}|\tau_i|^2\sum_{\alpha=0}^{n_i-1}\beta_{i,0}^{\alpha}\beta_{i,0}^{\alpha}=1 \tag{7.3.5}$$

$$\vec{\beta}_0'\cdot\vec{\beta}_k'=\sum_{i=0}^{m-1}|\tau_i|^2\sum_{\alpha=0}^{n_i-1}\beta_{i,0}^{\alpha}\beta_{i,k}^{\alpha}=0. \tag{7.3.6}$$

$$\vec{\beta}_0'\cdot\vec{\gamma}_k'=\sum_{i=0}^{m-1}|\tau_i|^2\sum_{\alpha=0}^{n_i-1}\beta_{i,0}^{\alpha}\gamma_{i,k}^{\alpha}=0. \tag{7.3.7}$$

$$\vec{\beta}_j'\cdot\vec{\gamma}_k'=\sum_{i=0}^{m-1}|\tau_i|^2\sum_{\alpha=0}^{n_i-1}\beta_{i,j}^{\alpha}\gamma_{i,k}^{\alpha}=0. \tag{7.3.8}$$

$$\vec{\beta}'_j \cdot \vec{\beta}'_k = \sum_{i=0}^{m-1} |\tau_i|^2 \sum_{\alpha=0}^{n_i-1} \beta_{i,j}^{\alpha} \beta_{i,k}^{\alpha} = \frac{1}{2}\delta_{jk}. \tag{7.3.9}$$

$$\vec{\gamma}'_j \cdot \vec{\gamma}'_k = \sum_{i=0}^{m-1} |\tau_i|^2 \sum_{\alpha=0}^{n_i-1} \gamma_{i,j}^{\alpha} \gamma_{i,k}^{\alpha} = \frac{1}{2}\delta_{jk}. \tag{7.3.10}$$

From this, we can see that these elements satisfy the FSM conditions. Therefore, the set $\mathcal{E}'$ corresponds to a Fisher Symmetric POVM with mn elements.

7.4 Adaptive Tomographic Method

We propose a three-stage adaptive pure state estimation method based on three FSMs and a single-shot measurement basis. This method also achieves the GMB and, apart from the purity of the unknown state, requires no other a priori information. In this way, we extend the applicability of FSMs to any unknown state.

Consider two minimal FSMs given by the POVMs $\mathcal{E}_{\pm} = \{|\varphi_{\pm}^{\alpha}\rangle\langle\varphi_{\pm}^{\alpha}|\}_{\alpha=1}^{n}$ with

$$|\varphi_{\pm}^{\alpha}\rangle = \beta_0^{\alpha}|0\rangle \pm \sum_{k=1}^{d-1} \beta_k^{\alpha}|k\rangle, \tag{7.4.1}$$

where the coefficients $\{\beta_k\}$ satisfy the conditions for being a FSM [**?**]. Each one of these FSMs estimates pure states in a neighborhood of the fiducial state $|0\rangle$. However, we showed that taken together, they can estimate an arbitrary pure state $|\Psi\rangle = a_0|0\rangle + \sum_{k=1}^{d-1} a_k e^{i\phi_k}|k\rangle$ with $a_k \geq 0$, except those orthogonal to the fiducial state $|0\rangle$.

This case can be avoided by introducing a single-shot measurement on an arbitrary basis $\{|k\rangle\}$, the outcome obtained from this measurement has non-null overlap with the state $|\Psi\rangle$, so it can be used as fiducial state to construct the FSMs $\mathcal{E}_{\pm}$. Notice that this adaptation does not ensure that the unknown state $|\Psi\rangle$ is close to the fiducial state, but permits to always obtain an estimate $|\tilde{\Psi}\rangle$ of $|\Psi\rangle$.

The estimate $|\tilde{\Psi}\rangle$ obtained from the FSMs $\mathcal{E}_{\pm}$ allows us to perform a near-optimal estimation of $|\Psi\rangle$ adapting the FSM $\mathcal{E}$. Let U be a unitary operator that transforms the fiducial state $|0\rangle$ into the estimate of the Two-FSMs, that is $|\tilde{\Psi}\rangle = U|0\rangle$. This unitary is constructed from the basis of the Hilbert space such that the first element

$|u_0\rangle$ corresponds to the estimated state $|\tilde{\Psi}\rangle$, this is done via the Gram-Schmidt procedure. In this way, the unitary U is of the form

$$U = |\tilde{\Psi}\rangle\langle 0| + \sum_{k=1}^{d-1} |u_k\rangle\langle k| . \tag{7.4.2}$$

Then, the unitary U adapts the FSM $\mathcal{E}$ into $\tilde{\mathcal{E}} = \{U|\varphi^\alpha\rangle\langle\varphi^\alpha|U^\dagger\}_{\alpha=1}^n$, which has $|\tilde{\Psi}\rangle$ as fiducial state. As long as the estimation with two-FSM is performed with sufficient sample size, the estimate $|\tilde{\Psi}\rangle$ is close enough to the unknown state $|\Psi\rangle$ to ensure that the adapted FSM $\tilde{\mathcal{E}}$ is optimal.

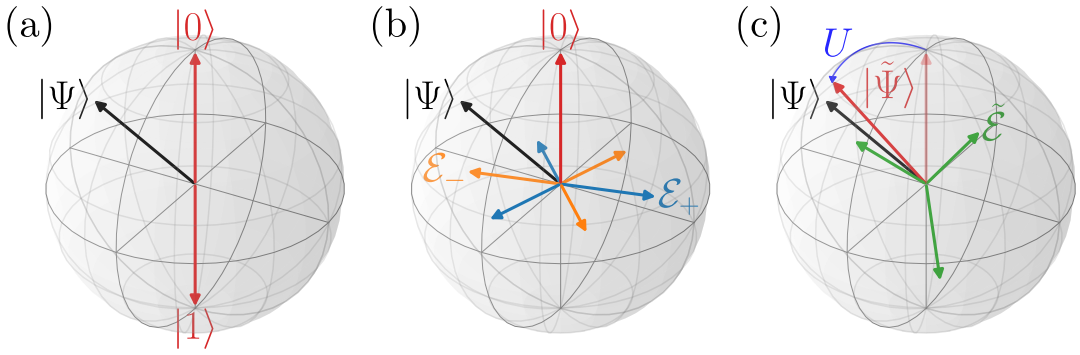

Figure 7.4.1: Schematic representation of the adaptive Fisher-Symmetric measurements for an unknown pure state $|\Psi\rangle$. (a) Single-shot measurement in the basis $\{|0\rangle, |1\rangle\}$ (red) for selecting a fiducial state. (b) Two-FSMs $\mathcal{E}_\pm$ (blue and orange) with fiducial state $|0\rangle$ (red). (c) Adapted FSM $\tilde{\mathcal{E}}$ (green) with fiducial state $|\tilde{\Psi}\rangle$ (red) estimated in the previous step. The unitary U (blue) is a change of basis such as $|\tilde{\Psi}\rangle = U|0\rangle$.

The complete estimation method reads as follows:

(0) Single-shot measurement on an arbitrary basis to select a fiducial state that has a nonzero overlap with the unknown pure state $|\Psi\rangle$, as shown in Fig. 7.4.1(a).

(1) Acquisition of data through FSM $\mathcal{E}_\pm$ to obtain probabilities $P_\pm^\alpha$. These probabilities are used to obtain an estimate of $|\Psi\rangle$ by solving the system of equations (7.1.19), (7.1.26) and (7.1.18). This preliminary estimate is refined through the maximum likelihood estimation leading to the first estimate $|\tilde{\Psi}\rangle$. This is shown in Fig. 7.4.1(b).

