ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea Magistrale in Fisica

An adaptive scheme for quantum state tomography

Relatore:
Prof.ssa Elisa Ercolessi

Presentata da:
Valerio Crescimanna

Correlatore:
Prof. Harald Weinfurter

Anno Accademico 2018/2019

ii

Sommario

Il processo di determinazione e ricostruzione dello stato di un sistema quantistico a partire dai risulati delle misure, meglio noto come tomografia quantistica, svolge un ruolo cruciale nel campo emergente delle tecnologie quantistiche. Oggi è possibile controllare sperimentalmente sistemi quantistici contenenti decine di qubit e realizzare misure di qualsiasi osservabile con grande precisione. Tuttavia, la tomografia quantistica richiede un numero di misure esponenzalmente grande in n per caratterizzare completamente uno stato incognito con n qubit. Una possibile soluzione a questo problema consiste nel realizzare una tomografia incompleta che sia in grado di fornire una buona stima dello stato sotto osservazione con poche misure. Questa tesi propone uno schema tomografico per stati di n qubit che ha lo scopo di migliorare la fedeltà tra lo stato ricostruito e quello effettivamente sotto osservazione. Lo schema, in particolare, individua quale misura realizzare a partire dalla conoscenza dello stato già acquisita con le misure precedentemente realizzate. Le prestazioni di questo schema sono state infine analizzate tramite simulazioni di tomografia quantistica con misure proiettive sia su stati separabili che su stati entangled. In entrambi i casi si osserva che lo schema adattivo qui proposto supera uno schema tomografico standard in termini di fedeltà dello stato riscostruito.

Abstract

The process of inferring and reconstructing the state of a quantum system from the results of measurements, better known as quantum state tomography, constitutes a crucial task in the emerging field of quantum technologies. Today it is possible to experimentally control quantum systems containing tens of entangled qubits and perform measurements of arbitrary observables with great accuracy. However, in order to complete characterize an unknown n-qubit state, quantum state tomography requires a number of measurements which grows exponentially with n. A possible way to avoid this problem consists in performing an incomplete tomographic procedure able to provide a good estimate of the true state with few measurements. This thesis proposes a scheme for n-qubit state tomography which aims to improve the fidelity between the reconstructed state and the target state. In particular, the scheme identifies the next measurement to perform based on the knowledge already acquired from the previous measurements on the experimental prepared state. The performance of this scheme was finally analyzed by means of simulations of quantum state tomography with product measurements as well as with entangled measurements. In both cases one observes that the here proposed adaptive scheme significantly outperforms a standard scheme in terms of the fidelity of the reconstructed state.

Contents

Introduction

Quantum information technologies have witnessed huge evolution in the last years. From the early stage of thought experiments they have developed into almost ready-to-use technology. Today, the wide range of prospective applications includes quantum computing, quantum cryptography, and quantum communications. In view of the many possible applications the question of efficient analysis of quantum systems is a fundamental task but an efficient characterization of quantum systems is still an unsolved problem. One of the major challenge concerns state estimation by which the unknown state of a quantum system is determined through a succession of different measurements. The quantum state is not observable in the same way classical states are [1]. First, one cannot perform multiple measurements on the same state [2] since a measurement destroys the information of the state, or at best perturbs it [3]. Second, it is not possible to clone the state one wants to study, so the measurements cannot be performed on exact copies of a state [4]. Moreover, a state is generally unknown and in principle one does not know whether it is pure or mixed or it fulfills some peculiar properties, therefore the best choice of basis and measurement sequence is not evident from the beginning [5, 6, 7]. In general the measurements are described by a set of positive operators. The measurement data obtained are then use to infer the quantum state of the system. Due to the formal similarities with medical non-invasive three-dimensional imaging, quantum state estimation is often referred as to quantum tomography. In this thesis the terms quantum state estimation and quantum state tomography are going to be used interchangeably both to refer to estimating the state using incomplete information, and to describe the situation where complete and possibly noise-free information about the state is assumed. Quantum tomography is a relevant tool in quantum information science being necessary for characterization of quantum states, gates, and measurement apparatuses. Actually, many of the notable advances in the field including demonstrations of entanglement of two [8], three [9], and four-photon states [10], quantum logic gate characterization [11], and cluster state quantum computing [12] used quantum state tomography as the main diagnostic and descriptive tool [13]. The first theoretical demonstration that the quantum state of light could be reconstructed from a series of measurements data arrived already 20 years ago thanks to Vogel and Risken [14]. Soon after, from the first experimental realization [15] many different methods of quantum tomography were proposed and implemented for various physical systems. Quantum state

1

tomography has by now been applied to nearly all candidate systems proposed for quantum information and computation including spontaneous parametric down conversion sources [16, 17], trapped ions [18], atomic ensemble quantum memories [19], atoms trapped in optical lattices [20], cavity QED systems [21], quantum dot-sources of entangled photons [22] and superconducting quantum bits [23]. Together with the progress on the experimental detection of the state to identify even the smallest and feeblest non-classical effects [24, 25], new mathematical algorithms have been proposed in order to improve the quality and the speed of the tomographic procedures. In fact, the original linear methods [26] based on the inverse transformation does not guarantee the positivity of the reconstructed state required by quantum theory. For this reason, in spite of its simplicity, the linear inversion is often replaced by many others statistically motivated methods. In particular, some of the most common reconstruction methods are least squares regression [27], maximum likelihood estimation [28, 29], methods based on compressed sensing [30, 31], and the Bayesian approach [32]. All these estimator, indeed, are built to yield a physical state, but in this thesis a particular emphasis is given to maximum likelihood. In fact it is regarded as an optimal method since it provides the physical state which maximizes the probability to find the observed measurement outcomes. On the other side, the disadvantage of maximum likelihood estimation is that it often provides rank deficient states. The frequentist's definition of classical estimation constitutes the leading idea of quantum estimation, a well-defined physical state is assigned for each measured quantum system which is prepared by the source [33]. A finite number of measurements on the copies of the system, does not allow the observer to reconstruct unambiguously the state. In general another observer would reconstruct another state after measuring his own copies [34]. The two estimates would start to approach each other and the target state as the number of measurements approaches infinity. However, the number of measurements to perform complete tomography, i.e., to reconstruct unambiguously the state up to statistical error, grows exponentially with the number of qubits or in general the size of the system. This means that even for a few qubits, full tomography is a resource-intensive procedure which becomes unfeasible when one wants to study many qubit states. In order to overcome this problem one might choose to perform incomplete tomography and reconstruct the state which approximates the target one as good as possible after a few measurements.

This thesis proposes an experimental scheme which aims to achieve this goal by selecting cleverly in an adaptive manner which measurements to use during the experiment. In particular this adaptive procedure has been compared both with the standard scheme with product measurements and with a scheme employing mutually unbiased bases measurement.

Regarding the structure of this thesis, the first chapter is devoted to an overview of the quantum mechanics formalism employed. The second chapter describes more in detail quantum state tomography with a brief presentation of most common tomographic schemes and estimation procedures. Finally in the third chapter the adaptive scheme is described and results are compared with those obtained with standard schemes.

Chapter 1

Quantum Formalism

In this chapter the basic formalism used in the rest of the thesis is introduced. In particular the concepts of quantum state and density matrix are presented as well as the definitions of fidelity and purity functions. Finally, the second section is devoted to the description of measurements on quantum states.

1.1 Quantum bits

The quantum bit or qubit, for short, in quantum computing is the counterpart to the binary digit of classical computing. From the physical point of view photons, spin$-\frac{1}{2}$ particles [35, 36] and two level atoms [37, 38] are common examples of two state quantum-mechanical systems able to encode the properties of qubits. The two basis states a qubit can assume are often referred to as $|0\rangle$ and $|1\rangle$, which straightforwardly corresponds to the logical levels 0 and 1 usually considered for a classical bit. Unlike the classical bit however, a quantum bit can be in an infinite number of others states than $|0\rangle$ and $|1\rangle$, that are all the coherent superpositions of these states. A general superposition state $|\psi\rangle$ in fact is given by the linear combination of the states $|0\rangle$, $|1\rangle$

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle, \qquad (1.1)$$

where α and β are complex numbers. From (1.1) one can view the general superposition state $|\psi\rangle$ as a vector in the two-dimensional Hilbert space $\mathcal{H}_2$ generated by the states $|0\rangle$ and $|1\rangle$. The basis with elements $|0\rangle$ and $|1\rangle$ is called computational basis. When one measures the state of the qubit $|\psi\rangle$ in the computational basis, according to the Born rule [39] one finds that the probability of outcome $|0\rangle$ is $|\alpha|^2 = |\langle 0|\psi\rangle|^2$ and the probability of outcome $|1\rangle$ is $|\beta|^2 = |\langle 1|\psi\rangle|^2$. Since probabilities have to sum to 1, the so-called probability amplitudes α and β have to fulfill the constraint

$$|\alpha|^2 + |\beta|^2 = 1. \qquad (1.2)$$

From (1.1) and (1.2), without loss of generality, one can rewrite $|\psi\rangle$ as

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle,\qquad(1.3)$$

with $\theta \in [0,\pi]$ and $\varphi \in [0,2\pi]$. The complex part of the probability amplitude associated to $|0\rangle$ is factorized in an overall phase which, in turn is ignored since it has no observable effects. The variables θ and φ just defined, unambiguously identify a point on the surface of a unitary three-dimensional sphere, which in this context is called Bloch sphere. Each point on the surface of the Bloch sphere can be then put in biunivocal relation with one specific state $|\psi\rangle$. As an example, in Fig.1.1 the state $|\psi\rangle$ of (1.3) is represented as a unit vector of polar coordinates θ and φ.
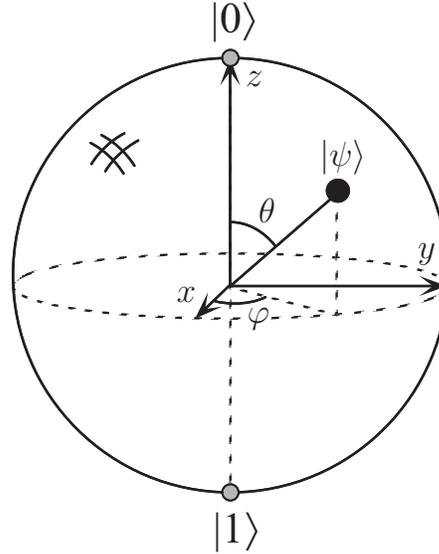

Figure 1.1: Bloch sphere representation of a pure state $|\psi\rangle$ (from [40].)

So far the discussion considered only single qubit states, but one might be interested in systems which are a collection of many two-level systems. In such a case the description made for single qubit states generalizes quite straightforwardly to more qubits. In the same way as for the single qubit case, for n qubit states one can define a computational basis in the Hilbert space $\mathcal{H}_n$. The elements of such a basis are built from all possible combinations of tensor products of n single qubits states $|0\rangle$ and $|1\rangle$, which form the computational basis of $\mathcal{H}_1$. It follows that the cardinality of the n-qubit basis is 2^n. For example, when $n = 2$ the $2^2 = 4$ basis elements are $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$, and a general two qubit superposition state reads

$$|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle,\qquad(1.4)$$

where the normalization constraint (1.2) generalizes for probability amplitudes α_{ij} to $\sum_{i,j=0}^{1} |\alpha_{ij}|^2 = 1$. For the sake of brevity, the tensor product of 2 qubit states was here and will be from now on indicated as $|ij\rangle \equiv |i\rangle \otimes |j\rangle$. Even though the Bloch sphere representation (1.3) is no longer usable for multiple qubits, one can keep in mind that $\mathcal{H}_n \equiv \mathbb{C}^{2^n}$. Each state in $\mathcal{H}_n$ can be unambigously identified with a vector in the compex coordinate space $\mathbb{C}^{2^n}$. For example one can identify the two single qubit basis elements as

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.5}$$

From (1.5) it is easy to parameterize every single qubit superposition state like (1.1) as

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{1.6}$$

The vectorial formalism is particularly useful since it makes evident how quantum operator act on physical states. In this representation in fact, each operators corresponds to a $2^n \times 2^n$ matrix and the application of such an operator on a state reduces to a product between a matrix and a column vector. As an example, starting again in single qubit state space $\mathcal{H}_1$, one might consider Pauli operators which play a fundamental role in quantum mechanics and throughout this thesis. They are defined as

$$I = |0\rangle\langle 0| + |1\rangle\langle 1|, \tag{1.7}$$
$$\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0|, \tag{1.8}$$
$$\sigma_y = i\left(|1\rangle\langle 0| - |0\rangle\langle 1|\right), \tag{1.9}$$
$$\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|, \tag{1.10}$$

but can be expressed in matricial form in the basis (1.5) as

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.11}$$

Pauli operators are Hermitian operator and fulfill the following commutation relations

$$[\sigma_i, \sigma_j] = \sigma_i\sigma_j - \sigma_j\sigma_i = 2\epsilon_{ijk}\sigma_k, \quad \{\sigma_i, \sigma_j\} = \sigma_i\sigma_j + \sigma_j\sigma_i = 2I\delta_{ij}. \tag{1.12}$$

with δ_{ij} the Kronecker symbol and ϵ_{ijk} Levi-Civita symbol. Moreover since the single qubit operator algebra is four dimensional and Pauli operators are four linearly independent operators, one can conclude that they form a basis of the operator space. Namely every operator A can be expressed as linear combination of Pauli operators.

$$A = a_0 I + \boldsymbol{a} \cdot \boldsymbol{\sigma}, \tag{1.13}$$

where $\boldsymbol{a} = (a_x, a_y, a_z)$ and $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. Moving from single qubit to n qubit states, one needs to consider a tensor product of n Pauli operators $\sigma_j \otimes \sigma_k \otimes \cdots \otimes \sigma_l$, from now on denoted by $JK\cdots L$ to not burden the notation.

Density matrix   The implicit assumption made so far was that just one pure superposition state $|\psi\rangle$ as for example defined in Eq. (1.3) for a single qubit describes the system completely. However this assumption does not hold in general. In fact a system can have a certain probability p_1 to be in the state $|\psi_1\rangle$, a probability p_2 to be in another state $|\psi_2\rangle$ and in general a probability p_i to be in the state $|\psi_i\rangle$. In such a case the system is usually described by the so-called density operator ρ

$$\rho = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \,, \tag{1.14}$$

with $\sum_i p_i = 1$, while the collection of the states with corresponding probabilities $\{p_i, |\psi_i\rangle\}$ is referred to as ensemble of states. The density operator or density matrix is positive by definition since its eigenvalues are the probabilities $p_i \geq 0$

$$\rho \geq 0. \tag{1.15}$$

Moreover ρ is Hermitian

$$\rho^\dagger = \sum_i p_i^* \left(|\psi_i\rangle\langle\psi_i|\right)^\dagger = \rho, \tag{1.16}$$

and has trace 1

$$
\begin{aligned}
\mathrm{tr}(\rho) &= \sum_{x \in \{0,1\}^n} \langle x| \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \, |x\rangle = \sum_i p_i \sum_{x \in \{0,1\}^n} \langle x|\psi_i\rangle \, \langle\psi_i|x\rangle \\
&= \sum_i p_i \sum_{x \in \{0,1\}^n} c_{x,i} c_{x,i}^* = \sum_i p_i = 1.
\end{aligned}
\tag{1.17}
$$

Here $|x\rangle$ denotes the element of the computational basis for n qubit states, while $c_{x,i}$ is the corresponding probability amplitudes for the $i-$th state $|\psi_i\rangle$. As any other operator, also the single qubit density matrix can be expressed as linear combination of Pauli operators,

$$\rho = \frac{1}{2} \left(I + T_x \sigma_x + T_y \sigma_y + T_z \sigma_z\right) = \frac{1}{2} \left(I + \boldsymbol{r} \cdot \boldsymbol{\sigma}\right) \tag{1.18}$$

with the so-called Stokes parameters or correlation values T_x, T_y, T_z such that $\boldsymbol{r} = (T_x, T_y, T_z) \in \mathbb{R}^3$ and $\|\boldsymbol{r}\|^2 = T_x^2 + T_y^2 + T_z^2 \leq 1$ in order to fulfill the constraints (1.16) and (1.17). Thanks to (1.18) one can see how it is possible to associate unambiguously every single qubit state to a vector $\boldsymbol{r}$. In the Bloch sphere representation this means that every state ρ is associated to a point of the sphere and that one recovers the already studied case of pure superposition states when $\|\boldsymbol{r}\|^2 = T_x^2 + T_y^2 + T_z^2 = 1$. The generalization of these results to more qubits is pretty straightforward; for n qubits (1.18) generalizes to

$$\rho = \frac{1}{2^n} \sum_{\mu_1,\dots\mu_n \in \{0,1,2,3\}} T_{\mu_1,\dots\mu_n} \sigma_{\mu_1} \otimes \cdots \otimes \sigma_{\mu_n} \tag{1.19}$$

with correlation parameters $T_{\mu_1,\dots\mu_n}$ and index $\mu_i = 0, 1, 2, 3$ respectively for the Pauli operator $I, \sigma_x, \sigma_y, \sigma_z$ acting on the $i-$th qubit. The correlation elements corresponding to the identity n qubit operator is fixed to be one $T_{0\dots0} = 1$ by the trace condition (1.17). All the 4^n correlations can be also arranged to form a $4 \times 4 \times \cdots \times 4$ tensor which is called correlation tensor. The inverse relation which allows to recover any correlation element $T_{\mu_1,\dots\mu_n}$ from the state ρ is

$$T_{\mu_1,\dots\mu_n} = \text{tr}\left(\sigma_{\mu_1} \otimes \cdots \otimes \sigma_{\mu_n}\rho\right) \equiv \text{tr}\left(\sigma_{\boldsymbol{\mu}}\rho\right), \tag{1.20}$$

Purity and fidelity Once the density matrix is defined, it is possible to define two related quantities which are commonly employed for a description of the states. These quantities are the purity and the fidelity. Given a state ρ the purity of this state γ is defined as

$$\gamma = \text{tr}\left(\rho^2\right), \tag{1.21}$$

which, for single a qubit state, is conveniently related to the Bloch lengh of the vector $\boldsymbol{r}$ thanks to (1.18) with

$$\gamma = \text{tr}\left[\frac{1}{4}\left(I + \boldsymbol{r} \cdot \boldsymbol{\sigma}\right)^2\right] = \frac{1}{4}\text{tr}\left[I + 2\left(\boldsymbol{r} \cdot \boldsymbol{\sigma}\right) + \|\boldsymbol{r}\|^2 I\right] = \frac{1}{2}\left(1 + \|\boldsymbol{r}\|^2\right). \tag{1.22}$$

Then, for single qubits, $\frac{1}{2} \leq \gamma \leq 1$, while in general $\frac{1}{2^n} \leq \gamma \leq 1$ with n being the number of qubits. A state is called pure ρ_p if it has purity one $\text{tr}\left(\rho_p^2\right) = 1$ and it is called mixed ρ_m otherwise $\text{tr}\left(\rho_m^2\right) < 1$; states with purity $\gamma = \frac{1}{2^n}$ are sometimes referred to as maximally mixed. Finally one can observe that superposition states as (1.1) and (1.4) are pure by definition

$$\gamma_s = \text{tr}\left(|\psi\rangle\langle\psi|\psi\rangle\langle\psi|\right) = \sum_{x \in \{0,1\}^n} \langle x|\psi\rangle\langle\psi|x\rangle = \sum_{x \in \{0,1\}^n} c_{x,i}c_{x,i}^* = 1, \tag{1.23}$$

with $x, c_{x,i}$ defined as in (1.17).