(2) The adapted FSM $\tilde{\mathcal{E}}$ allows us to obtain probabilities $\tilde{P}^\alpha$. These probabilities,

along with those from the previous step $P_{\pm}^{\alpha}$, are used to derive a final estimate $|\hat{\Psi}\rangle$ using maximum likelihood estimation, with the first estimate $|\tilde{\Psi}\rangle$ serving as the starting point for optimization. This is shown in Fig. 7.4.1(c).

The total number of measurement outcomes of our method scales linearly as $7d-3$ and does not resort to collective measurements on multiple copies of the unknown state.

7.5 Simulations

To analyze the proposed protocol and compare it with other tomographic methods, we performed numerical Monte Carlo simulations using Python language. For each dimension, 100 random unknown pure states were chosen, $\Omega = \{|\Psi\rangle\}_{i=1}^{100}$, according to a Haar uniform distribution [76] and for each of those states 10 state estimations are performed using the protocol. For the construction of the FSM elements, we follow 6.2. Since we need two FSM, to construct the other FSM we flip the sign of the coefficients of all states different from $|0\rangle$,

$$\left|\phi^{0}\right\rangle = \frac{1}{\sqrt{n}}\left[|0\rangle \mp e^{i\pi/4}\sum_{j=1}^{d-1}|j\rangle\right], \tag{7.5.1}$$

$$\left|\phi^{2k-1}\right\rangle = \frac{1}{\sqrt{n}}\left[|0\rangle \pm \sqrt{\frac{n}{2}}\,|k\rangle \mp \frac{e^{i\pi/4}}{\sqrt{n}+1}\sum_{j=1}^{d-1}|j\rangle\right], \quad k=1,...,d-1, \tag{7.5.2}$$

$$\left|\phi^{2k}\right\rangle = \frac{1}{\sqrt{n}}\left[|0\rangle \pm i\sqrt{\frac{n}{2}}\,|k\rangle \mp \frac{e^{i\pi/4}}{\sqrt{n}+1}\sum_{j=1}^{d-1}|j\rangle\right], \quad k=1,...,d-1. \tag{7.5.3}$$

And the measures are simulated by estimating the probability distribution as a multinomial distribution. The accuracy is given by the infidelity at the final estimation. For a better understanding of the results, we plot the infidelities versus ensemble used and infidelities versus the dimension of the unknown pure state at each stage.

Remembering the Gill-Massar bound for the infidelity; the infidelity can be written as a function of the dimension d and size of the ensemble N,

$$\mathcal{I} = \alpha\frac{(d-1)^{\gamma}}{N^{\beta}}. \tag{7.5.4}$$

To visualize the performance of the method, we calculate the infidelities of the estimations at each step. Then, perform curve fitting of the results considering the log mean infidelity function

$$\log_{10}(\bar{\mathcal{I}}) = \log_{10}(\alpha) - \beta \log_{10}(N) + \gamma \log_{10}(d-1) \tag{7.5.5}$$

to find the coefficients α, β, and γ, and observe how the method scales in comparison to the Gill-Massar bound. The bound is attained for $\alpha = \beta = \gamma = 1$. Notice that for the fittings of the first stage, N is the size of the ensemble size used at that step.

First, we determine the coefficient γ by fitting a linear function to the curves of $\langle\bar{\mathcal{I}}\rangle$ versus d for each value of N. Next, we obtain the coefficient β by fitting a linear function to the curves of $\langle\bar{\mathcal{I}}\rangle$ versus N for each value of d. Finally, to determine α, we perform a non-linear fit using the function in Eq. (7.5.5) to the curves of $\langle\bar{\mathcal{I}}\rangle$ versus N for each d. At this step, we fix the coefficients β and γ as the averages of the values obtained previously.

For the simulations, we consider three possible ensemble distributions among the measurements at each step.

- Distribution 2/3: this means that for the first step 2/3 of the total ensemble N is used, so $N/3$ of copies of the state are for each of the first two measurements. At the last step, $N/3$ of copies are used for the measurement. So, for this ensemble, each measurement uses the same fraction of the ensemble.
- Distribution 2/4: the total ensemble is divided equally between the two stages, so the first two measurements at the first step use $N/4$ of each. In the last step, the final measurement uses the remaining $N/2$ copies.
- Distribution 2/5: the ensemble is distributed so that for the first step, 1/5 of the ensemble is used at each of the measurements, and the 3/5 left of the ensemble is used for the last measurement, at the final step.

The code files necessary for the simulation and the simulated data used for this research can be found at https://github.com/CoVargasRo/FSM.

7.5.1 Distribution 2/3

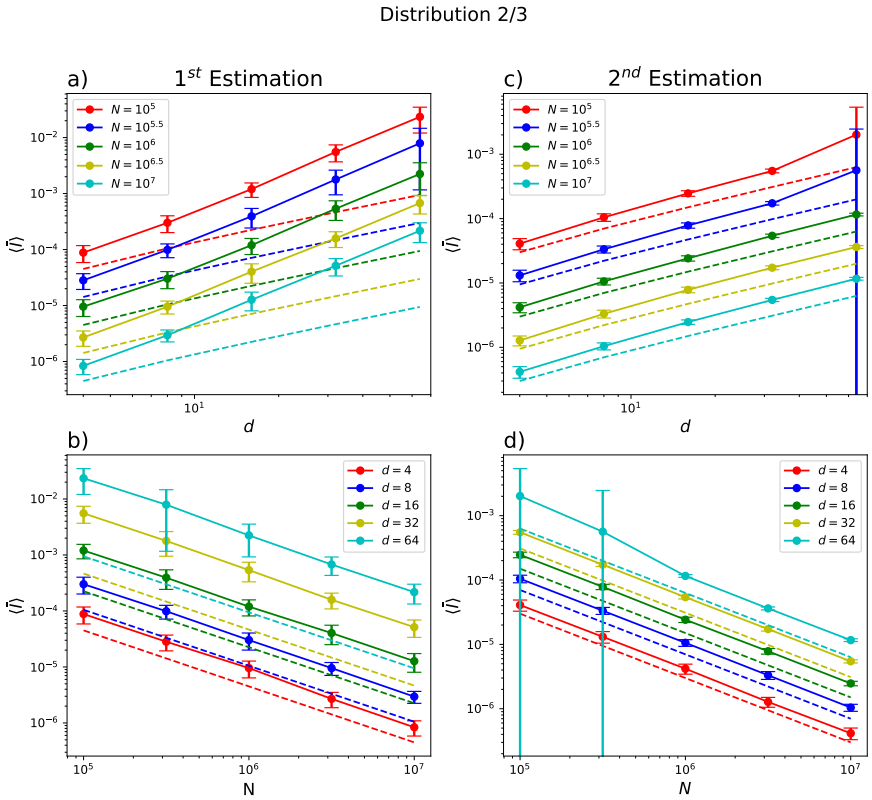

Figure 7.5.1: Average infidelity and standard deviation achieved in 10 estimations of 100 different unknown states using the ensemble distribution 2/3, for states of dimension $d = 4, 8, 16, 32$ and ensemble sizes $N = 10^5, 10^{5.5}, 10^6, 10^{6.5}, 10^7$. Insets a) and c) shows average infidelity vs state dimension, and insets b) and d) shows average infidelity vs ensemble size. The GMB is displayed (dashed line) for the first stage (insets a) and b)) as $3(d-1)/2N$ and for the final stage (insets c) and d)) as $(d-1)/N$.

Values and standard deviations of the coefficients α, β and γ entering in the lineal fit of the mean infidelity (7.5.5), generated by this method using the ensemble distribution of 2/3, are shown for each stage in the tables (7.5.1) and (7.5.2).

d	$\hat{\alpha} \pm \Delta\alpha$	$\hat{\beta} \pm \Delta\beta$	$log_{10}N$	$\hat{\gamma} \pm \Delta\gamma$
4	0.78 ± 0.02	1.01 ± 0.01	5	1.85 ± 0.08
8	0.562 ± 0.005	1.006 ± 0.005	5.5	1.86 ± 0.08
16	0.565 ± 0.008	0.988 ± 0.006	6	1.81 ± 0.09
32	0.64 ± 0.01	1.023 ± 0.008	6.5	1.83 ± 0.06
64	0.74 ± 0.02	1.03 ± 0.01	7	1.84 ± 0.06

Table 7.5.1: Coefficients for the first stage using the distribution 2/3.

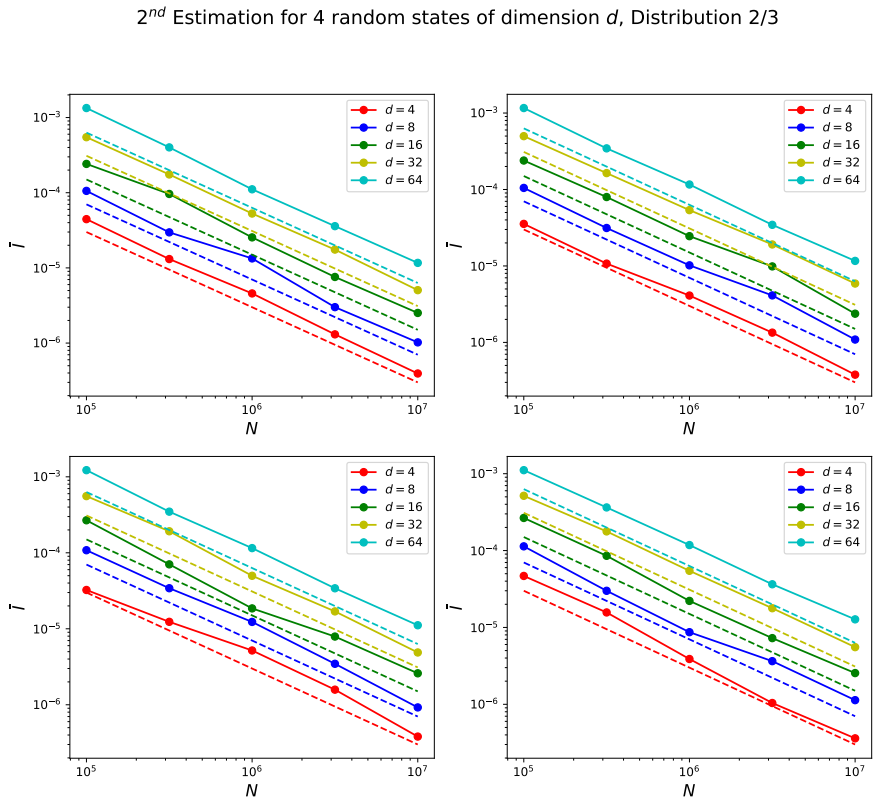

Figure 7.5.2: Mean infidelity achieved in 10 estimations for four different states of each dimension $d = 4, 8, 16, 32, 64$ using different ensemble sizes N.

d	$\hat{\alpha} \pm \Delta\alpha$	$\hat{\beta} \pm \Delta\beta$	$log_{10}N$	$\hat{\gamma} \pm \Delta\gamma$
4	1.165 ± 0.008	0.999 ± 0.005	5	1.25 ± 0.08
8	1.120 ± 0.003	1.001 ± 0.002	5.5	1.21 ± 0.06
16	1.093 ± 0.005	1.000 ± 0.003	6	1.092 ± 0.003
32	1.053 ± 0.003	1.002 ± 0.002	6.5	1.10 ± 0.01
64	1.2 ± 0.1	1.13 ± 0.04	7	1.098 ± 0.006

Table 7.5.2: Coefficients for the second stage using the distribution 2/3.

We can see that the coefficients are well approximated by $\alpha \approx 0.66$, $\beta = 1$ and $\gamma = 1.84$ at the first stage, and $\alpha \approx 1.1$, $\beta \approx 1$ and $\gamma \approx 1.15$ at the second stage. Finally, we can approximate the average infidelity at each stage as

$$\bar{\mathcal{I}}_1(|\Psi\rangle) \approx 0.66 \, \frac{(d-1)^{1.84}}{N}, \tag{7.5.6}$$

$$\bar{\mathcal{I}}_2(|\Psi\rangle) \approx 1.1 \, \frac{(d-1)^{1.15}}{N}. \tag{7.5.7}$$

7.5.2 Distribution 2/4

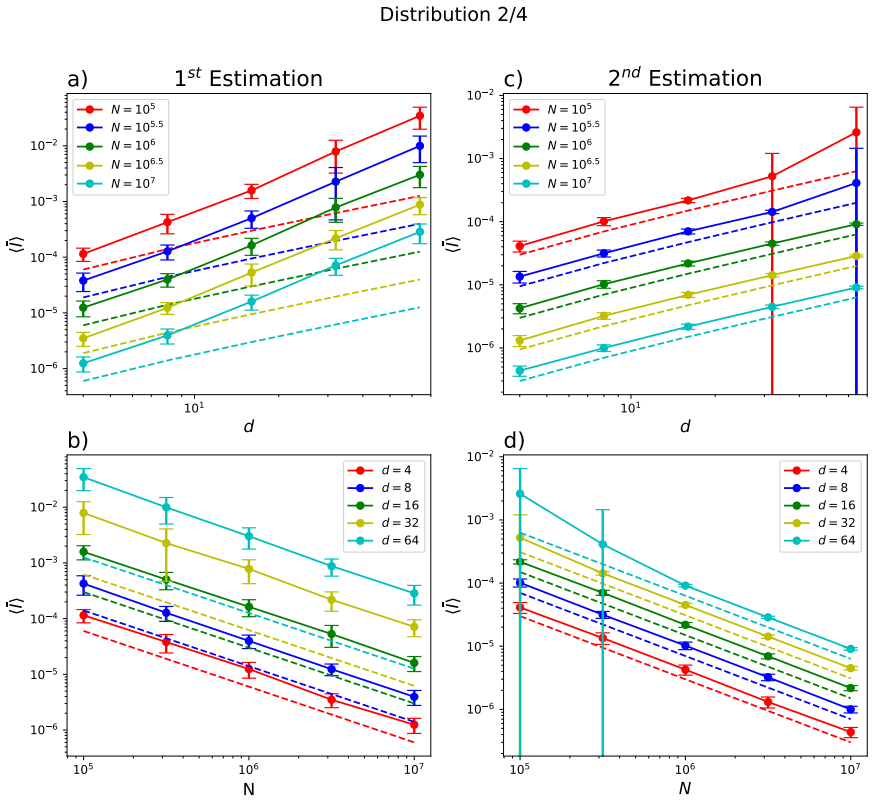

Figure 7.5.3: Average infidelity and standard deviation achieved in 10 estimations of 100 different unknown states using the ensamble distribution 2/4, for states of dimension $d = 4, 8, 16, 32$ and ensamble sizes $N = 10^5, 10^{5.5}, 10^6, 10^{6.5}, 10^7$. Insets a) and c) shows average infidelity vs state dimension, and insets b) and d) shows average infidelity vs ensemble size. The GMB is displayed (dashed line) for the first stage (insets a) and b)) as $2(d-1)/1N$ and for the final stage (insets c) and d)) as $(d-1)/N$.

Values and standard deviations of the coefficients α, β and γ entering in the lineal fit of the mean infidelity (7.5.5), generated by this method using the ensamble distribution of 2/4, are shown for each stage in the tables 7.5.3 and 7.5.4.

d	$\hat{\alpha} \pm \Delta\alpha$	$\hat{\beta} \pm \Delta\beta$	$log_{10}N$	$\hat{\gamma} \pm \Delta\gamma$
4	0.77 ± 0.02	0.99 ± 0.02	5	1.89 ± 0.08
8	0.560 ± 0.008	1.016 ± 0.006	5.5	1.84 ± 0.08
16	0.554 ± 0.006	0.993 ± 0.006	6	1.84 ± 0.08
32	0.66 ± 0.02	1.02 ± 0.01	6.5	1.83 ± 0.06
64	0.74 ± 0.03	1.04 ± 0.01	7	1.81 ± 0.08

Table 7.5.3: Coefficients for the first stage using the distribution 2/4.