The fidelity is a measure of the distance between two states. The fidelity between a state ρ and a state σ is defined as

$$F(\rho, \sigma) = \left(\text{Tr}\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right)^2. \tag{1.24}$$

The fidelity between two states is symmetric under exchange of the arguments $F(\rho, \sigma) = F(\sigma, \rho)$ and $F(\rho, \sigma) = 1$ iff $\rho = \sigma$. If one of the arguments is a pure state then the fidelity becomes

$$F(|\psi\rangle\langle\psi|, \rho) = \left(\text{Tr}\sqrt{|\psi\rangle\langle\psi|\rho|\psi\rangle\langle\psi|}\right)^2 = \langle\psi|\rho|\psi\rangle. \tag{1.25}$$

That is the fidelity is equal to the square root of the overlap between $|\psi\rangle$ and ρ. It follows that if both arguments are pure the fidelity further reduces to

$$F(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = |\langle\phi|\psi\rangle|^2. \tag{1.26}$$

As well as purity (1.22), one can express the fidelity between tho single qubit states in terms of Bloch vectors. In fact, given two states ρ and σ with corresponding Bloch vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, the fidelity reads [41]

$$F(\rho(\boldsymbol{u}), \sigma(\boldsymbol{v})) = \frac{1}{2}\left(1 + \boldsymbol{u} \cdot \boldsymbol{v} + \sqrt{1 - \|\boldsymbol{u}\|}\sqrt{1 - \|\boldsymbol{v}\|}\right). \tag{1.27}$$

1.2   Quantum measurement

Projective measurements or von Neumann measurements are the simplest form of quantum measurements [42]. By spectral theorem [43], for all Hermitian operator A, there exist a complete set of orthonormal eigenstates $|\lambda_n\rangle$ with eigenvalues λ_n such that A can be expressed through its spectral decomposition as

$$A = \sum_n \lambda_n |\lambda_n\rangle\langle\lambda_n|. \tag{1.28}$$

The probability that a measurement of A of state ρ provides λ_n as outcome is

$$P_{|\lambda_n\rangle} = \langle\lambda_n|\rho|\lambda_n\rangle = \mathrm{Tr}\left(\rho |\lambda_n\rangle\langle\lambda_n|\right) \equiv \mathrm{Tr}\left(\rho P_n\right). \tag{1.29}$$

The projector $P_n = |\lambda_n\rangle\langle\lambda_n|$, defined in the last equivalence of (1.29), generalizes for m degenerate eigenstates $\left|\lambda_n^1\right\rangle, \ldots, |\lambda_n^m\rangle$ as

$$P_n = \sum_{j=1}^{m} \left|\lambda_n^j\right\rangle\!\left\langle\lambda_n^j\right|. \tag{1.30}$$

One can prove that projection operators just defined, fulfill the following properties

$$\text{Hermiticity} \qquad\qquad P_n^\dagger = P_n; \tag{1.31}$$

$$\text{positiveness} \qquad\qquad P_n \leq 0; \tag{1.32}$$

$$\text{completness} \qquad\qquad \sum_n P_n = I; \tag{1.33}$$

$$\text{orthonormality} \qquad\qquad P_i P_j = P_i \delta_{ij}. \tag{1.34}$$

Now it is possible to see how the correlations $T_{\boldsymbol{\mu}}$ presented in (1.19) can be effectively measured from the state. Let's consider first single qubit correlations T_0, T_x, T_y, T_z. From (1.20)

$$T_0 = \mathrm{Tr}\left[\rho\left(|0\rangle\langle0| + |1\rangle\langle1|\right)\right] = P_{|0\rangle} + P_{|1\rangle}; \tag{1.35a}$$

$$T_x = \mathrm{Tr}\left[\rho\left(|0\rangle\langle1| + |1\rangle\langle0|\right)\right] = P_{\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)} - P_{\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)}; \tag{1.35b}$$

$$T_y = \mathrm{Tr}\left[\rho\left(i\,|1\rangle\langle0| - i\,|0\rangle\langle1|\right)\right] = P_{\frac{1}{\sqrt{2}}(|0\rangle+i|1\rangle)} - P_{\frac{1}{\sqrt{2}}(|0\rangle-i|1\rangle)}; \tag{1.35c}$$

$$T_z = \mathrm{Tr}\left[\rho\left(|0\rangle\langle0| - |1\rangle\langle1|\right)\right] = P_{|0\rangle} - P_{|1\rangle}; \tag{1.35d}$$

with Pauli operators $I, \sigma_x, \sigma_y, \sigma_z$ substituted with (1.7). One can conclude that correlations are derived by the projection outcomes of eigenstates of Pauli operators. In fact if one defines

$$|\pm\rangle \equiv \frac{1}{2}(|0\rangle \pm |1\rangle), \quad \text{and} \quad |\pm i\rangle \equiv \frac{1}{2}(|0\rangle \pm i|1\rangle), \tag{1.36}$$

then

$$\sigma_x |+\rangle = +|+\rangle, \quad \sigma_x |-\rangle = -|-\rangle, \tag{1.37a}$$
$$\sigma_y |+i\rangle = +|+i\rangle, \quad \sigma_y |-i\rangle = -|-i\rangle, \tag{1.37b}$$
$$\sigma_z |0\rangle = +|0\rangle, \quad \sigma_z |1\rangle = -|1\rangle. \tag{1.37c}$$

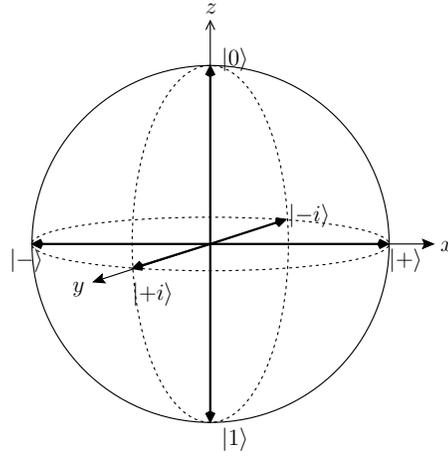The six eigenstates are displayed with Bloch sphere representation in Fig. 1.2

Figure 1.2: Pauli eigenstates in Bloch sphere representation (from [44]).

Experimental realization In this thesis polarization encoding for photons is mainly used to show experimental realization of tomographic measurements. Namely, horizontal polarization state $|H\rangle$ and vertical polarization state $|V\rangle$ play the role of two level of the quantum state such that

$$|H\rangle = |0\rangle, \quad |V\rangle = |1\rangle. \tag{1.38}$$

It follows that diagonal $|D\rangle$, antidiagonal $|A\rangle$, left $|L\rangle$ and right $|R\rangle$ polarizations read

$$|D\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |A\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \tag{1.39}$$

$$|R\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), \quad |L\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle). \tag{1.40}$$

States of (1.39) are equivalent to the eigenstates of Pauli operators σ_x and σ_y of (1.37) with $|D\rangle \equiv |+\rangle$, $|A\rangle \equiv |-\rangle$, $|R\rangle \equiv |+i\rangle$, $|L\rangle \equiv |-i\rangle$. The required experimental set-up to perform projection measurements on photons in polarization encoding consits of an half wave plate, a polarizing beam splitter and a single photon counting module. The incoming photon passes in the order through the half wave plate, the quarter wave plate, the polarizing beam splitter and is finally detected by the photon counting module. The whole apparatus is shown in Fig. 1.3. The function of wave plates is that of changing the polarization of incoming photons. The half wave plate rotates the Bloch vector defining the state of the photon by $180°$ around an axis which lies in the same plane of $|H\rangle$ and $|+\rangle$, tilted by an angle 2α respect to $|H\rangle$ state, with α tilt angle of the half wave plate itself. The $0°$ position is defined as the position where horizontal and vertical polarizations remain unchanged since the rotation axis concurs with the z-axis in the Bloch Sphere.

The experimental set-up one needs to measure a singe qubit encoded in such a way consists of a polarizing beamsplitter, an half wave plate HWP, a quarter wave plate QWP and a single photon counting module. The role of polarizing beamsplitter is to split the incoming unpolarized beam of photons into two outcoming polarized beams respectively with horizontal and vertical polarization. The operator for an half wave plate reads

$$\text{HWP}(\alpha) = \begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix}. \tag{1.41}$$

Similarly a quarter wave plate rotates by the Bloch vector $90°$ around an axis defined in the same way as the half wave plate. The operator associated to the quarter wave plate then is

$$\text{QWP}(\alpha) = \begin{pmatrix} \cos^2(\alpha) - i\sin^2(\alpha) & (1+i)\cos(\alpha)\sin(\alpha) \\ (1+i)\cos(\alpha)\sin(\alpha) & -i\cos^2(\alpha) + \sin^2(\alpha) \end{pmatrix}. \tag{1.42}$$

Experimental set-up including single photon counting module is shown in Fig. 1.3 while Fig. 1.4 illustrates action of wave plates on Bloch vector.

The angle α employed for transformations of Fig. 1.4, indeed are chosen to transform the eigenstates $|H\rangle$ and $|V\rangle$ of σ_z to the eigenstates of σ_x and σ_y. In fact this choice for α allows to measures the correlations as defined in (1.35). In Table 1.1 the waveplate angles for measuring the Stokes parameters using the setup shown in Fig. 1.4 are listed.

Stokes parameter	HWP	QWP
T_x	$22.5°$	$0°$
T_y	$0°$	$45°$
T_z	$0°$	$0°$

Table 1.1: Values of the angle α, one needs to rotate the crystal of HWP and QWP in order to measure the Stokes parameter T_x, T_y, T_z.
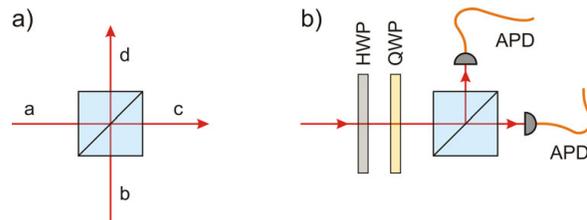
Figure 1.3: a) Polarizing beam splitter, with two incoming beams a and b, and two outcoming beams c and d. b) Half wave plate (gray) and quarter wave plate (yellow) acting on the incoming beam. The beam subsequently passes through a polarizing beam splitter. Finally orthogonal photons are sent to avalanche photo diodes to detect their polarization (from [45]).
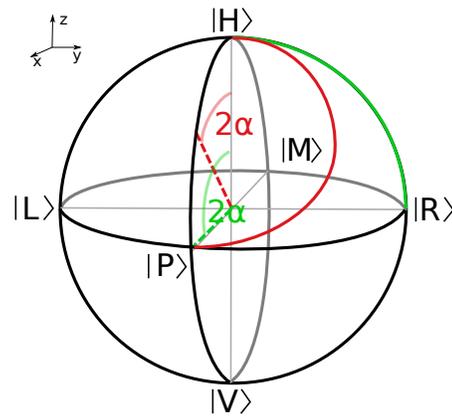

Figure 1.4: Action of half wave plate (red) and quarter waveplate (green), with tilt angle $\alpha = \pi$ on the state $|H\rangle$. Resulting states are $|P\rangle$ for the half wave plate and $|R\rangle$ for the quarter wave plate (from[46]).

Chapter 2

Quantum tomography

The knowledge of the state ρ is indeed the most complete knowledge one can gain about a physical system. One very important task in quantum mechanics consists in inferring the physical state through multiple measurements on the quantum system. In general, one cannot reconstruct completely the state with a single measurement operation. Moreover, while in classical physics it is always possible, at least in principle, to devise a procedure consisting of multiple measurements that fully recovers the physical state of each system, in quantum mechanics the same statement is no longer true. The impossibility to recover all the properties of a state by a finite amount of measurements is inherently related to fundamental features of quantum theory, namely its linearity and the Heisenberg uncertainty principle. On one hand linearity implies the no-cloning theorem [4], which forbids to create perfect copies of an arbitrary system, which one might desire to have in order to make multiple measurements on the same state. On the other hand, the uncertainty principle [47] states that one cannot perform an arbitrary sequence of measurements on a single system without disturbing it by inducing a back-action which modifies the state itself. Therefore, there is no way, even in principle, by which a quantum state of a single system may be inferred with arbitrary accuracy without having some prior knowledge on it [2]. In fact, in order to estimate the quantum state of a system one has to perform different measurements on many copies of the same statistical ensemble; in such a way one infers a little information at a time by probing a particular aspect of the state [48]. The standard approach for collecting information about a quantum state which works through a succession of projection measurements is known as quantum state tomography. Just like tomography scans in medicine each unknown quantum state is completely characterized by retrieving the information of a new dimension after each measurement in a different basis [49]. A classical representation for the tomographic process is given in Fig. 2.1. Estimating the state through successive projection measurements is indeed akin to building up an image of a complex object by making only simple projections of its shadow.

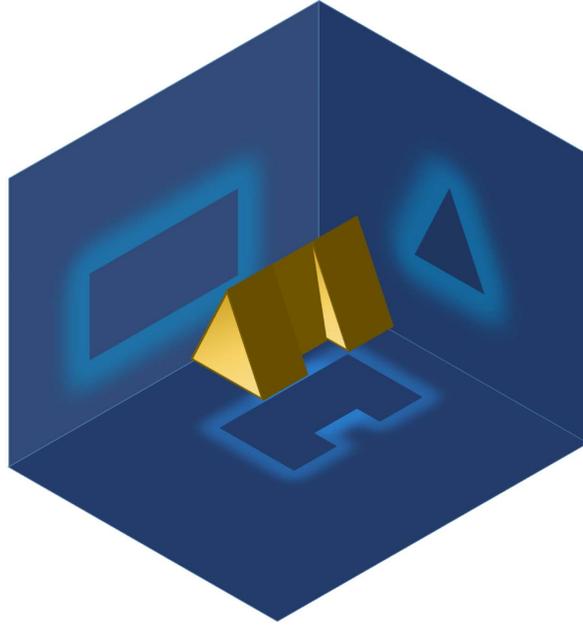In this analogy the unknown quantum state is represented by the shape of

Figure 2.1: Classical representation of tomography, each shadow provides a partial information on the shapes of three-dimensional volume (from [50]).

a three-dimensional object while directional beams of light producing shadows play the role of quantum measurements.

In spite of its fundamental relevance the first systematic approach to the problem of inferring the state of a quantum system from measurements was studied by Fano in the late 1950's [51]. Since then, however the interest devoted to quantum tomography has constantly increased. On one side, new developments in experimental techniques, especially in the fields of photodetection and nonlinear optical technology, resulted in a set of novel and beautiful experiments about quantum mechanics. On the other, promising techniques of error correction and purification would make possible fault tolerant quantum computing, long distance teleportation and cryptography motivating in such a way the increasing attention directed to quantum information technology. Quantum state estimation takes a fundamental role in this context, given its relevance to develop suitable purification protocols, and quantum characterization of communication channels.

2.1   Tomography for qubits

Single qubit tomography   Due to its simplicity it is particularly recommended to explain the basic principles of quantum state tomography using a

single qubit. Any 2×2 Hermitian density matrix of single qubit state is completely identified by three real parameters, this means that a sequence of three linearly independent measurements is enough to achieve a complete tomography. In general, the minimal set of required observables to get full knowledge of the state is referred as quorum [51]. In order to look in detail how the state is actually reconstructed, consider a general single qubit density matrix

$$\rho = \begin{pmatrix} 1+z & x-iy \\ y+iy & 1-z \end{pmatrix}, \tag{2.1}$$

with $x, y, z \in \mathbb{R}$ and $x^2 + y^2 + z^2 \leq 1$, such that the state satisfies the so-called quantum constraints (1.15) and (1.17)

$$\rho \geq 0 \quad \text{and} \quad \text{tr}(\rho) = 1. \tag{2.2}$$

The choice of real parameters x, y, z in Eq. (2.1) is not completely arbitrary but can be understood if one takes into account the Bloch sphere representation; in fact, thanks to (1.18), the same state ρ reads

$$\rho = \frac{1}{2} \left(\sigma_0 + y\sigma_1 + y\sigma_2 + z\sigma_3 \right), \tag{2.3}$$

that is X, Y and Z correspond to the correlation parameters of the state, and can be derived from ρ through (1.20)

$$x = \text{tr}(\sigma_x \rho), \quad y = \text{tr}(\sigma_y \rho), \quad z = \text{tr}(\sigma_z \rho). \tag{2.4}$$

Corresponding projective measurement are defined in (1.35). Moreover one can take advantage of orthonormality of eigenstates of Pauli operators (1.36) $\langle \psi | \psi_\perp \rangle = 0$ to halve the number of projection measurements

$$P_{|\psi\rangle} - P_{|\psi_\perp\rangle} = 2P_{|\psi\rangle} - 1. \tag{2.5}$$

One should observe howewer that the choice made here for the set of orthogonal eigenstates, which in Bloch sphere representation 1.2 corresponds to projections on the Cartesian axes $\boldsymbol{n} = \boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z$, is not unique, but makes calculations easier; on the other side any choice of three linearly independent directions allows to perform as well complete tomography on single qubit.

On the contrary, the same choice is not equivalent to others for tomography performed in laboratory where infinite ensembles cannot be considered and statistical errors come into play. In this last case making the same amount of measurements on the elements of basis $\{\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z\}$ or those equivalent by unitary transformation turns out to be the most uniform, precise and simple one among the possible measurement strategies [52].

Multi-qubit state tomography The procedure illustrated for single qubit states can be easily generalized to perform tomography on many qubit states. First, it is helpful to remark that $4^n - 1$ real parameters are required in order

to completely determine a general n qubit density matrix since it is a complex Hermitian $2^n \times 2^n$ matrix with unit trace. Analogously to the single qubit case, one observe that for multiple qubits as well, a link exists between the multiple-qubit correlation elements and measurements probabilities does exist. In fact combining (1.19) and (1.3) one has

$$
\begin{aligned}
T_{i_1,\ldots,i_n} &= \mathrm{Tr}\{(\sigma_{i_1} \otimes \cdots \otimes \sigma_{i_n})\rho\} \\
&= \frac{1}{2^n} \sum_{j_1,\ldots,j_n=0}^{3} \mathrm{Tr}\{\sigma_{i_1}\sigma_{j_1}\}\cdots\mathrm{Tr}\{\sigma_{i_1}\sigma_{j_1}\}T_{j_1,\ldots,j_n}.
\end{aligned}
\tag{2.6}
$$

Recalling that, for a single qubit

$$
T_i = P_{|\psi_i\rangle} - P_{|\psi_i^\perp\rangle}, \quad i = 1,2,3;
\tag{2.7}
$$

$$
T_0 = P_{|\psi\rangle} + P_{|\psi^\perp\rangle}, \quad \forall \psi,
\tag{2.8}
$$

one gets the following equation for n qubit systems

$$
T_{i_1,\ldots,i_n} = \left\langle \left(\hat{P}_{|\psi_{i_1}\rangle} \pm \hat{P}_{|\psi_{i_1}^\perp\rangle}\right) \otimes \cdots \otimes \left(\hat{P}_{|\psi_{i_n}\rangle} \pm \hat{P}_{|\psi_{i_n}^\perp\rangle}\right)\right\rangle,
\tag{2.9}
$$

where $\hat{P}$ indicates the measurement operator while angular bracket indicate the expectation value, such that the measurement outcomes $P \equiv \langle\hat{P}\rangle$. The plus sign is used for 0 index and the minus sign is used for a nonzero index. It follows that each local measurement, that is, any measurement which is tensor product of single qubit measurements, provides information over $2^n - 1$ correlation values at same time. Whenever one measures $T_{i_1,\ldots,i_n}$, from the same projection measurements one can calculate $\{T_{j_1,\ldots,j_n}|j_k = 0, i_k, \forall k = 1,\ldots,n\}$. In order to give a clarifying example consider tomography for a two-qubit system. Given that i_1 and i_2 are nonzero indices, the corresponding correlation element T_{i_1,i_2} reads

$$
\begin{aligned}
T_{i_1,i_2} &= \left\langle \left(\hat{P}_{|\psi_{i_1}\rangle} - \hat{P}_{|\psi_{i_1}^\perp\rangle}\right) \otimes \left(\hat{P}_{|\psi_{i_2}\rangle} - \hat{P}_{|\psi_{i_2}^\perp\rangle}\right)\right\rangle \\
&= P_{|\psi_{i_1}\rangle|\psi_{i_2}\rangle} - P_{|\psi_{i_1}\rangle|\psi_{i_2}^\perp\rangle} - P_{|\psi_{i_1}^\perp\rangle|\psi_{i_2}\rangle} + P_{|\psi_{i_1}^\perp\rangle|\psi_{i_2}^\perp\rangle},
\end{aligned}
\tag{2.10}
$$

but from the same measurement one gets also the following two correlations

$$
\begin{aligned}
T_{i_1,0} &= \left\langle \left(\hat{P}_{|\psi_{i_1}\rangle} - \hat{P}_{|\psi_{i_1}^\perp\rangle}\right) \otimes \left(\hat{P}_{|\psi_{i_2}\rangle} + \hat{P}_{|\psi_{i_2}^\perp\rangle}\right)\right\rangle \\
&= P_{|\psi_{i_1}\rangle|\psi_{i_2}\rangle} + P_{|\psi_{i_1}\rangle|\psi_{i_2}^\perp\rangle} - P_{|\psi_{i_1}^\perp\rangle|\psi_{i_2}\rangle} + P_{|\psi_{i_1}^\perp\rangle|\psi_{i_2}^\perp\rangle},
\end{aligned}
\tag{2.11}
$$

$$
\begin{aligned}
T_{0,i_2} &= \left\langle \left(\hat{P}_{|\psi_{i_1}\rangle} + \hat{P}_{|\psi_{i_1}^\perp\rangle}\right) \otimes \left(\hat{P}_{|\psi_{i_2}\rangle} - \hat{P}_{|\psi_{i_2}^\perp\rangle}\right)\right\rangle \\
&= P_{|\psi_{i_1}\rangle|\psi_{i_2}\rangle} - P_{|\psi_{i_1}\rangle|\psi_{i_2}^\perp\rangle} + P_{|\psi_{i_1}^\perp\rangle|\psi_{i_2}\rangle} + P_{|\psi_{i_1}^\perp\rangle|\psi_{i_2}^\perp\rangle}.
\end{aligned}
\tag{2.12}
$$

Finally one might notice that for n qubits, unlike the single qubit case, n detectors are not enough to completely determine a single correlations. While in single qubit case a no-click of the detector unambiguously determine a measurements the orthogonal eigenstate (2.5) in this case the ambiguity cannot be resolved but $2n$ are required to perform the tomography and determine the correlation elements [53].

2.2 Errors in quantum tomography

So far no kind of error naturally present in real world tomography was mentioned. In the previous section, for example, all measurements were supposed to be made on an infinite ensemble of states, so that one could always neglect statistical errors and perform a perfect reconstruction of the state. Tomography realized in the laboratory, however, does not fulfill this nice property. When dealing with quantum tomography actually one has to take different kinds of errors into account. First, errors in the measurement basis due to the imperfection of the experimental apparatus. This kind of errors results from the projection direction being not properly aligned with the expected one. An error induced in such a way depends on the state being measured, the totally mixed state for example has always probability $\frac{1}{2}$ to be measured whatever is the basis in which the projection is made, the tilted axes defining the projection direction has indeed to pass for the center of the Bloch sphere. Fig. 2.2 illustrates the deviation of the projection axes forming two cones. As a result the errors which arise for the estimated states become larger and larger with purity of states themselves.
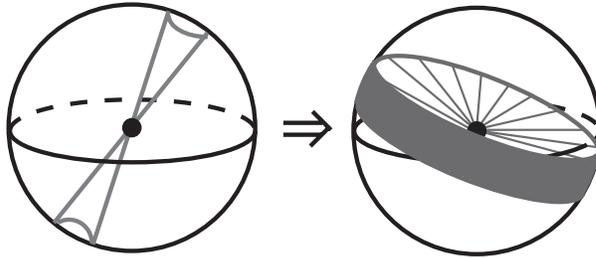

Figure 2.2: On the left the tilt of projection axes form a cone around the intended one. A measure of zero of expectation value which should correspond to a ring with projection axes as symmetry axes here turns into a band of uncertainty whose width increases with the purity of the state (from [53]).

Second, errors from experimental stability are addressed. The state production can in fact experience a drift, providing copies which do not belong anymore to the starting ensemble. If the drift is much faster than duration of the measurement the estimation procedure will provide only a totally mixed state independently of the state under examination. If the states provided by the sources change only slowly in time than also in this case the error is negligible since the timescale of drift exceeds timescale of measurements. The worst case comes into effect when the two timescales are almost equal since one faces an aliasing effect which leads to the reconstruction of a wrong state. In this third case indeed it might be very difficult to spot whether the procedure is affected by an error or not.

A possible solution for the two first causes consists in cutting data-collection time on the other side experimenter has to find a trade off with the third and last uncertainty sources that is counting statistics [53]. Actually one can never

avoid such uncertainty linked with the finite number of measurements, not only because it is impossible to realize infinite ensembles of states but it is often experimentally costly to perform a large amount of measurements. To clarify how statistical errors arise and affect the reconstruction of the state, it is helpful to consider again single qubit states. Probability $P_{|\psi_i\rangle}$ used in (1.35) experimentally corresponds to the ratio of clicks $N_{|\psi_i\rangle}$ each marking a detection of state $|\psi_i\rangle$ over the total number of measurement $N_i = N_{|\psi_i\rangle} + N_{|\psi_i^\perp\rangle}$. The expectation value for the correlation elements is then [52]

$$\langle T_i \rangle = \frac{N_{|\psi_i\rangle} - N_{|\psi_i^\perp\rangle}}{N_{|\psi_i\rangle} + N_{|\psi_i^\perp\rangle}}. \tag{2.13}$$

The angular brackets around the correlations are here used to distinguish the expectation value from the true one. The uncertainty over the correlation elements inevitably produces an uncertainty over the reconstructed state as is shown in Fig. 2.3. The Bloch's sphere representation, in fact makes clearer how the likelihood volume, that is the volume in which is most probable to find the states, evolves after each projective measurements along one specific axes. A
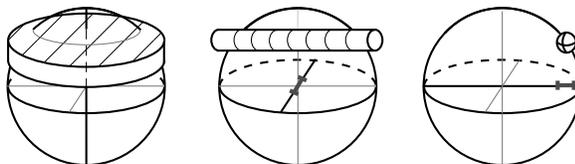

Figure 2.3: Evolution of uncertainty region of target state after projection measurements along a given direction. On the left a disk of uncertainty is obtained after projection measurements on the vertical axes, the thickness of the disk is given by the uncertainty on that direction. In the middle measurements along y axis restrict the uncertainty region to a cylinder. On the right, projection measurements on x axis finally provide a sphere of uncertainty. Here all measurements are supposed to have the same statistical error (from [40]).

statistical estimate of the error for the expectation value $\langle T_i \rangle$ is given by the width of a binomial distribution with the same expectation value

$$\Delta \langle T_i \rangle = \sqrt{\frac{N_{|\psi_i\rangle} N_{|\psi_i^\perp\rangle}}{N_i^3}}. \tag{2.14}$$

Each measurement behaves in fact like N tosses of a coin with bias $p_i = \frac{1}{2}(1 + \langle \sigma_i \rangle)$ and yields[54]

$$\langle p_i \rangle = p_i \pm \sqrt{\frac{p_i(1 - p_i)}{N}} \implies \langle T_i \rangle = T_i \pm \sqrt{\frac{1 - \langle T_i \rangle^2}{4N}} \tag{2.15}$$

This dependence of the variance on the reconstructed state itself is shown in Fig. 2.4.
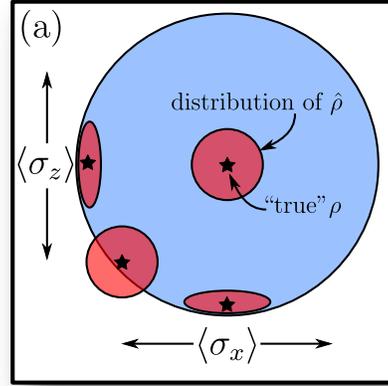
Figure 2.4: Red regions correspond to uncertainty region for measured correlations. Note that for a given number of measurements, uncertainty of a correlation decreases with the magnitude of expectation value of the correlation itself (from [54]).

2.3   State reconstruction

Fig. 2.3 shows clearly that after a finite yet complete set of tomographic measurements the state is not individuated unambiguously, but a region of uncertainty still remains. Every tomographic scheme is composed indeed of two steps, first measurements and then reconstruction of the state itself; indeed many procedures were created for this purpose, each one with some advantage and disadvantage but none can be regarded as absolutely the best. However one can identify some procedures more used than others; a prime example for these procedures is maximum likelihood reconstruction which is commonly employed thanks to its reliability in many cases of interests. In order to address the description of maximum likelihood reconstruction, it is crucial to introduce the other very significant estimation procedures. One can start considering linear inversion estimation which is the simplest procedure. It consists in setting the expectation values as calculated in (2.13) as effective correlation elements T_i of the state: $T_i = \langle T_i \rangle$. If one considers the probability $\mathcal{P}$ to have a certain set of occurrences $\{N_{|\psi_x\rangle}, N_{|\psi_x^\perp\rangle}, N_{|\psi_y\rangle}, N_{|\psi_y^\perp\rangle}, N_{|\psi_z\rangle}, N_{|\psi_z^\perp\rangle}\}$ for a given target state ρ one gets

$$\mathcal{P}(N_{|\psi_x\rangle}, N_{|\psi_x^\perp\rangle}, N_{|\psi_y\rangle}, N_{|\psi_y^\perp\rangle}, N_{|\psi_z\rangle}, N_{|\psi_z^\perp\rangle}|\rho) =$$
$$= \prod_{j=x,y,z} \binom{N_j}{N_{|\psi_j\rangle}} \left(\frac{1+T_j}{2}\right)^{N_{|\psi_j\rangle}} \left(\frac{1-T_j}{2}\right)^{N_{|\psi_j^\perp\rangle}}, \quad (2.16)$$

this means that, by liner inversion the reconstructed state turns out to be

the state which maximizes the probability $\mathcal{P}$ to get the measured number of clicks. Unfortunately, this state might be non-physical i.e. it might not fulfill the physicality constraints (2.2). In Fig. 2.3, for example part of the uncertainty region is outside the Bloch sphere which encloses all possible physical states. In order to clarify how this can happen one can look at a specific example for a single qubit state. An experimenter measures σ_x, σ_y, and σ_z once, and observe the result $+1$ for all of them. The density matrix one can reconstruct at this point is

$$\rho_{\mathrm{e}} = \frac{1}{2} \begin{pmatrix} 2 & 1-i \\ 1+i & 0 \end{pmatrix}, \tag{2.17}$$

where the correlation elements $T_i = \langle T_i \rangle = +1$ are plugged into (1.18). Thus, even if such a matrix is obtained from possible outcomes of measurements on physical states it has a negative eigenvalue $\lambda_- = \frac{1-\sqrt{3}}{2} \approx -0.336$. Although, with just one measurement in each direction, this simple example shows that setting straightforwardly Born's frequencies as correlation elements does not guarantee the positivity and so the physicality of the state. More generally if one performs N measurements in each direction, possible estimates of correlation elements form $(N+1) \times (N+1) \times (N+1)$ grid; this grid defines a cube usually called Bloch cube [55] which circumscribes the Bloch sphere but contains also non-physical states. Such a structure is shown in Fig. 2.5.
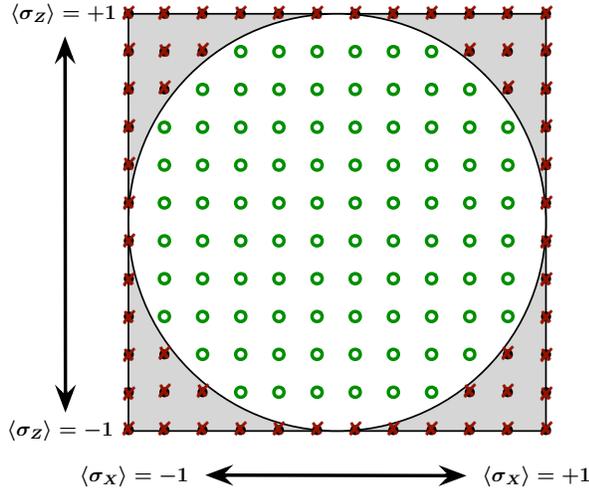

Figure 2.5: Grid of expectation values for the Stokes parameters forms a Bloch cube which circumscribes the Bloch sphere of physical state. The layer here reported is the one with $\langle \sigma_z \rangle = 0$, the combination of Stokes parameters which corresponds to physical states are marked with green circles, while those which produce unphysical states are crossed out in red (from [55]).

Of course, the probability to fall in the unphysical region decreases with the

number of measurements but increases with the purity of the target state, for example getting never lower then 50% even after an arbitrary large number of measurements for state of purity one. This problem gets even worse for higher dimensional Hilbert spaces. In order to keep the root-mean-square deviation $\Delta_2 = \sqrt{\mathrm{Tr}[(\rho_{\mathrm{tomo}} - \rho)^2]}$ constant, N must grow proportional to the dimension d since the number of independent parameters grows quadratically like $d^2 - 1$. Furthermore the number of eigenvalues equals the number of dimensions d, and with it the probability that at least one of them is negative. That being so, one can easily conclude that tomographic estimates of large systems through linear inversion are rarely non-negative. Such relevant problem forbids in fact the use of a simple and straightforward procedure like linear inversion and forces to look for other procedures able to provide only physical states [52].

State reconstruction fulfilling physicality constraints The most direct procedure one can think in order to reconstruct a physical state starting from any measurement outcomes for correlations consists in keeping the reconstructed state by linear inversion with Bloch vector $\boldsymbol{r}_{\mathrm{LI}}$ if this state is physical $\|\boldsymbol{r}_{\mathrm{LI}}\| \leq 1$; otherwise it provides the physical state whose Bloch vector $\boldsymbol{r}_e$ minimizes the distance with $\boldsymbol{r}_{\mathrm{LI}}$

$$\boldsymbol{r_e} = \begin{cases} \boldsymbol{r}_{\mathrm{LI}}, & \text{if } \|\boldsymbol{r}_{\mathrm{LI}}\| \leq 1, \\ \boldsymbol{r}_{\mathrm{LI}}/\|\boldsymbol{r}_{\mathrm{LI}}\| & \text{if } \|\boldsymbol{r}_{\mathrm{LI}}\| > 1. \end{cases} \tag{2.18}$$

Of course the distance to use is not fixed; for example one is free to choose any Shannon $p-$distance $\|\rho - \rho'\|_p = (|x - x'|^p + |y - y'|^p + |z - z'|^p)^{\frac{1}{p}}$. Trace distance $p = 1$ is among distances employed more often. However spherical symmetry of physical state space around the fully mixed state $\boldsymbol{r}_m = 0$ suggests the use of the Hilbert-Schmidt distance. i.e. $p = 2$.

An alternative procedure looks for the physical state ρ_e which maximizes the fidelity $F(\rho_e, \rho_{\mathrm{LI}})$ with the state reconstructed by linear inversion ρ_{LI}. More generally, trough spectral decomposition of ρ_{LI}

$$\rho_{LI} = \sum_{\lambda} f_\lambda \, |\phi_\lambda\rangle\langle\phi_\lambda| \,, \tag{2.19}$$

one can look for the state ρ_T which minimizes a given target function F of frequencies f_λ, the eigenvalues of ρ_{LI}

$$\rho_T = \operatorname*{argmin}_{\rho > 0}(F(\rho|f)). \tag{2.20}$$

The most common choices for the target function are the maximum likelihood [28]

$$F^{MLE} = -\sum_{\lambda} f_\lambda \log\left[\mathrm{Tr}(\rho \, |\phi_\lambda\rangle\langle\phi_\lambda|)\right] \tag{2.21}$$

and the least squares function [16]

$$F^{LS} = \sum_{\lambda} \frac{1}{w_\lambda} \left[f_\lambda - \mathrm{Tr}(\rho \, |\phi_\lambda\rangle\langle\phi_\lambda|)\right]^2. \tag{2.22}$$

Weights w_λ might be both fixed to $w = 1/f_\lambda$ or left free to change with the states like $w = 1/\operatorname{Tr}(\rho \,|\phi_\lambda\rangle\langle\phi_\lambda|)$. However minimizing these functions is not an easy task, and iterative algorithms are necessary to find the state which fulfills the required condition.

2.3.1   Iterative algorithms

Iterative fixed point algorithm   With iterative fixed point algorithm (2.21) one starts choosing completely mixed state $\rho^{(0)} = \frac{1}{2^n}\mathbb{1}_2^{\otimes n}$ as initial guess, this choice guarantees in fact to avoid zero eigenvalues. The density matrix $\rho^{(k+1)}$ at the $(k+1)$th iterative step is then calculated from $\rho^{(k)}$ by

$$\rho^{(k+1)} = \mathcal{N}\left[R(\rho^{(k)})\rho^{(k)}R(\rho^{(k)})\right], \tag{2.23}$$

where $\mathcal{N}$ is a normalization constant such that $\operatorname{Tr}\big(\rho^{(k+1)}\big) = 1$. The function $R(\rho)$ is instead defined as[29]

$$R(\rho) = \sum_\lambda \frac{f_\lambda}{\operatorname{Tr}(\rho \,|\phi_\lambda\rangle\langle\phi_\lambda|)}\,|\phi_\lambda\rangle\langle\phi_\lambda|. \tag{2.24}$$

Indeed one can easily proof that $F^{MLE}(\rho^{(k+1)}) < F^{MLE}(\rho^{(k)})$ [29] and approach in such a way a maximum likelihood reconstructed state with arbitrary precision.

Cholesky decomposition   Cholesky decomposition [56] relies on the fact that every non-negative Hermitian matrix can be decomposed into a product of a triangular matrix Λ and its conjugate transposed $\Lambda^\dagger$. So that

$$\Lambda = \begin{pmatrix} t_1 & 0 & 0 & \dots & 0 \\ t_{n+1} + it_{n+2} & t_2 & 0 & \dots & 0 \\ t_{n+3} + it_{n+4} & t_{n+5} + it_{n+6} & t_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{4^n-2^n-1} + it_{4^n-2^n} & t_{4^n-2^n+1} + it_{4^n-2^n+2} & 0 & \dots & t_n \end{pmatrix}. \tag{2.25}$$

Then, for any choice of $\boldsymbol{t} \in \mathbb{R}^{4^N}$ one can parametetrize any valid physical density matrix ρ with trace condition (1.17) thanks to

$$\rho(\boldsymbol{t}) = \frac{\Lambda^\dagger \Lambda}{\operatorname{Tr}(\Lambda^\dagger \Lambda)}. \tag{2.26}$$

This parameterization is then employed to transform the constrained optimization problem of (2.20) into an unconstrained one

$$\rho_e = \arg\min_{\boldsymbol{t}\in\mathbb{R}^{n^4}} F(\rho(\boldsymbol{t})|f). \tag{2.27}$$

Convex optimization   An alternative approach to find a physical minimum of the target function is to set an additional cost function term which inhibits unphysical solution; as example one can start with

$$\tilde{F}_T(\rho) = F_T(\rho) + \epsilon \log[\det(\rho)]. \tag{2.28}$$

The determinant of the density matrix coincides with the product of its eigenvalues. This means that starting from a physical state with non null eigenvalues one encounters states with at least one eigenvalue which becomes lower, and small eigenvalues approaching zero before passing to unphysical region with negative eigenvalues. The barrier $\log[\det(\rho)]$ which separates the physical from unphysical region however, avoids the crossing during the optimization and forces the optimal state to be physical. Once found such a state, it becomes possible to lower a little the barrier finding another state, and so on iteratively approaching the state which minimizes $F_T(\rho)$.

Direct and conjugate gradient   With the direct gradient algorithm [57] one arbitrarily fixes a starting physical state ρ_0, and moves a step towards the direction defined by the steepest gradient of the cost function. The length of the step is chosen in such a way to reach the minimum of the cost function along the direction of the step itself. This kind of approach, yet quite intuitive might need many steps which end up to create a zigzag path before approaching the effective minimum of the function; conjugate gradient [58]instead overcomes this problem. Starting again with a arbitrary physical state ρ_0 one moves a first step just like steepest gradient method. Now if $\boldsymbol{u}$ is the unit vector identifying the direction of the first step and $\boldsymbol{A}$ is the Hessian matrix, conjugate gradient method recommends to move the second step along direction $\boldsymbol{v}$ defined by

$$\boldsymbol{u} \cdot \boldsymbol{A} \cdot \boldsymbol{v} = 0, \tag{2.29}$$

that is vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ are conjugate. In such a way interfering directions of direct gradient are avoided and a minimum of cost function is reached in two steps. A comparison between the two mentioned algorithms, with astonishing improvement of conjugate gradient is also illustrated in Fig. 2.6
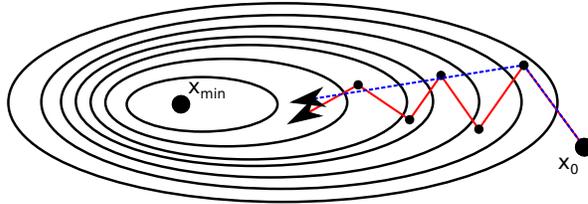

Figure 2.6: Comparison between conjugate gradient (dashed blue arrows) and steepest descent methods (red arrows). The contour represents the function to optimise over. (from [59])

Projected gradient descent   Projected gradient descent is an iterative pro-
cedure with two substeps. Starting with a well-chosen physical state, a step
is taken in the downhill direction of the optimization function, which has the
chance to result in a nonphysical matrix. Second, to bring the estimate back
to the constrained, physical space, one projects it to the closest point in the
solution space. This two step process is then repeated until the cost function
converges towards a sufficiently small value. Since the search is over a convex
set, as long as the cost function is a strictly convex function of ρ, there will be
a unique solution that minimizes it. Fig. 2.7 shows the evolution of the density
matrix estimate of a qubit through six iterations of projected gradient descent.
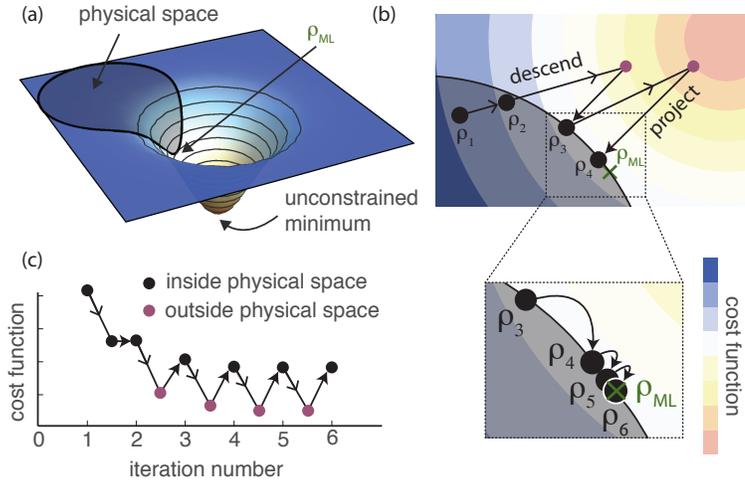

Figure 2.7: a) Graphical representation of convex cost function, with region of
physical states which does not include the unconstrained minimum of maximum
likelihood function. b) Reconstructed state for the first six steps of project
gradient procedure, red dots correspond to the unphysical states. Path of last
three physical state is highlighted in zoomed box. c)Trend of cost function for
physical and and unphysical states at each step (from [60]).

Conjugate gradient-Accelerated project gradient   The main advantage
of project gradient method relies on the fact that despite other maximum like-
lihood reconstruction algorithms it does not contain by construction any phys-
icality constraints which slow down the procedure especially when the estimate
approaches the true value. In fact, taking into account the Cholesky parma-
terization of the physical states $\rho = A^{\dagger}A/\operatorname{Tr}(A^{\dagger}A)$, and performing gradient
descent method of gradient descent in the factored space of unconstrained A

operators, for $F(A) \equiv F(A^\dagger A/\operatorname{Tr}(A^\dagger A))$, straightforward algebra yields [33]

$$\delta \widetilde{F}(A) = -\operatorname{Tr}\left(\delta A \frac{(R-1)A^\dagger}{\operatorname{Tr}(A^\dagger A)} + \text{h.c.}\right), \tag{2.30}$$

to linear order in δA and with $\nabla F(\rho) = -R$, the gradient of F. Choosing a small ϵ such that $\delta A = \epsilon A(R-1)$ one finds that the updated form of ρ_i is

$$\rho_{i+1} = \rho_i + \delta\rho_i = \rho_i + \epsilon[(R-1)\rho_i + \rho_i(R-1)], \tag{2.31}$$

to linear order in ϵ. When one has limited data and nearly a pure target state, maximum likelihood estimate is close to the boundary of the state space and ρ gets close to a rank-deficient state with at least one small eigenvalue because of the so-called low rank problem discussed further on. Yet ρ_i has unit trace, so its spectrum is highly asymmetric. $\delta\rho_i$ inherits this asymmetry, leading to a locally ill-conditioned problem and slow convergence. On the other side, far away from the boundary of the state space the Hessian of cost function is almost constant making the use of conjugate gradient advantageous. The conjugate gradient-accelerated product gradient combines the advantages of the schemes. In fact it employs conjugate gradient for first steps and switches to project gradient when the Hessian of cost function starts to change too much between iterations. Moreover, each step of project gradient method is made longer at each iteration. When the scalar product between the gradient of the cost function and the direction in which last step was made becomes negative, algorithm proceeds to an adaptive restart of momentum, which defines the length and the direction of the steps. The improvement of the conjugate gradient-adaptive project gradient method compared to other algorithms is shown evaluating the time required for the esimation of 8qubit W state in Fig. 2.8.