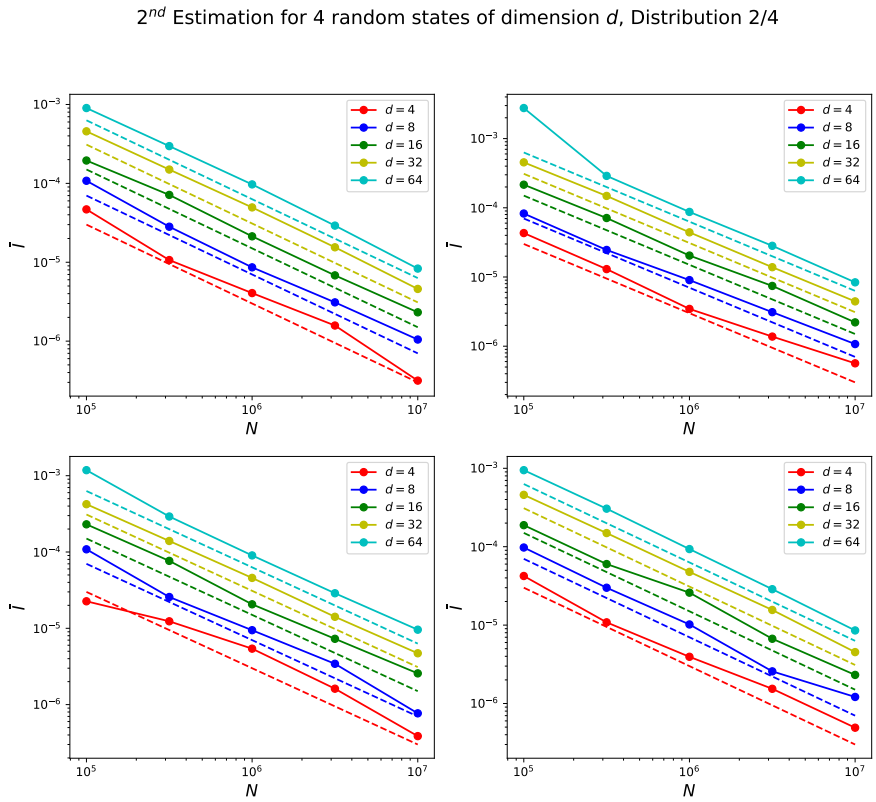

Figure 7.5.4: Mean infidelity achieved in 10 estimations for four different states of each dimension $d = 4, 8, 16, 32, 64$ using different ensemble sizes.

d	$\hat{\alpha} \pm \Delta\alpha$	$\hat{\beta} \pm \Delta\beta$	$log_{10}N$	$\hat{\gamma} \pm \Delta\gamma$
4	1.41 ± 0.01	0.991 ± 0.004	5	1.18 ± 0.09
8	1.45 ± 0.02	1.001 ± 0.009	5.5	1.05 ± 0.02
16	1.461 ± 0.004	1.000 ± 0.002	6	1.15 ± 0.09
32	1.468 ± 0.008	1.004 ± 0.003	6.5	1.00 ± 0.01
64	1.9 ± 0.3	1.12 ± 0.07	7	1.008 ± 0.008

Table 7.5.4: Coefficients for the second stage using the distribution 2/4.

We can see that the coefficients are well approximated by $\alpha \approx 0.66$, $\beta = 1$ and $\gamma = 1.84$ at the first stage, and $\alpha \approx 1.45$, $\beta \approx 1$ and $\gamma \approx 1$ at the second stage. Finally, we can approximate the average infidelity at each stage as

$$\bar{\mathcal{I}}_1(|\Psi\rangle) \approx 0.66 \frac{(d-1)^{1.84}}{N}, \tag{7.5.8}$$

$$\bar{\mathcal{I}}_2(|\Psi\rangle) \approx 1.45 \frac{(d-1)}{N}. \tag{7.5.9}$$

7.5.3 Distribution 2/5

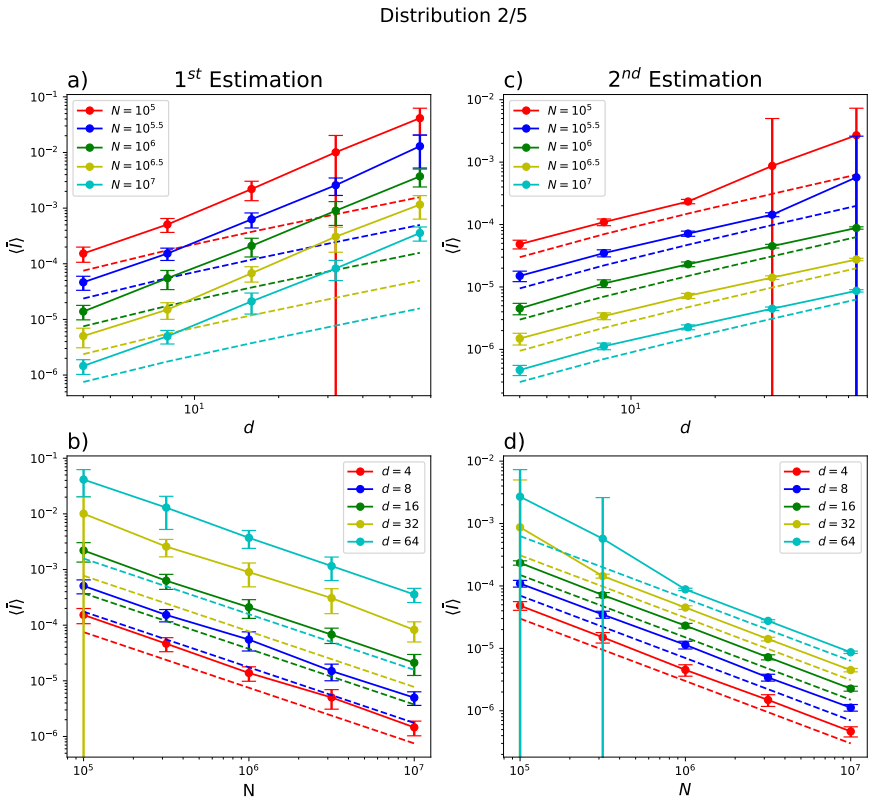

Figure 7.5.5: Average infidelity and standard deviation achieved in 10 estimations of 100 different unknown states using the ensamble distribution of 2/5, for states of dimension $d = 4, 8, 16, 32$ and ensamble sizes $N = 10^5, 10^{5.5}, 10^6, 10^{6.5}, 10^7$. Insets a) and c) shows average infidelity vs state dimension, and insets b) and d) shows average infidelity vs ensemble size. The GMB is displayed (dashed line) for the first stage (insets a) and b)) as $5(d-1)/2N$ and for the final stage (insets c) and d)) as $(d-1)/N$.

d	$\hat{\alpha} \pm \Delta\alpha$	$\hat{\beta} \pm \Delta\beta$	$log_{10}N$	$\hat{\gamma} \pm \Delta\gamma$
4	0.79 ± 0.02	1.00 ± 0.02	5	1.87 ± 0.08
8	0.56 ± 0.01	1.01 ± 0.02	5.5	1.85 ± 0.09
16	0.578 ± 0.009	1.00 ± 0.01	6	1.84 ± 0.05
32	0.65 ± 0.03	1.0 ± 0.3	6.5	1.83 ± 0.09
64	0.75 ± 0.02	1.036 ± 0.007	7	1.82 ± 0.07

Table 7.5.5: Coefficients for the first stage using the distribution 2/5.

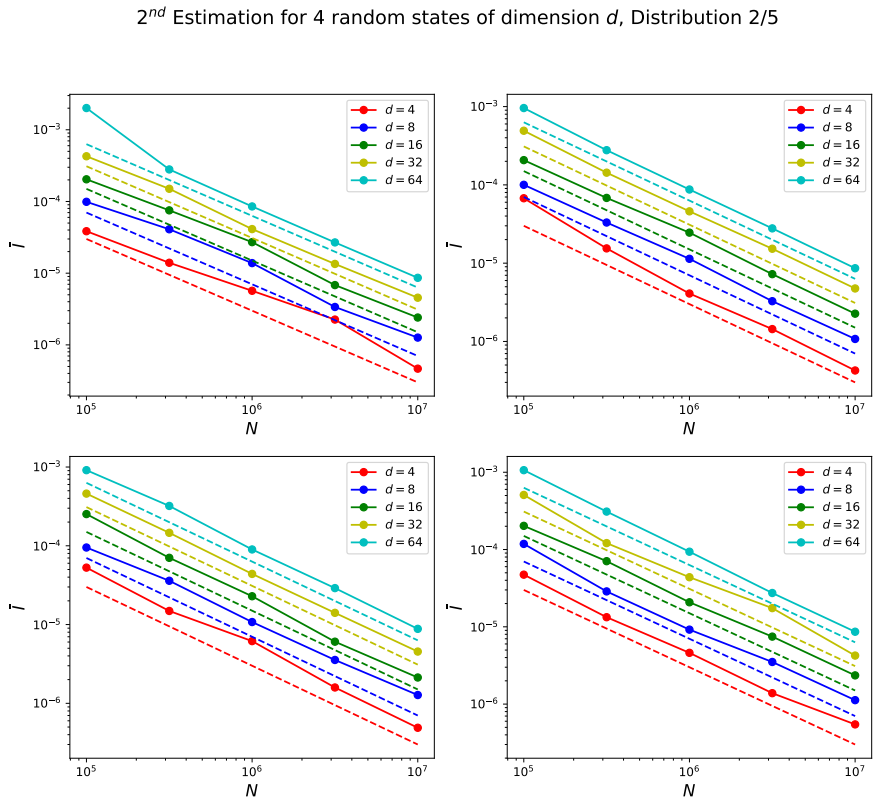

Figure 7.5.6: Mean infidelity achieved in 10 estimations for four different states of each dimension $d = 4, 8, 16, 32, 64$ using different ensemble sizes.

d	$\hat{\alpha} \pm \Delta\alpha$	$\hat{\beta} \pm \Delta\beta$	$log_{10}N$	$\hat{\gamma} \pm \Delta\gamma$
4	1.57 ± 0.02	1.007 ± 0.007	5	1.3 ± 0.1
8	1.58 ± 0.01	0.997 ± 0.006	5.5	1.1 ± 0.1
16	1.53 ± 0.01	1.005 ± 0.003	6	0.97 ± 0.02
32	1.7 ± 0.2	1.12 ± 0.06	6.5	0.957 ± 0.007
64	2.0 ± 0.5	1.26 ± 0.07	7	0.95 ± 0.01

Table 7.5.6: Coefficients for the second stage using the distribution 2/5.

Values and standard deviations of the coefficients α, β and γ entering in the lineal fit of the mean infidelity (7.5.5), generated by this method using the ensamble distribution of 2/5, are shown for each stage in the tables 7.5.5 and 7.5.6.

We can see that the coefficients are well approximated by $\alpha \approx 0.66$, $\beta = 1$ and $\gamma = 1.84$ at the first stage, and $\alpha \approx 1.45$, $\beta \approx 1$ and $\gamma \approx 1$ in the second stage.

Finally we can approximate the average infidelity at each stage as

$$\bar{\mathcal{I}}_1(|\Psi\rangle) \approx 0.66\frac{(d-1)^{1.84}}{N}, \tag{7.5.10}$$

$$\bar{\mathcal{I}}_2(|\Psi\rangle) \approx 1.6\frac{(d-1)}{N}. \tag{7.5.11}$$

7.6 Summary of results

Analyzing the infidelity plots for each ensemble division, we can conclude the following:

- From the plots of infidelity obtained for 4 different unknown states, we can say that the behavior of the infidelity is similar for all unknown states.
- The standard deviation is small at the final step (estimation) of the method, so we can infere that the behavior of the infidelity is similar for all unknown states. Except for the combination of small ensemble sizes and higher dimensions such as $N = 10^5$ with $d = 32, 64$ and $d = 64$ with $N = 10^5, 10^{5.5}$. In those cases, it is recommendable to use bigger ensamble sizes for optimal estimation with this method.
- Considering the points with small deviation, the average infidelity is close to the GMB.
- The average infidelity gets closer to the GMB from the first stage to the second, meaning that the first estimation is close enough to the real unknown state that the Fisher Symmetric measurement is optimal for the estimation at the final step.

Ensemble Division	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$
2/3	0.66	1.84	1
2/4	0.66	1.84	1
2/5	0.66	1.84	1

Table 7.6.1: Mean coefficients for the first stage.

Ensemble Division	$\bar{\alpha}$	$\bar{\beta}$	$\bar{\gamma}$
2/3	1.1	1.15	1
2/4	1.45	1	1
2/5	1.6	1	1

Table 7.6.2: Mean coefficients for the second stage.

We summarize the lineal fit results of the simulation data in the tables 7.6.1 and 7.6.2. In where we consider the mean values of the coefficients obtained. From these, we can conclude the following:

- In the First Stage, all the ensemble division have the same average performance in the estimation process. In where the mean infidelity can be aproximated as
$$\bar{\mathcal{I}}_1 \approx 0.66 \frac{(d-1)^{1.84}}{N}. \tag{7.6.1}$$

- At the Second Stage, all the ensemble divisions scales linearly with the inverse of the ensemble size,
$$\bar{\mathcal{I}}_2 \propto \frac{1}{N}. \tag{7.6.2}$$

- At the Second Stage, the division 2/3 scales with constant 1.1, less than the other two divisions. But, grows with the dimension with exponential 1.15, bigger than the other two that grows linearly with the dimension. Which, is best at higher dimensions.

- The division 2/4 grow with constant 1.45. The division 2/5 grows with constant 1.6. And both grows linearly with the dimension. So, between both the best ensemble division for the estimation is the division 2/4, with mean infidelity
$$\bar{\mathcal{I}}_2 \approx 1.45 \frac{d-1}{N}. \tag{7.6.3}$$

We can conclude from the simulations that, the infidelity obtained is very close to the Gill-Massar bound, meaning that the protocol is *near-optimal*. Also, for the best performance on infidelity, the distribution of the copies ensemble need to be such that for each stage $N/2$ copies are used.

Comparing the adaptive Fisher-symmetric method with the 5BB-QT (introduced in 5.4.5), we can highlight the following differences:

- The 5BB-QT method reconstructs the quantum state using measurements in 5 different bases. In contrast, the adaptive FSM employs a total of $7d-3$ measurements, which can be performed using 7 measurement bases.

- The infidelity of the 5BB-QT scales exponentially with the dimension, following a $d^{1.87}$ dependence. On the other hand, the infidelity of the adaptive FSM exhibits a linear growth with dimension, making it significantly more efficient in high-dimensional scenarios.

- For both methods, the infidelity scales linearly with the inverse of the

ensemble size, demonstrating similar behavior in this aspect.

- The prefactors governing the scaling with respect to dimension and ensemble size are 1.76 for the 5BB-QT and 1.45 for the adaptive FSM, indicating that the adaptive FSM achieves better scaling overall.

In summary, by increasing the number of measurement bases from 5 to 7, and using a two-stage adaptive protocol, the reconstruction is closer to the Gill-Massar bound.