Estimators described so far are indeed the most commonly employed ones in quantum tomography. However one should remark that does not exist any procedures which is the the best for every target state. Consider for example a four-party Greenberger-Horne-Zeilinger mixed state with white noise

$$\rho_0 = p\,|\text{GHZ}_4\rangle\langle\text{GHZ}_4| + (1-p)\frac{I}{16} \tag{2.32}$$

where $|\text{GHZ}_4\rangle = (|0000\rangle + |1111\rangle)/\sqrt{2}$. It can be proven [62] that when determing, e.g., the fidelity with respect the GHZ-state, the linear inversion procedure is more reliable than the least squares and the maximum likelihood, which conversely are shown to be biased. Mean fidelity of reconstructed states by linear inversion agree much better with the real value on average than those obtained with maximum likelihood and least squares optimization as shown in Fig.2.9.

In order to illustrate the sources of this bias one can see in Fig.2.10 how the physicality constraint forces the mean of the reconstructed states into and onto the Bloch sphere and hence systematically away from the target state; actually one can show [45, 62] that every reconstruction scheme for quantum
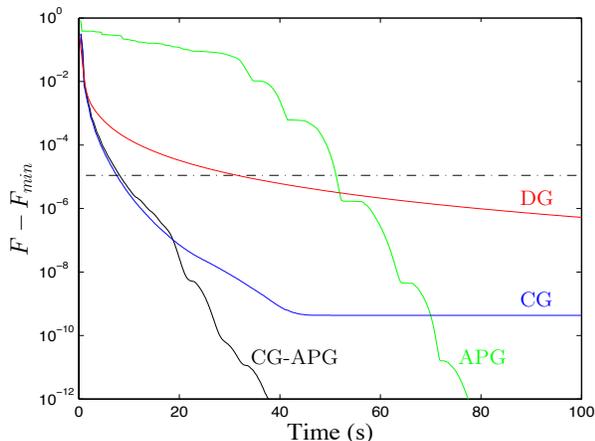
Figure 2.8: Performances of iterative algorithms. The deviation $F - F_{min} = \frac{1}{N} \log(\mathcal{L}\mathcal{L}_{max})$. Here $N = 3^8 \times 100$, since there are 100 copies for each of the 3^8 settings of the 8-qubit product-Pauli measurements [18] (from [61]).

state tomography that always yields a valid, non-negative, density operators is biased.

This kind of biasing problem relies more specifically on the fact that a procedure which takes the physicality constraint into account provides states with rank smaller than states reconstructed by linear inversion, this problem is usually addressed as low rank problem.

2.3.2   Low rank problem

Estimators manage to provide a physical state, specifically states which fulfill the positivity constraint (1.15) even when linear inversion offers a non-physical solution. To accomplish this task, all the eigenvalues of the unphysical density matrix are forced to be non negative, but in doing so, originally negative eigenvalues eventually go to zero. Because of this, estimators turn out to provide rank-deficient yet physical density matrices. Fig. 2.11, shows what happens to the maximum likelihood reconstructed state ρ_{MLE} when the density matrix obtained by linear inversion ρ_{tomo} is unphysical. Starting from the physical region where all eigenvalues are non-negative, the estimated state is moved towards the unphysical region where at least one eigenvalue is negative. Supposing that eigenvalues change smoothly near the hyperplane which separates the regions, at least one of them has to approach zero before the corresponding state reaches the unphysical region. As a consequence, all the states which lie on the edge have at least one null eigenvalue and are not full rank.

To understand why predicted zero eigenvalues are actually a problem in quantum state reconstruction one must have a look to probability theory. In fact a quantum state is nothing more than a prediction of the future. Every estimate
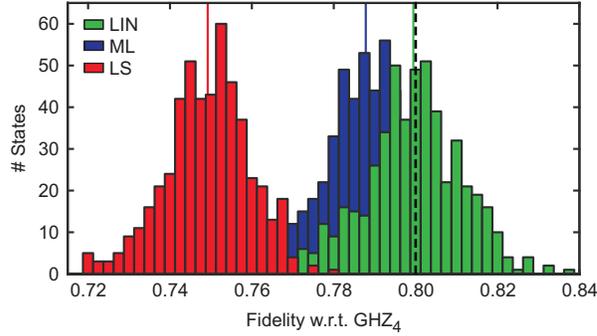
Figure 2.9: Histograms of fidelities between reconstructed states and GHZ_4 state, with least squares optimization in red, maximum likelihood in blue and linear inversion in green (from [62]).
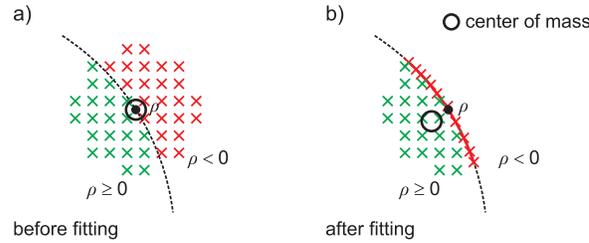

Figure 2.10: The state obtained by averaging expectation values of Stokes parameters fits the target one (a), if unphysical states are projected onto the physical space, the state provided by the averaging of the correlations does not fit with the true state (from [45]).

of a state is done on a finite amount of data and quantitative evaluation of an estimate is a matter for debate, especially when one considers low probabilities. One can start considering for example an estimator which never observed the event k over a finite number of observations. One should the hardly claim that the probability for his outcome $p_k = 0$ since p_k might be simply very small but not necessarily zero. For one trial it is quite trivial to see how such estimate is unjustified if $p_k \neq 1$. But actually, no matter how many data points the estimator has, one can always imagine a much larger data set in the future, which might debunk the prediction "$p_k = 0$". Even considering statistical uncertainty does not solve the problem. For instance, while $\hat{p} = 0.5$ is a decent estimate of $p = 0.51$, on the other side $\hat{p} = 0$ is not a good estimate for $p = 0.01$. The estimate $\hat{p} = 0.5$ could mean $\hat{p} = 0.5 \pm 0.01$. To report $\hat{p} = 0 \pm 0.01$, however, is nonsensical. This would mean "p is probably between -0.01 and 0.01" but p must be non-negative, an unconditionally better description is "p is probably between 0 and 0.01". This is not the only way of representing the probability p
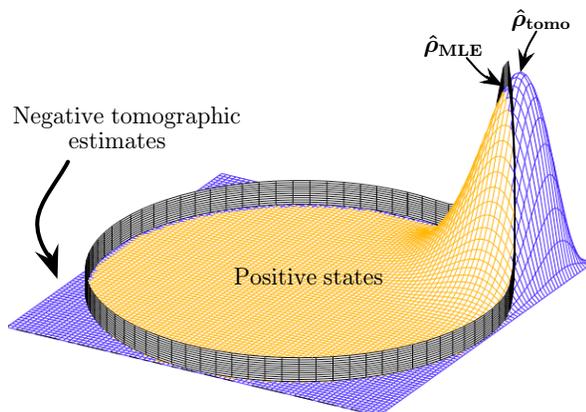
Figure 2.11: Maximum likelihood convex function. The state which maximize this function ρ_{tomo} is not physical. In fact it lies outside the physical region. If one constraints the maximization over the region of physical states then finds the maximum likelihood reconstructed state ρ_{MLE} (from [55]).

between 0 and 0.01 since the confidence of an estimator might be skewed towards one side of the interval. Anyway, 0 is only optimal when the confidence interval has zero width which implies absolute certainty about the outcomes of future measurements. If one accepts that zero probabilities are implausible, then each zero eigenvalue should be replaced by a small, but finite ϵ. However this choice brings two substantial problems. Firstly, what is ϵ, that is how is it related to the number of measurements, secondly, how does this fixing of small eigenvalues affects the larger eigenvalues of ρ. In fact increasing many small eigenvalues will require decreasing the largest ones. The only way to resolve this messy situation is to avoid zero eigenvalues in the first place. To do so one can think for instance of procedures which force full rank reconstruction of the states. An example for such procedures is the hedged estimator which is analogous to the maximum likelihood estimator but with a significant modification. A hedged estimator replaces the standard likelihood function $\mathcal{L}(\rho) = \mathrm{Pr}(\text{observed data}|\rho)$ with the product of $\mathcal{L}$ and an "hedging function"[63]

$$h(\rho) = \det(\rho)^{\beta}, \tag{2.33}$$

where $\beta = \frac{1}{2}$ is a positive constant chosen by the estimator. In Fig. 2.12 one can see how the grid of the reconstructed states for $\langle \sigma_z \rangle = 0$ is distorted for hedged maximum likelihood compared with the standard maximum likelihood and the linear inversion.

In order to have a quantitative estimation of how this second kind of procedures overcomes procedures affected by the low rank problem one can look at two simple estimators, one for each class. The first one is the so-called constrained least-squares $p-$distance estimator as defined in (2.18), the other estimator is the hedged estimator[65]
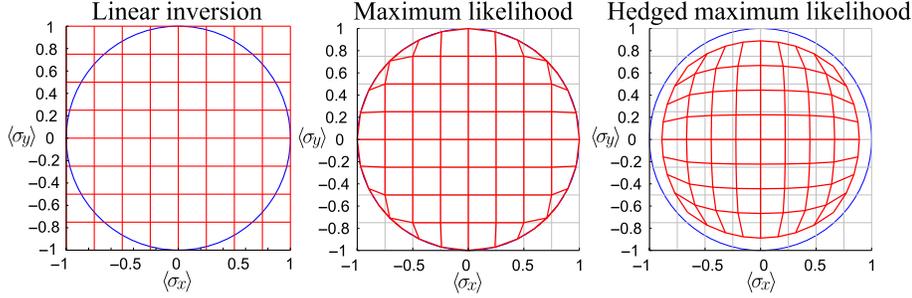
Figure 2.12: On the left it is shown the grid for pairs of expectation value $\langle\sigma_x\rangle$ and $\langle\sigma_y\rangle$ when $\langle\sigma_z\rangle = 0$, the blue circle contains those knots which correspond to physical states. In the middle, maximum likelihood reconstruction moves dots got by liner inversion into physical region, in particular all the states which were nonphysical are projected on the surface of the sphere, that is are made pure. On the right hedged maximum likelihood provides a physical state, which lies inside the physical region with no one on the edge for any value of expectation value for the correlations (from [64]).

$$r_e = \begin{cases} r, & \text{if } \|r\| \leq 1 \\ \sqrt{1-h}\,r/\|r\| & \text{if } \|r\| > 1 \end{cases} \tag{2.34}$$

this estimator indeed forces the state to be full rank (see Fig. 2.13). Now in order to compare these two estimators one has to fix the value of h. Here it is chosen so that expected Hilbert-Schmidt risk between the target state ρ and the estimated state ρ_e for a given set of data D

$$R(\rho, \rho_e) = \sum_D \Pr(\rho|D)\,\mathrm{tr}(\rho - \rho_e)^2; \tag{2.35}$$

is maximum in $r = (0, 1)$. In such a way one finds

$$h = \frac{1}{N} - \frac{1}{N^2}, \tag{2.36}$$

where N is the number of measurements in each direction [65]. Fig. 2.14 shows how the hedged estimator improves always in terms of Hilbert-Schmidt risk estimator reporting rank-deficient physical states.

2.3.3 Bayesian mean estimation

Bayesian estimation is another procedure often employed in order to overcome the low rank problem, moreover it provides natural errors to the reconstructed state. First one needs to define a probability distribution for physical states given a set of tomographic measurements. For example one can start from the maximum likelihood function and multiply it by a prior distribution over the
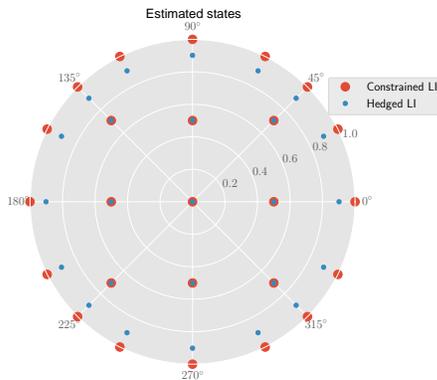
Figure 2.13: Distribution of 50 reconstructed states. Orange dots corresponds to states obtained from constrained estimator (2.18), while blue ones are those marking states obtained through hedged estimator (2.34), white circles indicates surfaces with same purity (from [65]).

state in order to get a measure $\mathcal{D}\rho = \pi(\rho)\mathrm{d}\rho$. Reconstructed state is then defined as the mean of the distribution over all the volume of physical states

$$\bar{\rho}_{\mathrm{BME}} = \frac{\int \rho \mathcal{L}(\rho) \mathcal{D}\rho}{\int \mathcal{L}(\rho) \mathcal{D}\rho}. \tag{2.37}$$

Defining the Bloch vector in analogous way, one can get the variance of its component as $(\Delta x_{\mathrm{BME}})^2 = \overline{x^2}_{\mathrm{BME}} - \bar{x}^2_{\mathrm{BME}}$ where

$$\overline{f}(x) = \frac{\int f(x)\mathcal{L}(\boldsymbol{r})\mathcal{D}\boldsymbol{r}}{\int \mathcal{L}(\boldsymbol{r})\mathcal{D}\boldsymbol{r}}. \tag{2.38}$$

Thus, on one side Bayesian mean estimation provides well-motivated error bars compatible with the reconstructed state ρ_{BME}, and is not affect by the low rank problem, in fact ρ_{BME}'s eigenvalues are never unjustifiably zero [55], on the other side there is a big disadvantage; it is unclear which is the best way to implement the integration over the physical states. Both the reconstructed state and its uncertainty in fact depend on the prior function set to define a measure over the space of physical states $\mathcal{D}\rho$ [32]. Such a handicap makes it indeed very challenging to implement this scheme for actual tomography.

2.3.4   Bootstrapping

Although errors with Bayesian mean estimation are obtained straightforwardly from theory, it is possible to evaluate uncertainty associated with the reconstructed state for any choice of the estimator and for any result provided by such estimator. The method which allows this error evaluation is called bootstrapping or case resampling procedure [66]. Uncertainty estimated with
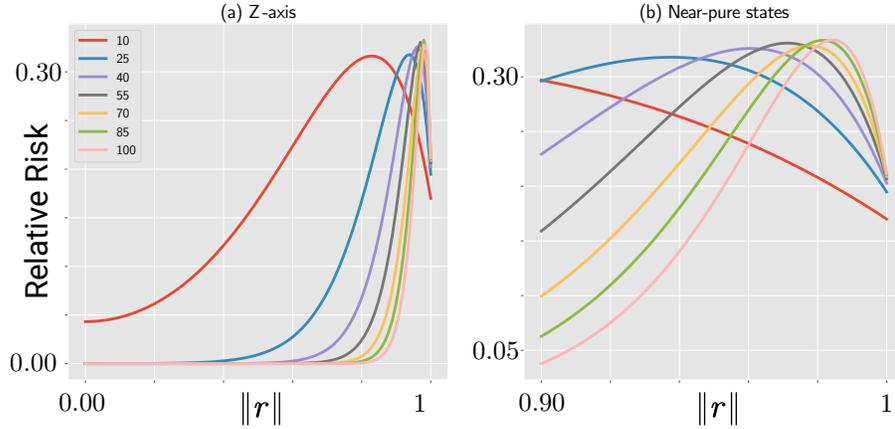
Figure 2.14: Relative Hilbert-Schmidt risk along the Z-axis, as a function of the radial coordinate $\|\boldsymbol{r}\|$. In (a) relative risk is calculated for after a number of measurements equal to $N = 10, 25, 40, 55, 70, 85, 100$. (b) shows with more detail the behaviour of the risk for higher valueso of $\|\boldsymbol{r}\|$ (from [65]).

bootstrapping contains however systematic biases for the different tomographic methods [52]. Once one has collected all experimental data and determined a density matrix, one can use this last one to simulate a new data set employing the same measurement operators and probability distributions (2.3). Then for each of these data sets one can perform state reconstruction so that one gets a distribution of density matrices. The argument which supports this procedure is that each data set generated might have been actually measured, and all the reconstructed density matrices, each with its probability might be the actual density matrix under examination.

2.4 Alternative tomographic schemes

The basic tomographic scheme discussed so far, represents often only a starting point for other procedures which aim to perform tomography in a more efficient way that is reaching a good precision for estimation with as little effort as possible. This section selects and describes tomographic protocols among the most promising ones in terms of robustness or simplicity of their experimental implementation.

2.4.1 Self-guided tomography

Self-guided quantum tomography [67] relies on a self-learning technique able to guide itself to an arbitrarily precise estimate of target state. In order to fulfill this task the algorithm of self-guided tomography has to evaluate at each step how far the reconstructed state is away from target one, and has to try

to indicate the direction the experimenter needs to perform next measurement. Namely, let ρ be the target state, σ the estimated state at the $i-$th step and m the measure chosen for evaluating the distance between these physical states. Then an expectation value is assigned to a function f of σ once fixed target state ρ

$$f(\sigma) = \langle m(\rho, \sigma) \rangle . \tag{2.39}$$

Of course one has to require that the distance measure m can be estimated from the experiment. It turns out that the efficiency of the algorithm does not depend on which measure the algorithm is considering, therefore the idea is to chose that one which requires less time to be measured, in order to speed up each step and the whole procedure accordingly; for example, if one considers pure states, one can use for metric the infidelity between the states [68]

$$m(\rho, \sigma) = 1 - F(\rho, \sigma). \tag{2.40}$$

Once the point of how to experimentally evaluate the distance among states was resolved, the task is to employ this information to provide an estimate state of σ which is nearer and nearer to ρ at each iterative step. To this end the algorithm makes use of the so-called simultaneous perturbation stochastic approximation [69]. Given a state σ it defines two proposal states σ_+ and σ_- which depends on σ trough

$$\sigma_\pm = \sigma \pm \beta \Delta, \tag{2.41}$$

where $\beta \in \mathbb{R}$ and $\Delta \in \mathcal{H}$, the Hilbert space of physical states. In such a way it becomes possible to define the gradient g of f as difference quotient between σ_- and σ_+, in particular at k-th step the gradient reads

$$g_k = \frac{f(\sigma_k + \beta_k \Delta_k) - f(\sigma_k - \beta_k \Delta_k)}{2\beta_k} \Delta_k. \tag{2.42}$$

Gradient g_k indicates direction where move the old estimate state to get a new one ensuring a lower value for f, in fact one has

$$\sigma_{k+1} = \sigma_k + \alpha_k g_k. \tag{2.43}$$

In order to explore all directions, Δ_k is randomly generated at each step while numerical parameters like α_k and β_k are tuned through extensive numerical simulations on many problems. Gradient g_k indicates the direction where to move the old estimated state to get a lower value for f. Fig. 2.15, shows the path iteration by iteration of reconstructed single qubit states on the Bloch sphere surface. On the left 10^2 measurements are performed at each step, on the right side 10^4. Surprisingly, one can see that performance of the tomography does not increase significantly with the number of measurements; this represents indeed the peculiar advantage of self-guided tomography.

A most interesting result of self guided tomography is visible also from Fig. 2.15; when estimated state converges on the target true state, the performance is roughly independent of the number of experiments shot for iteration. Moreover
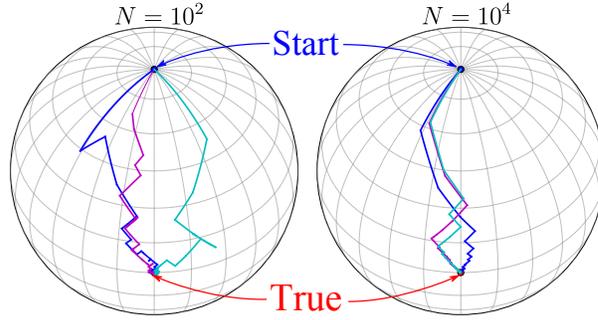
Figure 2.15: Paths with $k = 10^3$ steps for the reconstructed states in self-guided tomography. Starting point is a state here represented as the north pole of the Bloch sphere. Three sets of random direction, define each a different path leading near to the true state. On the left the gradient at each step is calculated after $N = 10^2$ measurements, on the right after $N = 10^4$ measurements (from [67]).

for a fixed total number of experiment shot, the procedure with less measurements and more steps guarantees an higher fidelity of the state. In Fig. 2.16 one can look at the trend of infidelities for selfguided tomography performed with 10, 10^2 or 10^4 measurements for iteration, with an inset which stresses the differences after the same total number of experiments shot.