Chapter 8

Conclusions and Outlook

The research presented in this thesis addresses the challenge of quantum pure state estimation using Fisher-symmetric measurements (FSM) and an adaptive estimation protocol. The method involves a projective measurement in a fixed basis, followed by three minimal Fisher-symmetric measurements, each consisting of $2d-1$ elements. The proposed method achieves an estimation infidelity that closely approximates the Gill-Massar bound, requiring no prior information beyond the purity of the quantum state. We also conclude that the optimal way to divide the total ensemble is by splitting it equally, half for each step. This protocol can be implemented by FSM via 2-basis, meaning that with $7d$ measurement outcomes, $2d$ more than the 5BB-QT, we achieve near optimality. There are several promising avenues for future exploration.

One natural extension would be to develop adaptive strategies aimed at minimizing estimation error for *mixed quantum states*. While the current protocol focuses on pure states, adapting it for mixed states would significantly broaden its applicability, especially in practical quantum information processing scenarios.

Additionally, this work primarily considers quantum states residing in discrete Hilbert spaces. It would be interesting to investigate whether Fisher-symmetric measurements can be generalized to *continuous Hilbert spaces*, and how such measurements could be applied to both pure and mixed states. This would represent a valuable step toward further generalizing the adaptive tomography protocol introduced here.

Another intriguing direction is the extension of the FSM-based adaptive technique

to *quantum process tomography*. Quantum tomography includes both state tomography and process tomography, the latter of which can be reformulated as a state estimation problem. Thus, it seems natural to explore whether the adaptive FSM method can be adapted to estimate unitary operations or quantum channels.

The experimental implementation of the proposed protocol also presents an important direction for future work. While the theoretical framework has been established, validating FSM-based adaptive tomography in *experimental settings* is essential for demonstrating its practical feasibility. This would involve assessing the protocol's robustness to experimental imperfections, such as noise and decoherence, which are inherent in real-world quantum systems.

In summary, this thesis provides a foundation for future research in adaptive quantum state estimation. Potential extensions include mixed-state tomography, the application of the method to continuous Hilbert spaces, quantum process tomography, and experimental implementations. These areas offer exciting opportunities for further exploration in the development of optimal quantum measurement strategies and their applications in advancing quantum technologies.

Appendix A

Simulation Code

In order to perform the simulations, we need to construct the elements of measurement, the FSMs. To do this, we follow the explicit construction given by Caves in [70], also shown in Section 6.2. This code is written in Python language.

```
import numpy as np

def FSM( d, phase=1 ): #with phase=1, the negative povm is build

    n = 2*d - 1   #number of elements
    #empty array of elements
    psi_k = np.zeros( (d, n), dtype=complex )

    psi_0 = np.ones(d, dtype=complex) #first element
    psi_0[1:] = - phase* np.exp( 1j*np.pi/4 ) #phases of the |k>
   states
    psi_0 = psi_0 / np.sqrt( n ) #normalization

    psi_k[:,0] = psi_0 #first element of the povm on the array

    for k in range(1,d):
    # odd elements
        psi = np.ones(d, dtype=complex)
        #phases of the |j> states
        psi[1:] = - phase*np.exp( 1j*np.pi/4 )/( np.sqrt(n) + 1 )
        #additional coef for the |k> state
        psi[k] += phase*np.sqrt(0.5*n)
        psi = psi / np.sqrt(n) #normalization

```

```
        #save the element on the array
        psi_k[:,2*k-1] = psi

    #even elements
        psi = np.ones(d, dtype=complex)
        #phases of the |j> states
        psi[1:] = - phase*np.exp( 1j*np.pi/4 )/( np.sqrt(n) + 1 )
        #additional coef for the |k> state
        psi[k] += phase*1j*np.sqrt(0.5*n)
        psi = psi / np.sqrt(n) #normalization

        #save the element on the array
        psi_k[:,2*k] = psi

    return psi_k
```

The two-stage adaptive FSMs protocol is as follows.

```
def TomoAnaliticaMLE(state,ensamble, shots_1 , shots_2):
    #state : unknown pure state
    #ensamble : number of copies of the state
    #shots_1: fraction of copies used in the first stage
    #shots_2: fraction of copies used in the second stage

    d = len(state) #dimension of the state

    #Step 0: measure the computational basis and choose the
    fiducial state
    BaseComp = np.eye(d, dtype=complex) #comp basis
    meas = SimMeas(state, BaseComp, shots=1) #meas comp basis

    #choose as fiducial state the detected state
    b0 = np.max(meas)
    fid = np.where(meas == b0, meas, 0)/b0
    # change of basis such that the fiducial state is |0>
    Base = BaseComp
    Base[:,0] = fid
    Base, R = np.abs(np.linalg.qr( Base )) #Gram-Schidth

    #Step 1:
    # buil FSM for fiducial state
    fsm_plus = FSM( d , -1 )
    fsm_plus_1 = Base@fsm_plus
```

```
25      measures_accumulated = fsm_plus_1
26
27      fsm_minus = FSM( d )
28      fsm_minus_1 = Base@fsm_minus
29      measures_accumulated = np.concatenate([measures_accumulated,
    fsm_minus_1], axis=1)
30
31      #simulate measures
32      probs_plus  = SimMeas(state, fsm_plus_1, shots_1*ensamble/2)
33      probs_minus = SimMeas(state, fsm_minus_1, shots_1*ensamble/2)
34
35      probs_accumulated = probs_plus
36      probs_accumulated = np.concatenate([probs_accumulated,
    probs_minus], axis=0)
37
38      #coefficients of the FSM elements
39      b0 = fsm_plus[0]
40      bk = fsm_plus[1:].real
41      ck = fsm_plus[1:].imag
42
43      # calculate the coefficients of the unknown state
44      Bk = np.sum( 0.5 * ( bk/b0 ) * ( probs_plus - probs_minus ),
    axis=1 )
45      Ck = np.sum( 0.5 * ( ck/b0 ) * ( probs_plus - probs_minus ),
    axis=1 )
46
47      angles = np.arctan2( Ck.real , Bk.real )
48
49      #Equation for b0^2
50      beta0 = np.zeros(2*d-1)
51
52      for i in range(2*d-1):
53          num = (b0[i])**2*np.sum(Bk**2 + Ck**2) - np.abs(np.sum((
    bk[:,i]-1j*ck[:,i])*(Bk + 1j*Ck)))**2
54          den = 4*(b0[i])**2 - 2*(probs_plus[i] + probs_minus[i])
55
56          beta0[i] = np.real(2*np.sqrt( num / den ))
57
58      mean_beta0 = np.mean(beta0) # choose the mean result of beta0
59
60      betak = np.sqrt( Bk**2 + Ck**2 ) / mean_beta0
61
```

```
62      #reconstruction of the state
63      state_0 = np.hstack( ( mean_beta0 , betak*np.exp( 1j*angles ) ))
64      state_hat = Base @ ( state_0 / np.linalg.norm( state_0 ) )
65
66      # refine the estimation via MLE
67      state_hat_est = MLE_pure( d, probs_accumulated, measures_accumulated, state_hat ) #agregar datos anteriores
68      Fid_1 = Fidelity(state_hat_est, state)
69
70      #Step 2: adapt the FSMs
71      Base = BaseComp
72      Base[:,0] = state_hat_est
73      Base, R = np.linalg.qr( Base ) #Gram-Schmidth
74      fsm = Base @ fsm_minus
75
76      measures_accumulated = np.concatenate([measures_accumulated, fsm], axis=1)
77
78      # simulated measure
79      probs_3 = SimMeas( state, fsm , shots_2*ensamble)
80      probs_accumulated = np.concatenate([probs_accumulated,probs_3], axis =0)
81
82      # MLE
83      state_est = MLE_pure( d, probs_accumulated, measures_accumulated, state_hat_est )
84
85      Fid_2 = Fidelity(state_est, state)
86
87      return Fid_1 , Fid_2
```

Bibliography

[1] C. Vargas, L. Pereira, and A. Delgado, "Near-optimal pure state estimation with adaptive fisher-symmetric measurements," 2024.