2.4.2 Permutationally invariant tomography

The protocol described here is suited especially for reconstruction of permutationally invariant states [70]. In fact every quantum state ρ can be decomposed in two parts, the permutationally invariant ρ_{PI} and non-permutationally invariant part $\bar{\rho}_{PI}$ such that

$$\rho = \rho_{PI} + \bar{\rho}_{PI}. \tag{2.44}$$

The definition of permutaionally invariant states ρ_{PI} is made in terms of permutation operator for qubits Π_i like

$$\rho_{PI} = \frac{1}{N!} \sum_i \Pi_i^\dagger \rho \Pi_i. \tag{2.45}$$

Permutation operators Π_i acts on n qubit states by switching the $i-$th qubit with the next one when $i \neq n$ and the last qubit with the first when otherwise, thus labeling each qubit with an index like $|\psi\rangle = |1, \ldots, i, \ldots, n\rangle$, Π_i reads

$$\Pi_i = |1, \ldots, i+1, i, \ldots, n\rangle\langle 1, \ldots, i, i+1, \ldots, n| . \tag{2.46}$$

Before proceeding with the discussion of the protocol, one might observe that convenience of the scheme relies on the fact that many states of physical interest
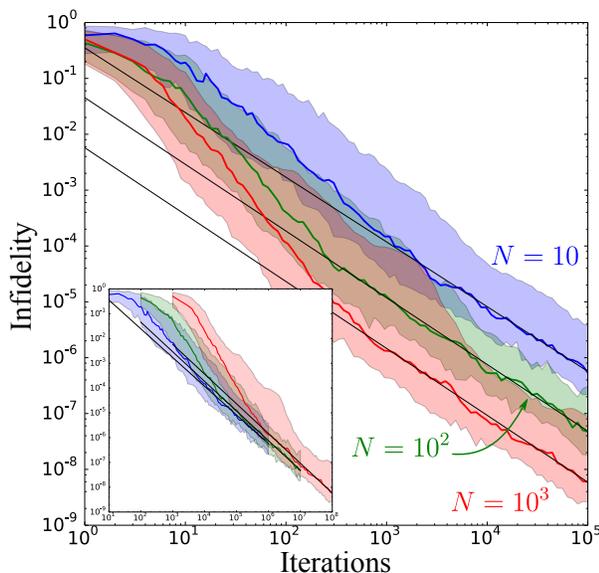
Figure 2.16: Trend of infidelity $1 - F(\rho_T, \rho_e)$ after a given number of iterations for $N = 10, 10^2, 10^3$ measuremtetnts. The thick lines correspond to the median while coloured regions correspond to interquartile ranges over 100 pure target states randomly distributed according to Haar measure. In the inbox same infidelities are compared with the total number of measurements. One can see that, for a given number of measurements strategy with less measurements for each iteration overcome the others, especially when few measurements are performed.(from [67])

are actually permutationally invariant states. Most common example are the Greenberger–Horne–Zeilinger state [71], which for n qubits reads

$$|\mathrm{GHZ}_n\rangle = \frac{|0\ldots 0\rangle + |1\ldots 1\rangle}{\sqrt{2}}, \qquad (2.47)$$

the W_n state [72]

$$|W_n\rangle = \frac{1}{\sqrt{n}} \left(|10\ldots 0\rangle + |01\ldots 0\rangle + \cdots + |00\ldots 1\rangle\right), \qquad (2.48)$$

and its generalization, the symmetric Dicke state [73]

$$\left|D_n^k\right\rangle = \frac{1}{\sqrt{C_n^k}} \sum_l \mathcal{P}_l \left(|1\rangle^{\otimes k} \otimes |0\rangle^{\otimes (n-k)}\right). \qquad (2.49)$$

where $\sum_l \mathcal{P}_l$ denotes the sum over all possible permutations and C_n^k is the binomial coefficient. Moreover, for many other states the non-permutationally

invariant part is negligible, thus one can conclude that just the determination of ρ_{PI} in many experimental realizations is already a fairly good approximation to ρ. In order to take advantage of the symmetry of permutationally invariant states, one can start observing that from (1.20) and (1.20), their correlation elements read

$$T_{i_1,\dots i_n}(\rho_{PI}) = \langle \sigma_{i_1} \otimes \cdots \otimes \sigma_{i_n} \rangle_{\rho_{PI}} = \frac{1}{n!} \sum_i \mathrm{Tr}\left(\rho \Pi_i \sigma_{i_1} \otimes \cdots \otimes \sigma_{i_n} \Pi_i^\dagger\right),$$
(2.50)

where the cyclic property of trace was employed. From (2.50), it turns out that the correlation elements $T_{i_1,\dots i_n}$ of permutationally invariant states are independent under swapping of local measurements. In other words, one measurement provides information on all the elements which are equivalent by indices permutation and as a result a complete tomography is achieved with less effort. Specifically, in order determine the number of independent correlation elements one has to count how many combinations exist for the 3 single qubit Pauli operators plus the identity in n slots. This is equivalent to looking at all possible combinations of three bounds in $n + 3$ slots as shown in Fig. 2.17, in fact each bound can be imagined as separating series of slots distinguished by Pauli operators filling each slot. The number of these combinations is given by the
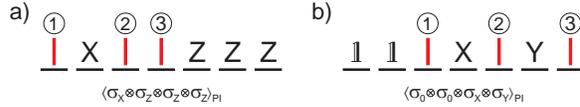

Figure 2.17: The number of parameters that need to be determined in order to identify unambigously a permutationally invariant state, is given by the number of possible distribution of the three red bars into $n + 3$ sites (from[45]).

binomial coefficient

$$\binom{n+3}{n} = \frac{(n+3)(n+2)(n+1)n!}{3!n!} = \frac{1}{6}(n^3 + n^2 + 11n + 6),$$
(2.51)

thus the necessary measurement effort to reconstruct a permutationally invariant state scales only polynomially with the number of qubits. Actually, since eigenstates of Pauli operators are eignestates of identity operators as well one can further reduce the number of measurements necessary to determine all permutationally invariant correlation elements. In particular an efficient approach consists in a set a global operators $A_j^{\otimes N}(j = 1, \dots D_N)$ where for each measurement setting, the same local operator is measured on all qubits, with $A_j = \boldsymbol{n}_j \cdot \boldsymbol{\sigma}$ while D_n is obtained similarly to the combinations of Eq.(2.51) but neglecting now all those arrangements containing the identity operator which results in

$$D_n = \binom{n+2}{n} = \frac{1}{2}(n^2 + n + 2).$$
(2.52)

With such a choice for operators, one has an increasing of required measurements which is only quadratic in the number of qubits, a factor n lower than what one had without this choice. In order to further improve the procedure one can optimize the set of measurement operators. In fact, even though this set can be chosen quite arbitrarily if one wants to achieve small errors, it is best to choose a uniform distribution for versors $\boldsymbol{n}_j$ which characterizes the operators. In Fig. 2.18 two examples for uniform spread of $\boldsymbol{n}$ for four and six qubit tomography are shown.
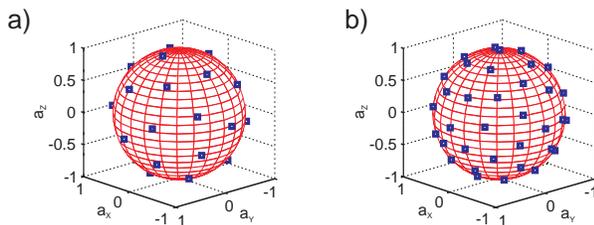

Figure 2.18: Optimal choice for measurements of local operators $A_j = \boldsymbol{n}_j \cdot \boldsymbol{\sigma}$, is got through a uniform distribution of D_n versors. This distribution of versors is here represented with blue squares on the Bloch sphere surface for 4qubit (a) and 6 qubit states (b) (from[45]).

2.4.3   Compressed sensing

Sometime a prior knowledge of the state may suggest the use of protocols otherwise ineffective, just like quantum state tomography with compressed sensing [30] when low rank states are under examination. Given a n qubit state of rank d, compressed sensing tomography allows to estimate the state with $m = cdr \log^2 d$ measurement settings and an error exponentially small in c. In fact, from a simple paramameter counting one can expect that $O(rd)$ could be enough to completely reconstruct the state but actually it is not clear how to achieve this performance in practice since mere knowledge of the rank of the state does not provide enough information on which measurements and efficiently reconstruct the density matrix.

2.4.4   Matrix product state tomography

Every N qubit state $|\psi\rangle$ can be defined through matrix product state representation [74]

$$|\psi\rangle = \sum_{i_1,\dots,i_n=0}^{1} \mathrm{Tr}\Big[A_{i_1}^{[1]} A_{i_2}^{[2]} \cdots A_{i_N}^{[N]}\Big] |i_1, i_2, \dots, i_N\rangle, \qquad (2.53)$$

where $A_{i_j}^{[k]}$ are $R \times R$ complex matrices with R bond dimension or local dimension of the state. A product matrix state of bond dimension R can be biseparable

into two states, each one with at most R qubits. Protocol starts by estimating, through standard state tomography, the reduced density matrix ρ_κ of the first $\kappa = \log_2(R) + 1$ qubits

$$\rho_\kappa = \mathrm{tr}_{\kappa+1,\dots,N}(\rho). \tag{2.54}$$

ρ_κ is then expressed through eigendecomposition

$$\rho_\kappa = \sum_{r=1}^{R} \lambda_r \, |\phi_r\rangle\langle\phi_r| \,, \tag{2.55}$$

please, note the sum goes to R instead of κ. In fact R is the bond dimension of ρ and thus of R_κ as well. From eq. (2.55) one see that there exists a density matrix with less qubits but same eigenvalues of ρ_κ. Then a local unitary operator acting on the first κ can disentangle the state ρ_κ

$$U = \sum_{s=0}^{1} \sum_{s'=0}^{2^\kappa-1} |s\rangle_1 \otimes |s'\rangle_{2,\dots,\kappa} \, \langle\phi_{s2^{\kappa-1}+s'+1}|_{1,\dots,\kappa} \,, \tag{2.56}$$

such that, if $|v\rangle$ is a pure state

$$U\,|\phi\rangle = |0\rangle_1 \, |v\rangle_{2,\dots,N} \,. \tag{2.57}$$

The protocol proceeds in the same way for qubits from the second to the $(\kappa+1)$th site, it considers a reduced κ qubit state and finds the operator U_2 able to disentangle it. Performing the procedures for $N - \kappa + 1$ states, the protocol provides a sequence of local unitary operators $U_1, \dots, U_{N-\kappa+1}$ and a separable state $U_{N-\kappa+1} \cdots U_1 |\phi\rangle = |0\rangle^{\otimes N-\kappa+1} \otimes |\eta\rangle$, with $|\eta\rangle$ $(\kappa-1)$ qubit state. Fig. 2.19 shows a scheme for this protocol. The original state is then constructed through the quantum circuit composed of unitary operations with small operation length and a pure qubit state.
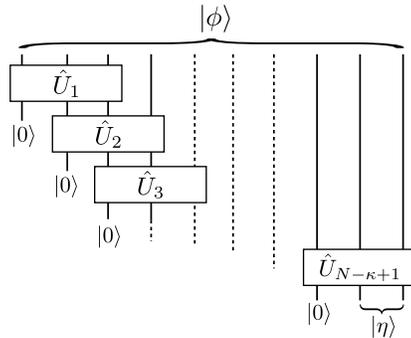

Figure 2.19: Sequence of unitary operators acts on three-qubit states. They end up to disentangle $|\phi\rangle$ into $|0\rangle^{\otimes N-\kappa+1} \otimes |\eta\rangle$, where $|\eta\rangle$ is a two-qubit pure states (from [74]).

Chapter 3

Adaptive scheme for tomography

This chapter presents a protocol for incomplete quantum state tomography. It adaptively selects the measurements to perform in order to guarantee a larger mean value of the fidelity between reconstructed and target state compared to the value one gets with a scheme where measurements at each step are chosen independently from the results of previous measurements. A specific adaptive procedure has been developed and tested both for standard tomography which exclusively employs product measurements and for tomography with entangled measurements on mutually unbiased bases. The target states used in the simulations are pure states of up to four qubits. The sampling of the states is discussed in the first section of this chapter.

3.1 Sampling of physical states

Sampling of physical states is indispensable in tomography simulation and comparison of different protocols, both to create a distribution of target states on which to perform the measurements but also to extract the reconstructed state at the end of the tomographic procedure. Moreover the adaptive scheme proposed in this thesis mainly relies on the sampling of a distribution of states with fixed constraints. In order to describe how Monte Carlo simulation [75] is employed in the sampling of quantum states, one can start by observing that a general measurement in quantum mechanics is a probability-operator measurement with outcomes $\Pi_1, \ldots \Pi_K$ which by definition are non-negative and with up to unity sum

$$\Pi_j \leq 0; \quad \sum_{j=1}^{K} \Pi_j = \mathbb{1}. \tag{3.1}$$

According to Born's rule the probability to get a detection at the $j-$th outcome for some measurements of a physical state ρ is

$$p_j = \mathrm{Tr}\left(\Pi_j \rho\right). \tag{3.2}$$

If state ρ is unambiguously determined by probabilities $p = (p_1, \ldots, p_K)$ then the positive operator-valued measures Π_j are called informationally complete. Best known example of informationally complete positive operator-valued measures are the symmetric ones [76], such Π_j fulfill

$$\mathrm{Tr}(\Pi_i \Pi_j) = \frac{d\delta_{ij} + 1}{d + 1}, \tag{3.3}$$

with dimension $d = 2^n$ of the n-qubit Hilbert space. The question of existence of symmetric informationally complete positive operator-valued measures in any dimension is still an open problem. However, for the single qubit case $d = 2$ one can rely on the Bloch sphere representation where symmetric operators correspond to the vertices of a tetrahedron as shown in Fig. 3.1. In particular they can be put in correspondence with the Bloch parameters $x = \langle \sigma_x \rangle$, $y = \langle \sigma_y \rangle$ and $z = \langle \sigma_z \rangle$ as

$$\Pi_1 = \frac{1}{4}\left[1 + \frac{1}{\sqrt{3}}\left(x + y + z\right)\right], \tag{3.4}$$

$$\Pi_2 = \frac{1}{4}\left[1 + \frac{1}{\sqrt{3}}\left(-x - y + z\right)\right], \tag{3.5}$$

$$\Pi_3 = \frac{1}{4}\left[1 + \frac{1}{\sqrt{3}}\left(x - y - z\right)\right], \tag{3.6}$$

$$\Pi_4 = \frac{1}{4}\left[1 + \frac{1}{\sqrt{3}}\left(-x + y - z\right)\right]. \tag{3.7}$$

More generally, one also can consider all the possible unitary rotations of the sphere

$$\Pi_j = \frac{1}{4}\left(\mathbb{1} + \boldsymbol{a}_j \cdot \boldsymbol{\sigma}_j\right), \tag{3.8}$$

with vectors $\boldsymbol{a}_j, j = 1, \ldots, 4$ corresponding to the four legs of the tetrahedron. Then, one can study how a state can be actually sampled starting from a distribution for these probabilities. Consider the simple case of a qubit state measured by a three-outcome trine positive operator with probabilities depending on Bloch parameters as

$$p_1 = \frac{1}{3}\left(1 + x\right), \quad p_{2,3} = \frac{1}{3}\left(1 - \frac{1}{2}x \pm \frac{\sqrt{3}}{2}y\right). \tag{3.9}$$

These probabilities further fulfill the constraint $\sum_{k=1}^3 p_k^2 \leq \frac{1}{2}$. All these constraints can be summarized in the function $w_{\mathrm{cstr}}(p)$ which is a product of step
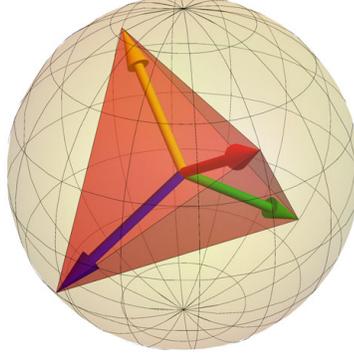
Figure 3.1: Symmetrically informationally complete measurements for single qubit states correspond to the vertices of a thetraedron in the Bloch sphere representation (from [77]).

functions η_{p_k} and delta functions and it is vanishing if p is not permissible. For example, for the basic constraints, we have

$$w_{\text{basic}}(p) = \eta_{p_1} \eta_{p_2} \cdots \eta_{p_K} \delta \left(1 - \sum_k p_k \right), \tag{3.10}$$

as factor in

$$w_{\text{cstr}}(p) = w_{\text{basic}}(p) w_{\text{qu}}(p), \tag{3.11}$$

with $w_{\text{qu}}(p)$, referring to others possible quantum constraints, in the current example of trine measurement $\sum_{k=1}^{3} p_k^2 \leq \frac{1}{2}$. More generally if one uses $\rho = \sum_k p_k \Lambda_k$, and writes $\Lambda_k = a\Pi_k + b$, trace condition (1.17) naturally leads to [78]

$$\sum_k p_k^2 \leq \frac{1 - b}{a}. \tag{3.12}$$

For example, the qubit problem with the tetrahedron measurement has the physicality constraint

$$\sum_{k=1}^{4} p_k^2 \leq \frac{1}{3}. \tag{3.13}$$

Thanks to global constraints in (3.11), one sees that the volume element of the probability space is

$$(\mathrm{d}p) = \mathrm{d}p_1 \mathrm{d}p_2 \cdots \mathrm{d}p_K w_{\text{cstr}}(p), \tag{3.14}$$

while the volume element of the infinitesimal vicinity of state ρ turn out to be

$$(\mathrm{d}\rho) = (\mathrm{d}p) w_0(p), \tag{3.15}$$

where $w_0(p)$ is called prior density. The prior density is arbitrarily chosen. One common option for prior density is the so-called primitive prior; with this choice the density is set to be uniform over the physical probability space p

$$w_{\text{primitive}}(p) = 1. \tag{3.16}$$

Anther common choice used when no external prior information is available [79] is the Jeffreys prior [80]

$$w_{\text{Jeffreys}}(p) = \frac{1}{\sqrt{p_1 p_2 \cdots p_K}}. \tag{3.17}$$

In general, it is not straightforward to sample directly from the target distribution $w_{\text{t}}(p) = w_{\text{constr}}(p)w_0(p)$ of probabilities which ensure physicality of the states, but since one can sample over the probability space with a known reference distribution $w_{\text{r}}(p)$, it is possible to approach the target distribution by means of rejection sampling. Such a sampling belongs to the class of independence sampling where points are generated independently of one another. The factor $r(p)$ that relates the target distribution to the reference distribution like

$$w_{\text{t}}(p) = w_{\text{r}}(p)r(p) \tag{3.18}$$

can be regarded as the ratio $r(p) = w_{\text{t}}(p)/w_{\text{r}}(p)$. With the rejection sampling method, first the probabilities p are uniformly sampled from the space of probabilities with the only constraints of their positivity and unit sum, as specified by the factor $w_{\text{basic}}(p)$ of (3.10); then one draws many sample points according to the chosen reference distribution $w_{\text{r}}(p)$, and discards or accepts points in such a way that the remaining sample points are distributed according to the target distribution $w_{\text{t}}(p)$. At the end one finds that sample points $p^{(j)}$ are accepted with probability equal to the so-called acceptance ratio

$$a = \frac{r(p^{(j)})}{R}, \tag{3.19}$$

where $R = \max_p\{r(p)\}$. Rejection sampling requires to discard points in accordance with the acceptance ratio thus one ends up with fewer sample points than the initial set drawn from $w_{\text{r}}(p)$. For example, for a three-outcome trine measurement on the single qubit of (3.9), only $\pi/\sqrt{27} \approx 60.5\%$ of the points sampled from are physical (see Fig. 3.2)

The yield decreases as the dimensionality of the system increases: for a nine-outcome measurement on a qubit-pair only about 10% of the points are physical [81]; and for the sixteen-outcome only one in $5 \cdot 10^4$ candidate points is accepted.