[2] D. J. Griffiths and D. F. Schroeter, *Introduction to Quantum Mechanics*, 3rd ed. Cambridge University Press, 2018.

[3] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[4] M. J. S. Morris H. DeGroot, *Probability and Statistics (4th Edition)*, 4th ed. Addison Wesley / Pearson, 2011.

[5] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, 1st ed., ser. Publications of the Scuola Normale Superiore / Monographs Scuola Normale Superiore. Scuola Normale Superiore Pisa, 2011.

[6] J. Bae and L.-C. Kwek, "Quantum state discrimination and its applications," *J. Phys. A: Math. Theor.*, vol. 48, no. 8, p. 083001, 2015, published 30 January 2015.

[7] S. L. Braunstein and C. M. Caves, "Statistical distance and the geometry of quantum states," *Phys. Rev. Lett.*, vol. 72, pp. 3439–3443, May 1994.

[8] S. Straupe, "Adaptive quantum tomography," *JETP Letters*, vol. 104, pp. 510–522, 2016.

[9] L. I. P. Valenzuela, "Adaptive standard quantum tomography in high dimensions," Master's thesis, Universidad de Concepción, 2018.

[10] J. Liu, H. Yuan, X.-M. Lu, and X. Wang, "Quantum fisher information matrix and multiparameter estimation," *Journal of Physics A: Mathematical and Theoretical*, vol. 53, no. 2, p. 023001, dec 2019.

[11] V. Giovannetti, S. Lloyd, and L. Maccone, "Quantum-enhanced measurements: Beating the standard quantum limit," *Science*, vol. 306, no. 5700, pp. 1330–1336, 2004.

[12] T. Nagata, R. Okamoto, J. L. O'Brien, K. Sasaki, and S. Takeuchi, "Beating the standard quantum limit with four-entangled photons," *Science*, vol. 316, no. 5825, pp. 726–729, 2007.

[13] R. Okamoto, H. F. Hofmann, T. Nagata, J. L. O'Brien, K. Sasaki, and S. Takeuchi, "Beating the standard quantum limit: phase super-sensitivity of n-photon interferometers," *New Journal of Physics*, vol. 10, no. 7, p. 073033, jul 2008.

[14] G. Y. Xiang, B. L. Higgins, D. W. Berry, H. M. Wiseman, and G. J. Pryde, "Entanglement-enhanced measurement of a completely unknown optical phase," *Nature Photonics*, vol. 5, no. 1, pp. 43–47, 2011.

[15] J. A. Jones, S. D. Karlen, J. Fitzsimons, A. Ardavan, S. C. Benjamin, G. A. D. Briggs, and J. J. L. Morton, "Magnetic field sensing beyond the standard quantum limit using 10-spin noon states," *Science*, vol. 324, no. 5931, pp. 1166–1168, 2009.

[16] S. Massar, "Collective versus local measurements on two parallel or antiparallel spins," *Phys. Rev. A*, vol. 62, p. 040101, Sep 2000.

[17] E. Bagan, M. A. Ballester, R. D. Gill, R. Muñoz Tapia, and O. Romero-Isart, "Separable measurement estimation of density matrices and its fidelity gap with collective protocols," *Phys. Rev. Lett.*, vol. 97, p. 130501, Sep 2006.

[18] R. D. Gill and S. Massar, "State estimation for large ensembles," *Phys. Rev. A*, vol. 61, p. 042312, Mar 2000.

[19] M. Muñoz, L. Pereira, S. Niklitschek, and A. Delgado, "Complex field formulation of the quantum estimation theory," 2022.

[20] Z. Huangjun, "Quantum state estimation and symmetric informationally complete poms," Ph.D. dissertation, National University of Singapore, 2012.

[21] C. W. Helstrom, "Quantum detection and estimation theory," *Journal of Statistical Physics*, vol. 1, no. 2, pp. 231–252, 1969.

[22] A. Czerwinski, "Selected concepts of quantum state tomography," *Optics*, vol. 3, no. 3, pp. 268–286, 2022.

[23] S. A. Metwalli and R. Van Meter, "Testing and debugging quantum circuits," *IEEE Transactions on Quantum Engineering*, vol. 5, pp. 1–15, 2024.

[24] J. Zhang, S. S. Hegde, and D. Suter, "Fast quantum state tomography in the nitrogen vacancy center of diamond," *Phys. Rev. Lett.*, vol. 130, p. 090801, Feb 2023.

[25] A. Farooq, U. Khalid, J. ur Rehman, and H. Shin, "Robust quantum state tomography method for quantum sensing," *Sensors*, vol. 22, no. 7, 2022.

[26] R. Jozsa, "Fidelity for mixed quantum states," *Journal of Modern Optics*, vol. 41, no. 12, pp. 2315–2323, 1994.

[27] H. Lange, M. Kebrič, M. Buser, U. Schollwöck, F. Grusdt, and A. Bohrdt, "Adaptive Quantum State Tomography with Active Learning," *Quantum*, vol. 7, p. 1129, Oct. 2023.

[28] Y. Quek, S. Fort, and H. K. Ng, "Adaptive quantum state tomography with neural networks," *npj Quantum Information*, vol. 7, no. 1, p. 105, 2021.

[29] D. H. Mahler, L. A. Rozema, A. Darabi, C. Ferrie, R. Blume-Kohout, and A. M. Steinberg, "Adaptive quantum state tomography improves accuracy quadratically," *Phys. Rev. Lett.*, vol. 111, p. 183601, Oct 2013.

[30] L. Pereira, L. Zambrano, J. Cortés-Vega, S. Niklitschek, and A. Delgado, "Adaptive quantum tomography in high dimensions," *Phys. Rev. A*, vol. 98, p. 012339, Jul 2018.

[31] G. I. Struchalin, E. V. Kovlakov, S. S. Straupe, and S. P. Kulik, "Adaptive quantum tomography of high-dimensional bipartite systems," *Phys. Rev. A*, vol. 98, p. 032330, Sep 2018.

[32] L. Pereira, D. Martínez, G. Cañas, E. S. Gómez, S. P. Walborn, G. Lima, and A. Delgado, "High-accuracy adaptive quantum tomography for high-dimensional quantum systems," 2020.

[33] Z. Hradil, "Quantum-state estimation," *Phys. Rev. A*, vol. 55, pp. R1561–R1564, Mar 1997.

[34] J. Shang, Z. Zhang, and H. K. Ng, "Superfast maximum-likelihood reconstruction for quantum tomography," *Phys. Rev. A*, vol. 95, p. 062336, Jun 2017.

[35] M. Paris and J. Rehacek, *Quantum State Estimation*, ser. Lecture Notes in Physics. Springer Berlin Heidelberg, 2004.

[36] Y. Teo, *Introduction To Quantum-state Estimation*. World Scientific Publishing Company, 2015.

[37] R. Bhatia, *Matrix Analysis*. Springer New York, NY, 1996.

[38] J. Řeháček, Z. Hradil, and M. Ježek, "Iterative algorithm for reconstruction of entangled states," *Phys. Rev. A*, vol. 63, p. 040303, Mar 2001.

[39] J. Řeháček, Z. c. v. Hradil, E. Knill, and A. I. Lvovsky, "Diluted maximum-likelihood algorithm for quantum tomography," *Phys. Rev. A*, vol. 75, p. 042108, Apr 2007.

[40] E. Prugovečki, "Information-theoretical aspects of quantum measurement," *International Journal of Theoretical Physics*, vol. 16, no. 5, pp. 321–331, 1977.

[41] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, "Symmetric informationally complete quantum measurements," *Journal of Mathematical Physics*, vol. 45, no. 6, pp. 2171–2180, 06 2004.