Pure states   In this thesis all the tomography simulations are realized on pure target states. In fact pure states can be sampled uniformly according to Haar measure [82] while for mixed states it becomes less obvious to find a fair distribution to sample. Moreover observation made in permutationally invariant tomography 2.4.2 regarding most common state under examination hold here
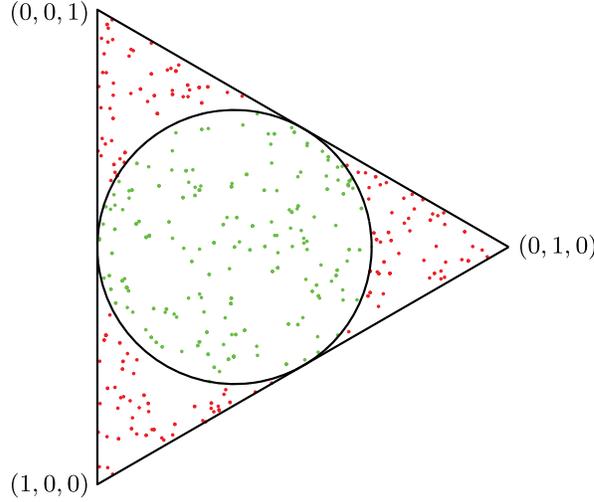
Figure 3.2: Distribution of probabilities in the probability space (p_1, p_2, p_3). The triangle contains all the triples of probabilities which fulfill $\sum_i p_i = 1$. The circle inscribed in the triangle delimits those which fulfill the quantum constraint $\sum_i p_i \leq \frac{1}{2}$ as well (from [75]).

as well for pure states which are indeed often employed. In order to sample pure states like $|\psi\rangle = \sum_{j=1}^{d} \psi_j \, |c_j\rangle$ one can start by observing that Born's probabilistic interpretation of the state vector requires it to be normalized

$$\||\psi\rangle\| \equiv \sqrt{\langle\psi|\psi\rangle} = \sqrt{\sum_{j=1}^{d} |\psi_j|^2}. \tag{3.20}$$

As the values $|\psi_j|^2$ are non-negative and sum up to one, they form a probability distribution $p_j \equiv |\psi_j|^2$. Thus one can write

$$|\psi_j\rangle = \sum_{j=1}^{d} \sqrt{p_j} \exp\{i\theta_j\} \, |c_j\rangle, \tag{3.21}$$

where $\psi_j = |\psi_j| \exp(i\theta_j)$ and phase $\theta_j \in [0, 2\pi]$ was used. The task now is reduced to create an unbiased random discrete probability distribution starting from a random number generator yielding random numbers with uniform distribution in $[0, 1]$. In order to do so, the following procedures can be applied: a random variable q_1 is picked from a uniform distribution of probabilities in the interval $[0, 1]$, afterwards a second variable q_2 is extracted from a uniform distribution in the interval $[0, 1 - q_1]$ and so on for d probabilities, with the $j-$th variable q_j picked from the interval $[0, 1 - \sum_{k=1}^{j-1} q_k]$. Variable q_j so far generated do not form a uniform distribution in the $d-$dimensional volume since

by construction $q_1 > \cdots > q_i > q_{i+1} > \cdots > q_n$, however one can consider a random permutation $\{k_1, \ldots, k_d\}$ of the indices $\{1, \ldots, d\}$ to create the desired unbiased distribution

$$\{p_1, \ldots, p_d\} = \{q_{k_1}, \ldots, q_{k_d}\}. \tag{3.22}$$

Fig.3.3 shows how probabilities are distributed for $d = 3$. Once got the distribution of probabilities $\{p_1, \ldots, p_d\}$, it is sufficient to plug their values into Eq.(3.20) and pick a random value for the phase θ from a uniform distribution between $[0, 2\pi]$ in order to sample all pure states $|\psi_j\rangle$.
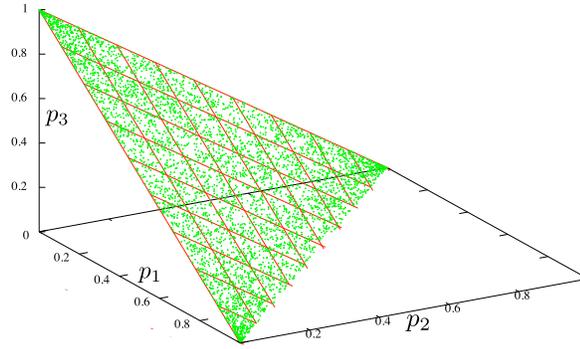

Figure 3.3: Uniform distribution of $5 \cdot 10^3$ probabilities vectors $\boldsymbol{p} = (p_1, p_2, p_3)$ which fulfill the constraint $\sum_i p_i = 1$ generated through normalization method (from [83]).

This last method was used indeed to generate all target states in the simulations described in this chapter. Fig. 3.4 ishows the distribution of 1000 states generated in such a way.

Mixed states A uniform distribution of physical states on the surface of the Bloch sphere is regarded as an optimal distribution to sample pure states. On the other side, for mixed state, it not as obvious as for pure states which distribution the sampling should follow. Even though this reasoning led to the choice to compare all the schemes over pure target states, state reconstruction picks from mixed ones as well. A straightforward method to sample mixed states is made up of just two steps. It starts generating a given distribution of $2n$-qubit pure states and then traces out n qubits from each of these ones; Fig. 3.5(a) shows an example for a distribution of single qubit mixed states generated with this method.

Another commonly used method [84] arises from the observation that every physical state can be spectrally decomposed as

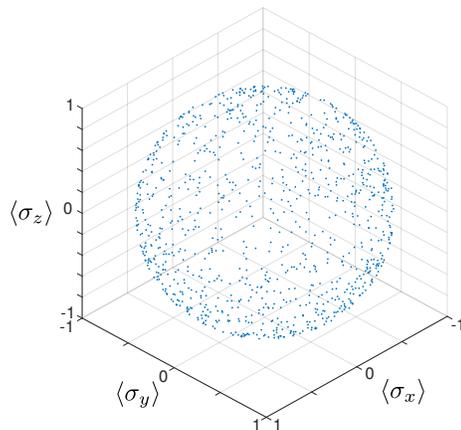$$\rho = \sum_{j=1}^{d} r_j \, |r_j\rangle\langle r_j| \,, \tag{3.23}$$

Figure 3.4: One thousand pure states in the Bloch sphere representation are distributed according to the Haar measure.

thanks to (1.17) one can conclude that $d - 1$ real parameters have to be used in order to identify all the eigenvalues $\{r_1, \dots, r_d\}$ of the state. Moreover, any two bases of the vector space $\mathbb{C}^d$ are connected by an unitary matrix U through the relation $|r_j\rangle = U |c_j\rangle$ with $\{|c_j\rangle\}$ computational basis. Therefore the sampling of the density matrices reduces to the creation of a random discrete probability distribution $\{r_1, \dots, r_d\}$ and of a random unitary $d \times d$ matrix U. The number of independent parameters is $d^2 - 1$ since it is equal to the sum of $d - 1$ independent parameters of probabilities distribution $\{r_1, \dots, r_d\}$ and $d^2 - d$ independent parameters of a unitary $d \times d$ matrix. The parametrization of random unitary matrices is non unique [85], in Fig. 3.5(b) the states are sampled with Hurwitz parametrization [86] which for single qubit states reads

$$U = \begin{pmatrix} e^{i\psi} \cos \phi & e^{i\chi} \sin \phi \\ -e^{-i\chi} \sin \phi & e^{-i\psi} \cos \phi \end{pmatrix}, \tag{3.24}$$

where $\phi = \arcsin \xi$, while $\xi \in [0, 1)$, and $\phi, \chi \in [0, 2\pi)$ are uniformly generated variables.

Finally one can recall Cholesky parameterization of Sec. 2.3, which allows to generate any physical state just starting from a distribution of vectors $\boldsymbol{t} \in \mathbb{R}^{4^n}$. In the example of Fig. 3.5(c), t_i are variable piked from a normal distribution with mean 0 and standard deviation equal to 10^2.

3.2 Standard tomography

This section describes the algorithm built to simulate the working and test performances of a tomographic scheme which is commonly used to investigate n qubit states in polarization encoding. In fact, such standard tomography serves as basis for adaptive tomography and more notably, as benchmark for
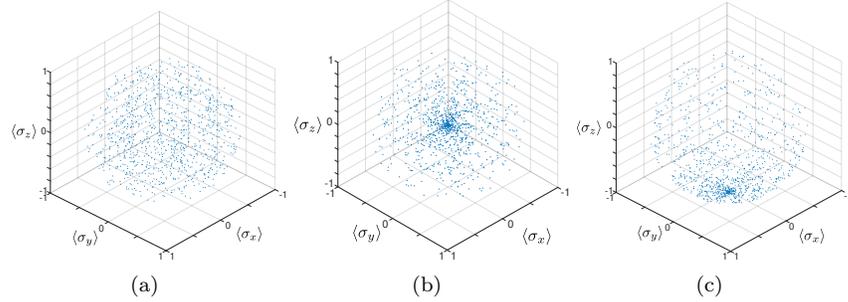
Figure 3.5: One thousand mixed state in Bloch sphere representation sampled with three different algorithms.

the efficiency of this last one. From now on, all the simulations are realized for $2, 3$ and 4 qubit states but for the sake of simplicity, examples refer only to two qubit states. The algorithm starts by sampling a n qubit pure state ρ_T from a uniform distribution in the Haar measure [82]; ρ_T acts as target state one wants to examine. Standard tomography works only with separable measurements, then each correlation element is obtained from the product of projection measurements on single qubit states just as shown in (2.9). The first correlation elements to be measured is picked arbitrarily from those with only non-zero indices; others are in fact derived from the former as shown in (2.10) and (2.11). Here, the operator $XX \equiv \sigma_x \otimes \sigma_x$ is the first to be measured. Then, the algorithm set up the four eigenstates of XX, which correspond to the four possible tensor products of the two single qubit eigenstates. Then the algorithm simulates an infinite number of projection measurements on these eigenstates, so that the three correlations elements are derived without uncertainty. The operators IX and XI indeed are measured simultaneously with XX since they share the same eigenstates. Maximum likelihood reconstruction completes the simulation of the first step of standard quantum state tomography. Finally, the fidelity between the reconstructed state and the target state is calculated in order to measure the performance of the scheme. The second step is implemented in the same way, but the set of parameters to be measured is chosen in such a way that avoids redundant information; in particular one seeks to perform projective measurements on a different set of single qubit eigenstates. In the present case the correlation element YY is used but every correlation which did not contain the operator X is in principle equally acceptable. The same reasoning forces the third correlation element to be ZZ. From the forth step on however, one has necessarily some redundancy of measurements since all three possible projections are already performed; this means that for two qubit states and from the forth step on, the six correlation elements with only non zero indices are a priori equivalent. In the present case XY and YX are chosen for the forth and fifth step respectively. For n qubit states with $n \geq 3$, this is no longer true since one should avoid already performed projections which provide

information about already known correlation elements with m non-null indices for $m = 2, \ldots n - 1$ as well. Once one has completed the simulation for one target state, the whole procedure is applied in the same way for all the others, such that one gets a distribution of fidelities at each step. In particular one can observe a growth of the mean fidelity of the reconstructed states after each step. The results one finds in such a way, however are on average independent on the specific order chosen for the measurements since the result since all the simulation are performed on a uniform distribution of target states.

3.3 Adaptive tomography

One might wonder if there exists any less arbitrarly approach of choosing correlation elements to measure in order to increase the fidelity between the maximum likelihood reconstructed state and the target state. This section offers a proposal for a scheme where measurements are indeed chosen adaptively according to results of preceding steps. The adaptive scheme is constructed from the standard one such that the first three steps are actually identical to the latter; in fact since the information one can get from few correlation elements is really small, it is inefficient to introduce adaptivity at earlier step, especially when one is considering target states in high dimensional Hilbert space. After three steps, the experimenter knows $3 \cdot (2^n - 1)$ correlation elements of the target state. The algorithm then creates through Cholesky parameterization a distribution of n_{samples} physical states $\{\rho'_i | i = 1, \ldots, n_{\text{samples}}\}$ which lie on the subspace defined by fixing the already measured parameters. The starting parameters $\boldsymbol{t}_i$ used for the Cholesky parametrization are arbitrarily chosen to provide a uniform sampling of pure state according to the Haar measure. Each ρ'_i built in this way belongs to the same subspace of ρ_T and one might expect that they share similar properties; for example one can simulate a tomographic step on the states ρ'_i and look how fidelities of maximum likelihood reconstructed states increase after a given measurement. Specifically, the algorithm picks a state ρ'_j, calculates the expectation value of M_k and performs maximum likelihood reconstruction starting from parameters obtained by the first three measurement on the true state. The expectation value is calculated in such a way that it reconstructs the state $\rho_{j,k}^{\text{ML}}$. Finally one computes fidelity between ρ'_j and $\rho_{j,k}^{\text{ML}}$: $F_{j,k} = F(\rho_{j,k}^{\text{ML}}, \rho'_j)$. The procedures applies identically for all $j = 1, \ldots, n_{\text{samples}}$ and $k = 1, \ldots, 3^n$. The amount of information each measurement M_k provides is estimated in this scheme by the sum of fidelities $F_{j,k}$ over all states ρ'_j

$$F_k = \sum_{j=1}^{n_{\text{samples}}} F_{j,k}. \tag{3.25}$$

The best measurement $M_{k_{\text{next}}}$ is thus the one which provides the maximum value of F_k

$$M_{k_{\text{next}}} = \operatorname*{argmax}_{k'} F_k. \tag{3.26}$$

The measurement $M_{k_{\text{next}}}$ is then performed on the target state at the forth step. It appears that the fidelity of the reconstructed state computed after measurement M_l^* increases on average compared to the fidelity one gets when measurement is chosen in advance as in standard tomography. The whole procedure is schematized in the flowchart of Fig. 3.6.
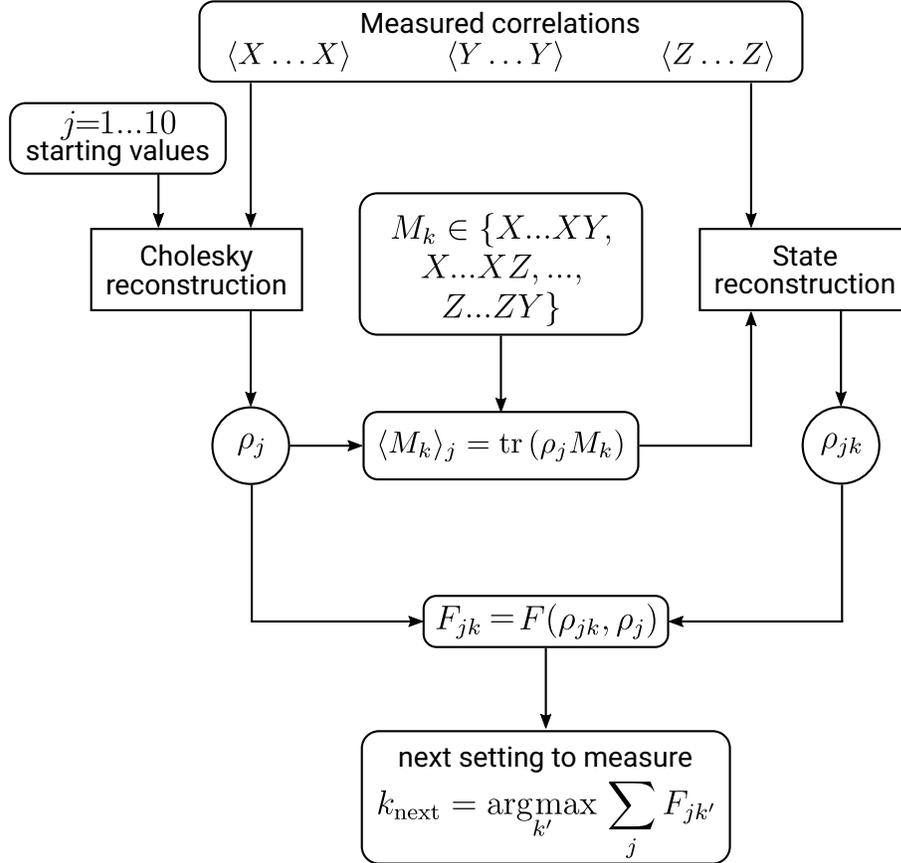
Figure 3.6: Flowchart of the adaptive scheme for quantum tomography. State reconstruction algotithm so far considered was the conjugate gradient-accelerated product gradient [61] but the scheme can be straightforwardly generalized with any other algorithm. Similar reasoning holds true for the Cholesky parametrization used for the sampling of the states with given constraints.

The same adaptive procedure is applied to the following steps. Actually, two conflicting phenomena, turn out to play a role in the efficiency of the adaptive scheme depending on the step one is considering; on one side the performance of adaptive tomography grows with the number of measurements, since states ρ_j' resemble target states better and better, on the other side, the fidelities computed in standard tomography grow as well approaching the maximum value

of 1 and thus bound the difference between schemes stronger at each step. Trends of distributions of fidelities for two-qubit states are shown with the boxplot method [87, 88] in Fig. 3.7 both for standard and adaptive scheme. Each box contains 50% of the occurrences. The plotted whisker extends to the up to 1.5 the height of the box. All the point outside the whiskers are considered outliers and are marked with a red cross. The notch departs symmetrically from the median and extend to $\pi/\sqrt{N}$ times the height of the box, where N is the number of computed fidelities. For the forth and the fifth step, the graph shows a comparison for standard (yellow) and adaptive (light blue) scheme.

Figure 3.7: Trend of fidelity after each measurement for two-qubit target states. Red horizontal line marks the median of a distribution of 2000 fidelities. The sequence of sets of correlations elements measured is $\{XX, XI, IX\}$, $\{YY, YI, IY\}$, $\{ZZ, ZI, IZ\}$, $\{XY, XI, IY\}$, $\{XY, YI, IX\}$. For the adaptive procedure in the light blue columns the order of measurements at the last two steps is not fixed. The distributions of fidelities for the lats two steps is diplayed with more detail in Fig. A.1 and Fig. A.2 of Appendix A.

One can observe that the number of outliers decreases significatively with the adaptive procedure and the mean of fidelities computed over 2000 states at the forth step goes from 95.96% with standard scheme to 99.84% with adaptive one while the means of fidelities computed from at the fifth step goes from 99.83% to

100.00%. Remarkably in this case the adaptive protocol allows to reach already at the forth step a mean value for the fidelity of the reconstructed state that is even larger than the value one findes at the fifth step with standard tomography. Analogous simulations were performed for three-qubit and four-qubit states. The mean of fidelities for 3000 three-qubit target states turns out to be 95.24% at the forth step and 98.61% after the fifth measurement for the standard scheme while with the adaptive one they are 99.90% and 100.00%. Again, if one look at the mean value of the fidelities the adaptive scheme overcome already at the forth step the results obtained at the fifth step with the standard scheme. Fig. 3.8 illustrated these results for three-qubit states.
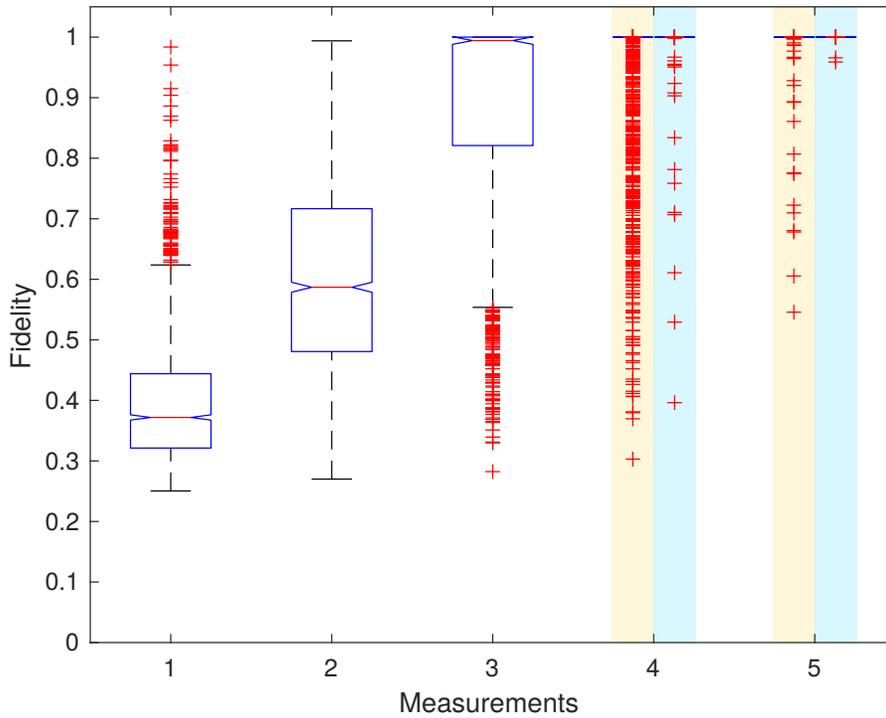

Figure 3.8: Trend of fidelity after each measurement for three-qubit target states. Red horizontal line marks the median of a distribution of 3000 fidelities. The sequence of full correlations measured is XXX, YYY, ZZZ, ZYX, XZY. For the adaptive procedure in the light blue columns the full correlations to measure at the last two steps are not fixed. The distributions of fidelities for the lats two steps is diplayed with more detail in Fig. A.3 and Fig. A.4 of Appendix A.

Finally for four-qubit states the tomographic scheme has been expanded to a sixth step, since the fidelity computed up to the fifth step it is not as large as for two-qubit and three-qubit cases, then one can look at an appreciable

difference between the schemes at the sixth step as well. Indeed one finds that with standard tomography the means of fidelities over 1000 target states are 49.11%, 64.35%, 75.72%, at forth, fifth and sixth step, respectively. When adaptive protocol is used the means of fidelities for four-qubit states at the last three steps are 54.95%, 74.66%, 85.76%, this means that one has an increasing of fidelity is on average large than 10% at the fifth and last step. Distribution of fidelities computed at each step for four-qubit states are displayed in Fig. 3.9.
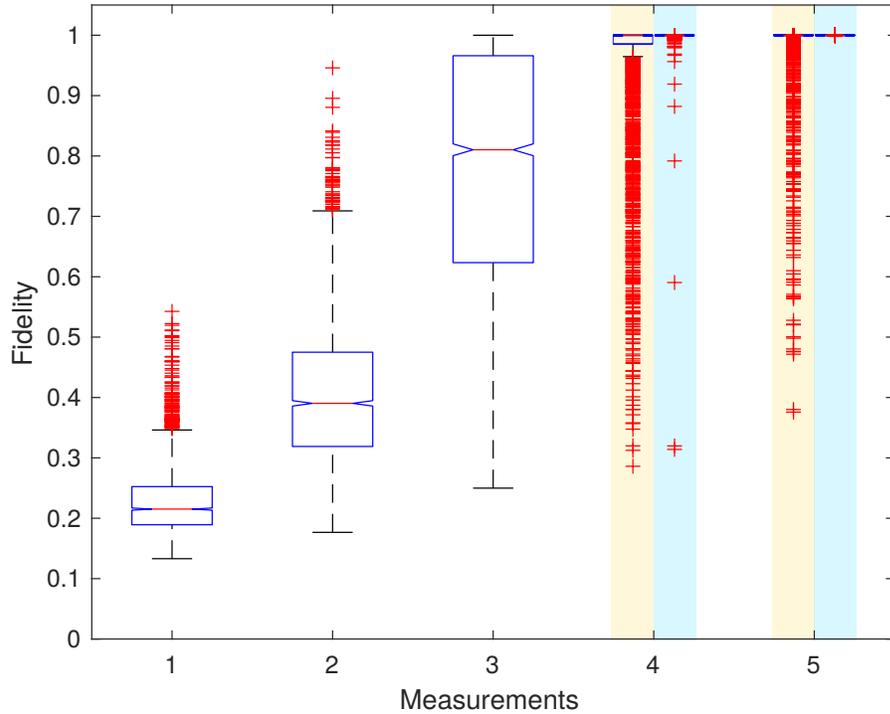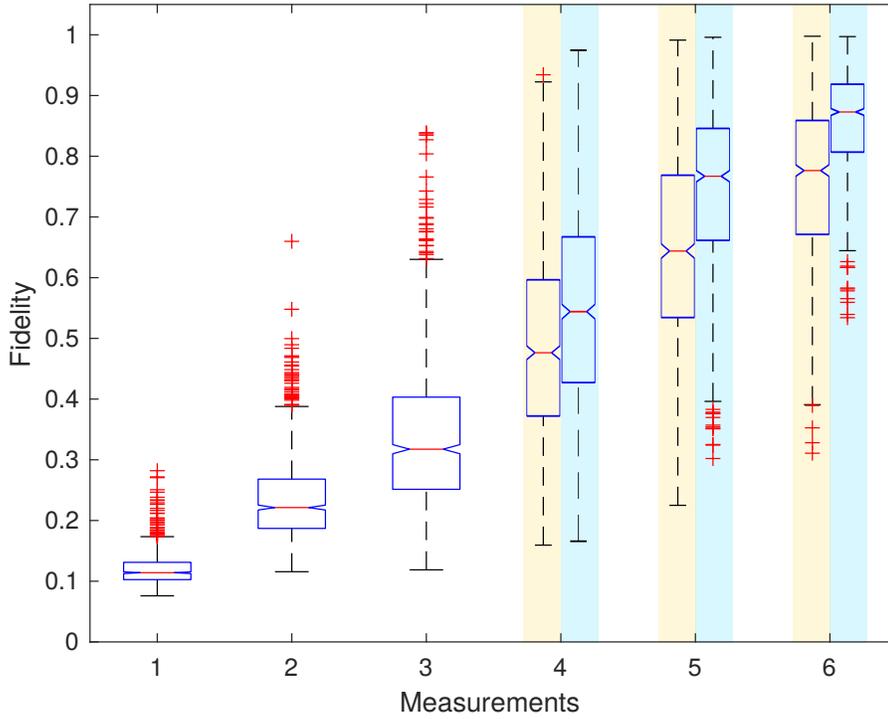

Figure 3.9: Trend of fidelity after each measurement for four-qubit target states. Red horizontal line marks the median of a distribution of 1000 fidelities. The sequence of full correlations measured is $XXXX$, $YYYY$, $ZZZZ$, $ZXYZ$, $YZXY$, $XYZX$. For the adaptive procedure in the light blue columns the full correlations to measure at the last two steps are not fixed.

The results of mean fidelities computed after each step both for standard and adaptive tomography with product measurements are summarized in table 3.1.

No.	Two qubits		Three qubits		Four qubits	
1	39.53%		22.72%		12.00%	
2	60.34%		40.71%		23.46%	
3	89.05%		77.78%		34.00%	
4	95.96%	99.84%	95.24%	99.90%	49.11%	54.95%
5	99.83%	100.00%	98.61%	100.00%	64.35%	74.66%
6					75.72%	85.77%

Table 3.1: Mean value of the fidelities obtained from tomography with separable measurements. The means are calculated over 2000 two-qubit states, 3000 three-qubit states and 1000 four-qubit states. The first three rows refer to results shared by standard and adaptive scheme, from the forth step on the left column corresponds to the value obtained with standard scheme while the right column corresponds to the value obtained with adaptive scheme.

3.3.1   Correlation complementarity

The adaptivity of the new scheme relies on interdependence and constraints existing between anticommuting operators which can be summarized in the concept of correlation complementarity [89]. Consider a set of traceless and trace-orthogonal dichotomic Hermitian operators A_k just like generalized Pauli operators considered so far, and let α_k real numbers in the range $[-1, 1]$ be their expectation values over a state ρ

$$\alpha_k \equiv \mathrm{Tr}(A_k \rho). \tag{3.27}$$

At this point one can arrange the observables in disjoint sets of anticommuting observales $S_j = \{A_1^{(j)}, A_2^{(j)}, \dots\}$ such that

$$\{A_k^{(j)}, A_l^{(j)}\} \equiv A_k^{(j)} A_l^{(j)} + A_l^{(j)} A_k^{(j)} = \mathbb{1}(1 + \delta_{kl}) \quad \forall k, l \in \{1, \dots, |S_j|\}. \tag{3.28}$$

Operators $A_k^{(j)}$ and their corresponding expectation values α_{jk} are then employed to define another operator F like

$$F \equiv \sum_{k=1}^{|S_j|} \alpha_{jk} A_k^{(j)} = \boldsymbol{\alpha}_j \cdot \boldsymbol{A}_j. \tag{3.29}$$

Then, making use of the anticommutation rules (3.28) one finds that the variance of F with respect to the state ρ is given by

$$
\begin{aligned}
\langle F^2 \rangle - \langle F \rangle^2 &= \mathrm{Tr}\left(\rho F^2\right) - \left[\mathrm{Tr}\left(\rho F\right)\right]^2 \\
&= \mathrm{Tr}\left(\rho \sum_{k=1}^{|S_j|} \alpha_{jk}^2 A_k^{(j)} A_k^{(j)}\right) - \left[\mathrm{Tr}\left(\rho \sum_{k=1}^{|S_j|} \alpha_{jk} A_k^{(j)}\right)\right]^2 \\
&= \sum_{k=1}^{|S_j|} \alpha_{jk}^2 \,\mathrm{Tr}\left(\rho A_k^{(j)^2}\right) - \left[\sum_{k=1}^{|S_j|} \alpha_{jk} \,\mathrm{Tr}\left(\rho A_k^{(j)}\right)\right]^2 \\
&= \sum_{k=1}^{|S_j|} \alpha_{jk}^2 - \left[\sum_{k=1}^{|S_j|} \alpha_{jk}^2\right]^2 = |\boldsymbol{\alpha}_j|^2 \left(1 - |\boldsymbol{\alpha}_j|^2\right)
\end{aligned}
\tag{3.30}
$$

Finally, the non-negativity of the variance leads to a constraining equation for the sum of the expectation values of anticommuting operators

$$
|\boldsymbol{\alpha}_j|^2 = \sum_{k=1}^{|S_j|} \alpha_{jk}^2 < 1.
\tag{3.31}
$$

Eq. (3.31), indeed comes on top of already measured correlation elements to restrict the subspace of parameters where states ρ' are sampled and enables to guess which is operator is best to measure in next step.

3.4 Tomography with mutually unbiased bases

Product measurements as those in (2.9) were performed so far on an arbitrarily large set of copies of one ensemble to derive the correlation element $T_{i_1,\ldots,i_n}$ of the state. These measurements project the state onto a product of single qubit eigenstates of Pauli operator σ_i with $i = 1, 2, 3$; it is worth remembering in fact that correlation elements with zero indices cam be derived without additional measurement effort. After three product measurements however one faces necessarily some redundancy such that a complete tomography is reached only after 3^n measurements. In order to avoid redundancy and lower this value one could group all $4^n - 1$ operators in disjoint and maximal sets of commuting operators. Commuting operators in fact share common eigenstates, this means that at least in principle they can be measured simultaneously.

Luckily, it turns out that such arrangement does exist, in particular it is possible to form $2^n + 1$ sets of $2^n - 1$ commuting operators even though some of these sets share non separable eigenstates which turn out to be much more challenging to measure. To see how n-qubit Pauli operators are actually arranged the theory of mutually unbiased bases turns out to be very helpful.

3.4.1   Mutually unbiased bases

Let $\mathcal{H}_d$ be a Hilbert space for n-qubit states, with total dimension of the space $d = 2^n$, and let $A = 1, 2, \dots$ denote basis sets in $\mathcal{H}_d$ with elements $|A, \alpha\rangle$, $\alpha = 1, 2, \dots d$. Two bases A and B are said mutually unbiased [90, 91] if a system prepared in any element of A has a uniform probability distribution of being found in any element of B

$$|\langle A, \alpha | B, \beta \rangle|^2 = d^{-1} \quad (A \neq B) \tag{3.32}$$

where individual bases are chosen to be orthonormal,

$$\langle A, \alpha | A, \beta \rangle = \delta_{\alpha\beta}. \tag{3.33}$$

In a $d-$dimensional Hilbert space, measurements within a particular basis set can yield only $d-1$ independent probabilities, but a physical state ρ is specified by $d^2 - 1$ real parameters. Hence one needs $d + 1$ distinct basis sets to provide the required total number of independent probabilities. Actually, one can show that the $4^n - 1$ generalized Pauli operators can be partitioned in $2^n + 1$ subsets each consisting of $2^n - 1$ commuting elements [92]. In fact Eq. (3.32) may be expressed in terms of projectors

$$P_\alpha^A = |A, \alpha\rangle\langle A, \alpha|, \tag{3.34}$$

as

$$\mathrm{Tr}\!\left(P_\alpha^A P_\beta^B\right) = 2^{-n}. \tag{3.35}$$

Now an operator $\mathcal{O}_a^A$ can be defined through its spectral decomposition

$$\mathcal{O}_a^A = \sum_{\alpha=1}^{2^N} \epsilon_{a\alpha} P_\alpha^A \tag{3.36}$$

where $\epsilon_{a\alpha}$ is defined as $2^n \times 2^n$ matrix made of 2^n orthogonal row vectors, of which one for example $a = 2^n$ has all components equal to $+1$ and others have an equal number of $+1$s and -1s. Components of each vector with index a corresponds to the eigenvalues of $\mathcal{O}_a^A$ Ignoring the row vector with 2^n elements being $+1$s which defines identity operator, one can conclude that for each $2^n + 1$ unbiased bases, others $2^n - 1$ rows which define an equal number of operators. These operators are mutually commuting since they shares common eigenstates, moreover they have eigenvalues ± 1 and are traceless by construction since these eigenvalues sum up to zero. Pauli matrices one wants to consider have exactly these properties. To complete the proof there is nothing left but to show that these operators $\{\mathcal{O}_a^A | a = 1, \dots, 2^n - 1 \wedge A = 1, \dots, 2^n + 1\}$ are unitary transformed Pauli operators. This can be easily shown using the fact that the set of operator $\mathcal{O}_a^A$ form an orthogonal set. Indeed for $A \neq B$

$$\mathrm{Tr}\!\left(\mathcal{O}_a^A \mathcal{O}_b^B\right) = \sum_{\alpha\beta} \epsilon_{a\alpha}\epsilon_{b\beta} \,\mathrm{Tr}\!\left(P_\alpha^A P_\beta^B\right) = 0, \tag{3.37}$$

where the property $\sum_\alpha \epsilon_{a\alpha} = 0$ for $a \neq 2^n$ was employed, while for $A = B$ one gets

$$\text{Tr}\left(\mathcal{O}_a^A \mathcal{O}_b^B\right) = \sum_{\alpha\beta} \epsilon_{a\alpha}\epsilon_{b\beta} \text{Tr}\left(P_\alpha^A P_\beta^B\right) = \sum_\alpha \epsilon_{a\alpha}\epsilon_{b\alpha} = 2^n \delta_{ab}. \tag{3.38}$$

The task now is to find a theory for grouping generalized n-qubit Pauli operators into $2^n + 1$ sets of mutually commuting elements. The partitioning of Pauli operators is not a trivial task at all, for example one might arrange first Pauli operators in such a way that it is impossible to group the remaining ones in sets of $2^n - 1$ mutually commuting elements [93]. As an instance, $2-$qubit Pauli operators can be arranged only in six sets of mutually unbiased bases, the complex structure of such sets is graphically shown in Fig. 3.10, with integers from 1 to 15 labelling all possible sets of commuting operators. The indices of each vertex corresponds to one set of three commuting operators as reported in table 3.2.
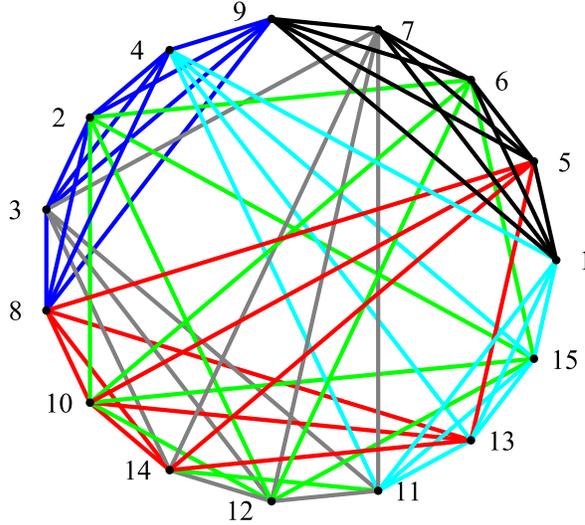
Figure 3.10: Six sets of mutually unbiased bases. Each vertex of the pentadecagon is labelled by a number corresponding to a set of three commuting operators like reported in table 3.2. Segments link mutually unbiased sets of operators. Five disjoint sets of operators are mutually linked by segments of the same colour (from [94]).

In order to see how to obtain these sets, one can start considering single qubit states; in that case each of the $2^1 + 1 = 3$ sets contains only $2^1 - 1 = 1$ operators, Pauli matrices are indeed unique elements of these sets, their eigenstates shown in Fig. 3.11 form in fact 3 mutually unbiased basis.

Note that any unitary operation preserves angles between axes of transformed operators, so one can redefine properly coordinates to have a swap for

No.	Commuting set	No.	Commuting set	No.	Commuting set
1	ZZ, ZI, IZ	6	YZ, ZX, XY	11	XY, XI, IY
2	ZY, ZI, IY	7	ZY, YX, XZ	12	YY, XX, ZZ
3	YZ, YI, IZ	8	ZZ, YX, XY	13	YZ, XX, ZY
4	YY, ZX, XZ	9	XX, XI, IX	14	ZX, ZI, IX
5	YY, YI, IY	10	XZ, XI, IZ	15	YX, YI, IX

Table 3.2: All possible sets of three commuting Pauli operators for two qubits states.
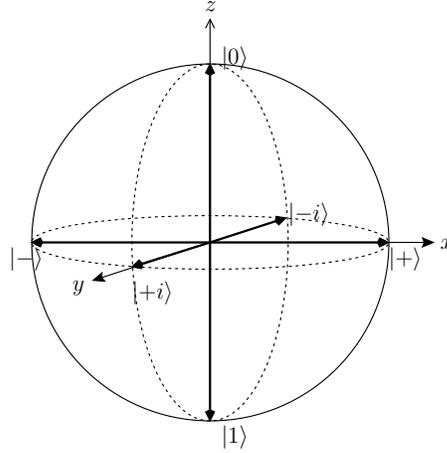
Figure 3.11: Elements of mutually unbiased bases on the Bloch Sphere (from [44]).

sets of Pauli matrices. The structure of the mutually unbiased bases is indeed invariant under any unitary transformation. This is akin to saying that only one mutually unbiased based structure exists in the two-dimensional Hilbert space. The same result holds as well for two qubit operators although, the extra feature of entanglement appears in addition. To see how the possible MUBs shown in Fig. 3.10 are obtained, one can consider 2 qubits basis as that reported in table 3.3.

No.	Commuting set	Sep.
1	XX, XI, IX	2
2	YY, YI, IY	2
3	ZZ, ZI, IZ	2
4	XY, ZX, YZ	1
5	YX, XZ, ZY	1

Table 3.3: Five sets of three operators defining a (3,2) MUB.

It can be easily observed that the first three bases are fully separable; that means that the operators belonging to the same set commute even when considering the single qubit Hilbert space separately. The last two bases on the contrary are not separable, in fact the common eigenstates of sets 4 and 5, respectively are maximally entangled states. Such a structure for mutually unbiased bases is labelled by the pair $(3, 2)$ indicating three biseparable and two nonseparable bases. Note that each separable basis has two operators containing the identity while a nonseparable basis cannot have any operator containing the identity and that there must be six entries, $IX, IY, \ldots, ZI$ containing the identity in the table, we can conclude that the $(3, 2)$ is the only possible construction in this space. In fact starting from three nonseparable sets it is not possible to arrange other operators in sets with three commuting elements. Any unitary transformation, local or nonlocal, will yield an isomorphic table with respect to the separability, except, for some row permutations. Moreover, the table is uniquely defined by the four entries in the two first columns of the first two rows. All other operators $O_{r,c}$ are determined by the relations $O_{r,c} = O_{r,c-2}O_{r,c-1}$, and $O_{r,c} = \mathcal{O}_{2,c}O_{1,c+r-3}$ for $r > 2$, where r and c denote the row and the column of the operator, respectively, and must be taken modulo four [93]. For more than two qubits the unitary transformation for one MUB does not allow to reach all others possible sets of bases and other transformations become necessary. 3 qubit Pauli operators can in fact form $(3, 0, 6), (1, 6, 2), (2, 3, 4)$ and $(0, 9, 0)$ mutually unbiased bases where the first index counts threeseparable bases, the second biseparable bases and the third nonseparable ones [95].

Similarly to quantum state tomography with product measurements we simulate and test tomography on mutually unbiased based, neglecting counting statistics error and computing fidelity of maximum likelihood reconstructed state measurement after each measurement for two, three and four-qubit pure target states. In particular, we fix mutually unbiased bases on which project since in principle they are all equivalent. For two-qubit states the mutually unbiased bases used are those reported in table 3.3. They are listed according to the order they were measured. Analogously the mutually unbiased bases used for three-qubit and four-qubit states are listed in table 3.4 and table 3.5.

In order to further increase the fidelity of the reconstructed state one may join advantages coming from mutually unbiased bases and adaptivity. In this case keeping fixed the sets of mutually unbiased bases, adaptivity helps in finding the best permutation for these sets of commuting operators. For example, the first three measurements for two-qubit states can be fixed to be $\{XX, IX, XI\}$; $\{YY, IY, YI\}$; $\{ZZ, IZ, ZI\}$ just like those listed in table 3.2. Adaptive procedure estimate which measurement among $\{XY, ZX, YZ\}$ and $\{YX, XZ, ZY\}$ can provide the estimated state which is more faithful compared with the target state. In order to show how good are the predictions of the adaptive procedure one can take into account the difference between the fidelity that the experimenter would calculate if the forth measurements was $\{XY, ZX, YZ\}$ and if the forth measurement was $\{YX, XZ, ZY\}$, this value is then compared with the difference expected by the adaptive scheme. One finds that the correlation between these two variables is of 87.4%. Fig. 3.12 shows a scatter plot of the

No.	Commuting set
1	$XXX, IZZ, ZIZ, ZZI, XYY, YXY, YYX$
2	$YYY, IXX, XIX, XXI, YZZ, ZYZ, ZZY$
3	$ZZZ, IYY, YIY, YYI, ZXX, XZX, XXZ$
4	$XYZ, XII, IYI, IIZ, XYI, XIZ, IYZ$
5	$YZX, YII, IZI, IIX, YZI, YIX, IZY$
6	$ZXY, ZII, IXI, IIY, ZXI, ZIY, IXY$
7	$ZYX, XZI, XIY, IZY, ZXZ, YXX, YYZ$
8	$XZY, YXI, YIZ, IXZ, XYX, ZYY, ZZX$
9	$YXZ, ZYI, ZIX, IYX, YZY, XZZ, XXY$

Table 3.4: Nine sets of mutually unbiased bases used for the simulation with three-qubit target states. The first five sest are listed according to the order in which they are measured.

No.	Commuting set
1	$XXXX, IIXI, IXII, XIII$
2	$YYYY, YZIZ, ZYZI, YXIX$
3	$ZZZZ, IIZI, IZII, ZIII$
4	$YYXY, YZZI, ZXII, XYZY$
5	$YYYX, YIII, IYZZ, YXZX$
6	$YXXX, XZZZ, ZXZI, XYIX$
7	$XYYY, YIZI, IYII, YYIX$
8	$YXYX, XZII, ZYII, YXZY$
9	$XYXX, YZII, ZXZZ, XXIY$
10	$YYXX, YIZZ, IXIZ, XYIY$
11	$YXYY, XIIZ, IYIZ, YXIY$
12	$XXXY, XZIZ, ZXIZ, XXZX$
13	$XYXY, YIIZ, IXZI, XXZY$
14	$YXXY, XIZI, IXZZ, XYZX$
15	$XXYX, XIZZ, IYZI, YYZY$
16	$XXYY, XZZI, ZYZZ, YYIY$
17	$XYYX, YZZZ, ZYIZ, YYZX$

Table 3.5: Seventeen sets of mutually unbiased bases used for the simulation with four-qubit target states. The first six sets are listed according to the order in which they are measured. Actually each set contains fifteen operators but here only the first four are displayed. All the other operators in fact can be retrieved as products of a power of the first four of the same set.

results for $2 \cdot 10^4$ pure target states.