[42] T. Durt, C. Kurtsiefer, A. Lamas-Linares, and A. Ling, "Wigner tomography of two-qubit states and quantum cryptography," *Phys. Rev. A*, vol. 78, p. 042338, Oct 2008.

[43] Z. E. D. Medendorp, F. A. Torres-Ruiz, L. K. Shalm, G. N. M. Tabia, C. A. Fuchs, and A. M. Steinberg, "Experimental characterization of qutrits using symmetric informationally complete positive operator-valued measurements," *Phys. Rev. A*, vol. 83, p. 051801, May 2011.

[44] N. Bent, H. Qassim, A. A. Tahir, D. Sych, G. Leuchs, L. L. Sánchez-Soto, E. Karimi, and R. W. Boyd, "Experimental realization of quantum tomography of photonic qudits via symmetric informationally complete positive operator-valued measures," *Phys. Rev. X*, vol. 5, p. 041006, Oct 2015.

[45] W. M. Pimenta, B. Marques, T. O. Maciel, R. O. Vianna, A. Delgado, C. Saavedra, and S. Pádua, "Minimum tomography of two entangled qutrits using local measurements of one-qutrit symmetric informationally complete positive operator-valued measure," *Phys. Rev. A*, vol. 88, p. 012112, Jul 2013.

[46] J. Schwinger, "Unitary operator bases," *Proceedings of the National Academy of Sciences*, vol. 46, no. 4, pp. 570–579, 1960.

[47] I. D. Ivonovic, "Geometrical description of quantal state determination," *Journal of Physics A: Mathematical and General*, vol. 14, no. 12, p. 3241, dec 1981.

[48] W. K. Wootters and B. D. Fields, "Optimal state-determination by mutually unbiased measurements," *Annals of Physics*, vol. 191, no. 2, pp. 363–381, 1989.

[49] A. B. Klimov, C. Muñoz, A. Fernández, and C. Saavedra, "Optimal quantum-state reconstruction for cold trapped ions," *Phys. Rev. A*, vol. 77, p. 060303, Jun 2008.

[50] S. N. Filippov and V. I. Man'ko, "Mutually unbiased bases: tomography of spin states and the star-product scheme," *Physica Scripta*, vol. 2011, no. T143, p. 014010, feb 2011.

[51] R. B. A. Adamson and A. M. Steinberg, "Improving quantum state estimation with mutually unbiased bases," *Phys. Rev. Lett.*, vol. 105, p. 030406, Jul 2010.

[52] G. Lima, L. Neves, R. Guzmán, E. S. Gómez, W. A. T. Nogueira, A. Delgado, A. Vargas, and C. Saavedra, "Experimental quantum tomography of photonic qudits via mutually unbiased basis," *Opt. Express*, vol. 19, no. 4, pp. 3542–3552, Feb 2011.

[53] S. T. Flammia, A. Silberfarb, and C. M. Caves, "Minimal informationally complete measurements for pure states," *Foundations of Physics*, vol. 35, no. 12, pp. 1985–2006, 2005.

[54] D. Goyeneche, G. Cañas, S. Etcheverry, E. S. Gómez, G. B. Xavier, G. Lima, and A. Delgado, "Five measurement bases determine pure quantum states on any dimension," *Phys. Rev. Lett.*, vol. 115, p. 090401, Aug 2015.

[55] L. Zambrano, L. Pereira, and A. Delgado, "Improved estimation accuracy of the 5-bases-based tomographic method," *Phys. Rev. A*, vol. 100, p. 022340, Aug 2019.

[56] L. Zambrano, L. Pereira, D. Martínez, G. Cañas, G. Lima, and A. Delgado, "Estimation of pure states using three measurement bases," *Phys. Rev. Appl.*, vol. 14, p. 064004, Dec 2020.

[57] L. Pereira, L. Zambrano, and A. Delgado, "Scalable estimation of pure multi-qubit states," *npj Quantum Information*, vol. 8, p. 57, 2022.

[58] R. Okamoto, M. Iefuji, S. Oyama, K. Yamagata, H. Imai, A. Fujiwara, and S. Takeuchi, "Experimental demonstration of adaptive quantum state estimation," *Phys. Rev. Lett.*, vol. 109, p. 130404, Sep 2012.

[59] D. Mahler, L. A. Rozema, A. Darabi, J. Combes, C. Ferrie, R. Blume-Kohout, and A. M. Steinberg, "Experimental demonstration of adaptive tomography," in *Frontiers in Optics 2012/Laser Science XXVIII*. Optica Publishing Group, 2012, p. FTh4B.3.

[60] Z. Hou, H. Zhu, and G. e. a. Xiang, "Achieving quantum precision limit in adaptive qubit state tomography." *npj Quantum Inf 2, 16001* ., 2016.

[61] D. G. Fischer, S. H. Kienle, and M. Freyberger, "Quantum-state estimation by self-learning measurements," *Phys. Rev. A*, vol. 61, p. 032306, Feb 2000.

[62] T. Hannemann, D. Reiss, C. Balzer, W. Neuhauser, P. E. Toschek, and C. Wunderlich, "Self-learning estimation of quantum states," *Phys. Rev. A*, vol. 65, p. 050303, May 2002.

[63] F. Huszár and N. M. T. Houlsby, "Adaptive bayesian quantum tomography," *Phys. Rev. A*, vol. 85, p. 052120, May 2012.

[64] A. Utreras-Alarcón, M. Rivera-Tapia, S. Niklitschek *et al.*, "Stochastic optimization on complex variables and pure-state quantum tomography," *Scientific Reports*, vol. 9, p. 16143, 2019.

[65] L. Zambrano, L. Pereira, and S. e. a. Niklitschek, "Estimation of pure quantum states in high dimension at the limit of quantum accuracy through complex optimization and statistical inference." *Scientific Reports*, vol. 10, 2020.

[66] M. Rambach, M. Qaryan, M. Kewming, C. Ferrie, A. G. White, and J. Romero, "Robust and efficient high-dimensional quantum state tomography," *Phys. Rev. Lett.*, vol. 126, p. 100402, Mar 2021.

[67] A. Kalev and I. Hen, "Fidelity-optimized quantum state estimation," *New Journal of Physics*, vol. 17, no. 9, p. 093008, sep 2015.

[68] C. Granade, C. Ferrie, and S. T. Flammia, "Practical adaptive quantum tomography," *New Journal of Physics*, vol. 19, no. 11, p. 113017, nov 2017.

[69] J. Řeháček, B.-G. Englert, and D. Kaszlikowski, "Minimal qubit tomography," *Phys. Rev. A*, vol. 70, p. 052321, Nov 2004.

[70] N. Li, C. Ferrie, J. A. Gross, A. Kalev, and C. M. Caves, "Fisher-symmetric informationally complete measurements for pure states," *Phys. Rev. Lett.*, vol. 116, p. 180402, May 2016.

[71] M. Hayashi, *Asymptotic Theory of Quantum Statistical Inference*. WORLD SCIENTIFIC, 2005.

[72] A. Fujiwara, "Strong consistency and asymptotic efficiency for adaptive quantum estimation problems," *Journal of Physics A: Mathematical and General*, vol. 39, no. 40, p. 12489, sep 2006.

[73] C. Ferrie, "Self-guided quantum tomography," *Phys. Rev. Lett.*, vol. 113, p. 190404, Nov 2014.

[74] A. Farooq, M. A. Ullah, J. ur Rehman, K. Lee, and H. Shin, "Self-guided quantum state learning for mixed states," *Quantum Information Processing*, vol. 21, no. 7, p. 243, 2022.

[75] H. Zhu and M. Hayashi, "Universally fisher-symmetric informationally complete measurements," *Phys. Rev. Lett.*, vol. 120, p. 030404, Jan 2018.

[76] F. Mezzadri, "How to generate random matrices from the classical compact groups," *Notices of the American Mathematical Society*, vol. 54, no. 5, pp. 592 – 604, May 2007.