For tomography via mutually unbiased bases one finds results which are analogous to those found with for tomography with separable measurements. In particular for the first three steps one cannot see any difference since the measurements with mutually unbiased bases includes product measurements as
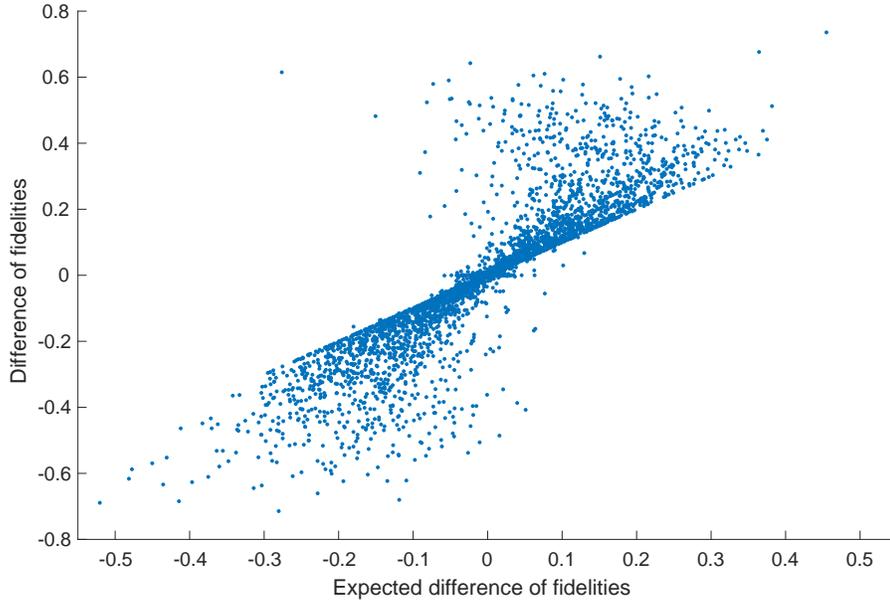
Figure 3.12: Expected difference between fidelity computed after the measure of $\{XY, YZ, ZX\}$ with the fidelity computed after the measure of $\{YX, ZY, XZ\}$ is compared with the difference one would actually obtain performing these measurements on the true state. For most of the states, if the expected difference is positive negative then the difference one actually measures on the true state. This means that the adaptive protocols correctly suggests the set of correlation to measure. Only for a small percentage of the state the measurement suggested by the adaptive procedure would provide a worse result for the fidelity of the reconstructed state.

one can see looking at table 3.3. From the forth step on the mean of the fidelities obtained with entangled measurements is appreciably larger than those obtained by employing only product measurements both for standard and adaptive scheme. Morover, here the simulation for two qubit states was realized up to the forth step. With this scheme, in fact, the quorum is reached at the fifth step and all the simulation would have provide fidelity equal to one. The results of mean fidelities computed after each step both for standard and adaptive tomography with entangled measurements are reported in table 3.6 .

No.	Two qubits		Three qubits		Four qubits	
1	40.01%		22.37%		12.03%	
2	61.21%		39.98%		23.70%	
3	89.64%		77.69%		34.07%	
4	98.51%	99.96%	98.95%	99.95%	54.79%	61.10%
5	100.00%	100.00%	99.88%	100.00%	71.95%	80.35%
6					82.36%	88.68%

Table 3.6: Mean value of the fidelities obtained from tomography with entangled measurements. The means are calculated over 3000 two-qubit states, 3000 three-qubit states and 1100 four-qubit states. The first three rows refer to results shared by standard and adaptive scheme, from the forth step on the left column corresponds to the value obtained with standard scheme while the right column corresponds to the value obtained with adaptive scheme. Note that at the five steps all the fifteen correlations of two-qubit states are already measured, thus the target state is recovered with fidelity one.

3.4.2   Experimental realization

In order to experimentally realize measurements on mutually unbiased bases, one needs to perform projective measuremenrts on entangled states like Bell states. Such measurements in, e.g., the basis of Bell states is in general a challenging task but there are few cases in which such measurements can be actually performed. Especially in low dimensions, $d = 4$ some linear optics setup is capable to perform such measurement. For two qubit states

For example for 2 qubits one can both employ one photon with 2 degrees of freedom like polarization and path degree of freedom [96], or consider two photons [13]: in this last case, however, since it is not possible to detect Bell states with photons in polarization encoding, one ends up with reduced efficiency to use this scheme [97, 98]. For more qubits physical realization becomes even more challenging; in these cases one has to rely on simulation of results performing mutually unbiased tomography on qu*d*its with through propagation modes of single photons [99].

Chapter 4

Conclusion and outlooks

This thesis described an experimental scheme for quantum state tomography able to improve the fidelity between the reconstructed state and the target one. The adaptive scheme cleverly selects the next measurements to perform according to the knowledge already acquired with the previous measurements on the true state. The adaptive algorithm here employed mainly relies on the sampling of physical states which are in agreement with the measurement outcomes held by the experimenter. The idea is that these state live in a subspace of the general Hilbert space of n-qubit states, and share common properties with the target state. In particular if one focuses on the correlation elements, some of them are fixed by the measurement already performed since any statistical error is neglected in the simulations, and many other correlation are constrained anyway according to anticommutation relations between Pauli operators. The conjugate gradient-adaptive projected gradjent algorithm was employed because of its speed, it used maximum likelihood as estimator function but one might observe that without statistical error this choice was indeed equivalent to any other physical estimator. In order to sample a distribution of physical mixed state in agreement with fixed parameters the algorithm used Cholesky parametrization with starting values uniformly distributed according to Haar measure in pure state space. All the target states chosen for the simulation were picked from a uniform distribution of n-qubit pure states with $n = 2, 3, 4$. Results for the difference of fidelities shows that there is an improvement in fidelity independently from the step considered which grows with the number of qubits both for standard tomography and for tomography via mutually unbiaesed bases. Since the idea behind the working of the adaptive procedure does not apparently rely on the particular schemes used one might study other cases by applying the same algorithm to other tomographic schemes, target states and estimators. Finally the promising results suggest a further research on the theoretical background possibly looking at the correlation complementarity which might explain and improve the efficiency of the adaptive procedure and show the potential of the scheme for many qubit states when computer simulation become more challenging.

Appendix A

Histograms for the comparison of simulations

Here are reported the histograms displaying a comparison of the distribution of the fidelities obtained with standard and adaptive scheme for tomography with separable measurements. In particular, Fig. A.1 and Fig. A.2 show the distributions of fidelities found at the forth step for two-qubit state tomography. Similarly, Fig. A.3 and Fig. A.4 show the distributions of fidelities found at the forth step for three-qubit state tomography.
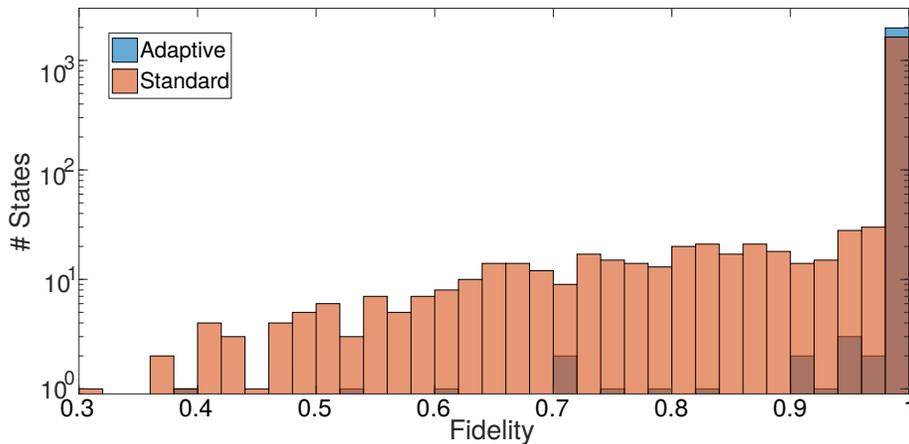

Figure A.1: Histogram of the fidelities obtained with standard scheme (orange) and adaptive scheme (light blue) at the forth step of two-qubit state tomography with product measurements.

Figure A.2: Histogram of the fidelities obtained with standard scheme (orange) and adaptive scheme (light blue) at the fifth step of two-qubit state tomography with product measurements. The leftmost bin collects all the occurences with fidelity smaller than 0.9997.

Figure A.3: Histogram of the fidelities obtained with standard scheme (orange) and adaptive scheme (light blue) at the forth step of three-q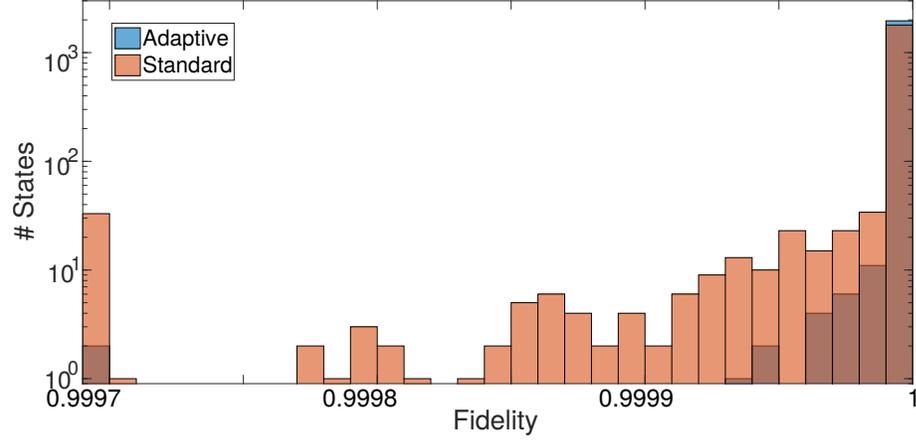ubit state tomography with product measurements. The inset zooms into the interval between 0.996 and 1 and refers only to the results obtained with the adaptive scheme. The leftmost bin collects all the occurences with fidelity smaller than 0.996.
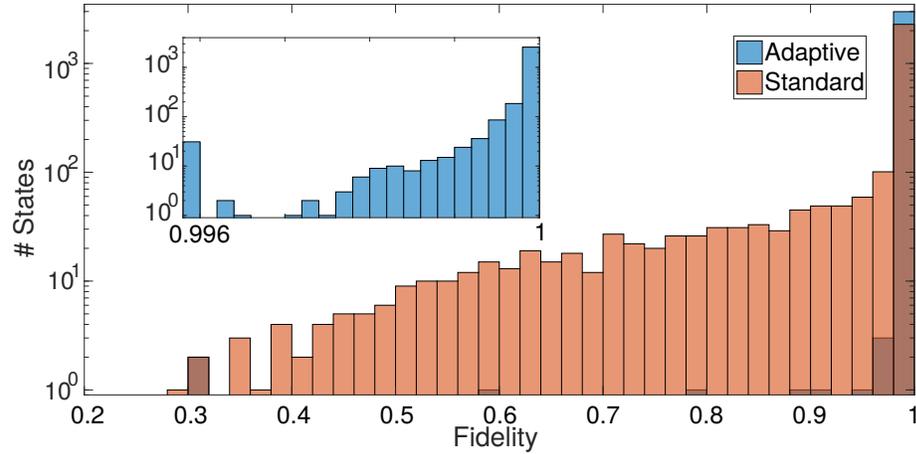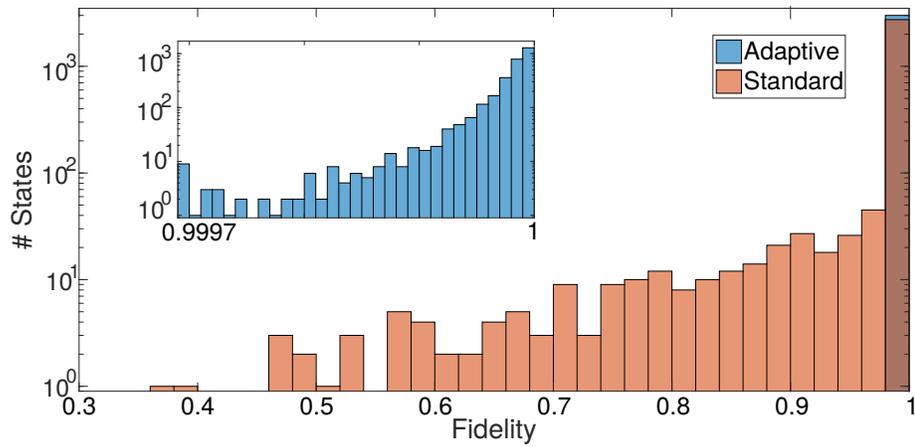
Figure A.4: Histogram of the fidelities obtained with standard scheme (orange) and adaptive scheme (light blue) at the fifth step of three-qubit state tomography with product measurements. The inset zooms into the interval between 0.9997 and 1 and refers only to the results obtained with the adaptive scheme. The leftmost bin collects all the occurences with fidelity smaller than 0.9997.

Bibliography

[1] J. Řeháček, D. Mogilevtsev, and Z. Hradil. "Tomography for quantum diagnostics". In: *New Journal of Physics* 10.4 (Apr. 2008), p. 043022. DOI: 10.1088/1367-2630/10/4/043022. URL: https://doi.org/10.1088%2F1367-2630%2F10%2F4%2F043022.

[2] G. M. D'Ariano and H. P. Yuen. "Impossibility of Measuring the Wave Function of a Single Quantum System". In: *Phys. Rev. Lett.* 76 (16 Apr. 1996), pp. 2832–2835. DOI: 10.1103/PhysRevLett.76.2832. URL: https://link.aps.org/doi/10.1103/PhysRevLett.76.2832.

[3] Shunlong Luo. "Using measurement-induced disturbance to characterize correlations as classical or quantum". In: *Phys. Rev. A* 77 (2 Feb. 2008), p. 022301. DOI: 10.1103/PhysRevA.77.022301. URL: https://link.aps.org/doi/10.1103/PhysRevA.77.022301.

[4] William K Wootters and Wojciech H Zurek. "A single quantum cannot be cloned". In: *Nature* 299.5886 (1982), p. 802. DOI: 10.1038/299802a0. URL: https://www.nature.com/articles/299802a0.

[5] Andrew G. White et al. "Nonmaximally Entangled States: Production, Characterization, and Utilization". In: *Phys. Rev. Lett.* 83 (16 Oct. 1999), pp. 3103–3107. DOI: 10.1103/PhysRevLett.83.3103. URL: https://link.aps.org/doi/10.1103/PhysRevLett.83.3103.

[6] Nicolas Brunner et al. "Testing the Dimension of Hilbert Spaces". In: *Phys. Rev. Lett.* 100 (21 May 2008), p. 210503. DOI: 10.1103/PhysRevLett.100.210503. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.210503.

[7] Martin Hendrych et al. "Experimental estimation of the dimension of classical and quantum systems". In: *Nature Physics* 8.8 (2012), pp. 588–591. ISSN: 1745-2481. DOI: 10.1038/nphys2334. URL: https://doi.org/10.1038/nphys2334.

[8] A. G. White et al. "Exploring Hilbert space: Accurate characterization of quantum information". In: *Phys. Rev. A* 65 (1 Dec. 2001), p. 012301. DOI: 10.1103/PhysRevA.65.012301. URL: https://link.aps.org/doi/10.1103/PhysRevA.65.012301.

[9] K. J. Resch, P. Walther, and A. Zeilinger. "Full Characterization of a Three-Photon Greenberger-Horne-Zeilinger State Using Quantum State Tomography". In: *Phys. Rev. Lett.* 94 (7 Feb. 2005), p. 070402. DOI: 10.1103/PhysRevLett.94.070402. URL: https://link.aps.org/doi/10.1103/PhysRevLett.94.070402.

[10] N. Kiesel et al. "Experimental Observation of Four-Photon Entangled Dicke State with High Fidelity". In: *Phys. Rev. Lett.* 98 (6 Feb. 2007), p. 063604. DOI: 10.1103/PhysRevLett.98.063604. URL: https://link.aps.org/doi/10.1103/PhysRevLett.98.063604.

[11] J. L. O'Brien et al. "Demonstration of an all-optical quantum controlled-NOT gate". In: *Nature* 426.6964 (2003), pp. 264–267. ISSN: 1476-4687. DOI: 10.1038/nature02054. URL: https://doi.org/10.1038/nature02054.

[12] P. Walther et al. "Experimental one-way quantum computing". In: *Nature* 434.7030 (2005), pp. 169–176. ISSN: 1476-4687. DOI: 10.1038/nature03347. URL: https://doi.org/10.1038/nature03347.

[13] R. B. A. Adamson and A. M. Steinberg. "Improving Quantum State Estimation with Mutually Unbiased Bases". In: *Phys. Rev. Lett.* 105 (3 July 2010), p. 030406. DOI: 10.1103/PhysRevLett.105.030406. URL: https://link.aps.org/doi/10.1103/PhysRevLett.105.030406.

[14] K. Vogel and H. Risken. "Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase". In: *Phys. Rev. A* 40 (5 Sept. 1989), pp. 2847–2849. DOI: 10.1103/PhysRevA.40.2847. URL: https://link.aps.org/doi/10.1103/PhysRevA.40.2847.

[15] D. T. Smithey et al. "Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum". In: *Phys. Rev. Lett.* 70 (9 Mar. 1993), pp. 1244–1247. DOI: 10.1103/PhysRevLett.70.1244. URL: https://link.aps.org/doi/10.1103/PhysRevLett.70.1244.

[16] Daniel F. V. James et al. "Measurement of qubits". In: *Phys. Rev. A* 64 (5 Oct. 2001), p. 052312. DOI: 10.1103/PhysRevA.64.052312. URL: https://link.aps.org/doi/10.1103/PhysRevA.64.052312.

[17] M. W. Mitchell et al. "Diagnosis, Prescription, and Prognosis of a Bell-State Filter by Quantum Process Tomography". In: *Phys. Rev. Lett.* 91 (12 Sept. 2003), p. 120402. DOI: 10.1103/PhysRevLett.91.120402. URL: https://link.aps.org/doi/10.1103/PhysRevLett.91.120402.

[18] H. Häffner et al. "Scalable multiparticle entanglement of trapped ions". In: *Nature* 438.7068 (2005), pp. 643–646. ISSN: 1476-4687. DOI: 10.1038/nature04279. URL: https://doi.org/10.1038/nature04279.

[19] K. S. Choi et al. "Mapping photonic entanglement into and out of a quantum memory". In: *Nature* 452.7183 (2008), pp. 67–71. ISSN: 1476-4687. DOI: 10.1038/nature06670. URL: https://doi.org/10.1038/nature06670.

[20] S. H. Myrskog et al. "Quantum process tomography on vibrational states of atoms in an optical lattice". In: *Phys. Rev. A* 72 (1 July 2005), p. 013615. DOI: 10.1103/PhysRevA.72.013615. URL: https://link.aps.org/doi/10.1103/PhysRevA.72.013615.

[21] Hao Yuan, Zheng-Wei Zhou, and Guang-Can Guo. "Quantum state tomography via mutually unbiased measurements in driven cavity QED systems". In: *New Journal of Physics* 18.4 (Apr. 2016), p. 043013. DOI: 10.1088/1367-2630/18/4/043013. URL: https://doi.org/10.1088%2F1367-2630%2F18%2F4%2F043013.

[22] R. M. Stevenson et al. "A semiconductor source of triggered entangled photon pairs". In: *Nature* 439.7073 (2006), pp. 179–182. ISSN: 1476-4687. DOI: 10.1038/nature04446. URL: https://doi.org/10.1038/nature04446.

[23] Matthias Steffen et al. "Measurement of the Entanglement of Two Superconducting Qubits via State Tomography". In: *Science* 313.5792 (2006), pp. 1423–1425. ISSN: 0036-8075. DOI: 10.1126/science.1130886. eprint: https://science.sciencemag.org/content/313/5792/1423.full.pdf. URL: https://science.sciencemag.org/content/313/5792/1423.

[24] Alexei Ourjoumtsev, Rosa Tualle-Brouri, and Philippe Grangier. "Quantum Homodyne Tomography of a Two-Photon Fock State". In: *Phys. Rev. Lett.* 96 (21 June 2006), p. 213601. DOI: 10.1103/PhysRevLett.96.213601. URL: https://link.aps.org/doi/10.1103/PhysRevLett.96.213601.

[25] J. S. Neergaard-Nielsen et al. "Generation of a Superposition of Odd Photon Number States for Quantum Information Networks". In: *Phys. Rev. Lett.* 97 (8 Aug. 2006), p. 083604. DOI: 10.1103/PhysRevLett.97.083604. URL: https://link.aps.org/doi/10.1103/PhysRevLett.97.083604.

[26] U. Leonhardt, P.L. Knight, and A. Miller. *Measuring the Quantum State of Light*. Cambridge Studies in Modern Optics. Cambridge University Press, 1997. ISBN: 9780521497305. URL: https://books.google.de/books?id=wmsJy1A%5C_cyIC.

[27] Bo Qi et al. "Quantum State Tomography via Linear Regression Estimation". In: *Scientific Reports* 3.1 (2013), p. 3496. ISSN: 2045-2322. DOI: 10.1038/srep03496. URL: https://doi.org/10.1038/srep03496.

[28] Z. Hradil. "Quantum-state estimation". In: *Phys. Rev. A* 55 (3 Mar. 1997), R1561–R1564. DOI: 10.1103/PhysRevA.55.R1561. URL: https://link.aps.org/doi/10.1103/PhysRevA.55.R1561.

[29] Jaroslav Rehacek et al. "Diluted maximum-likelihood algorithm for quantum tomography". In: *Phys. Rev. A* 75 (4 Apr. 2007), p. 042108. DOI: 10.1103/PhysRevA.75.042108. URL: https://link.aps.org/doi/10.1103/PhysRevA.75.042108.

[30] David Gross et al. "Quantum State Tomography via Compressed Sensing". In: *Phys. Rev. Lett.* 105 (15 Oct. 2010), p. 150401. DOI: `10.1103/PhysRevLett.105.150401`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.105.150401`.

[31] Steven T Flammia et al. "Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators". In: *New Journal of Physics* 14.9 (Sept. 2012), p. 095022. DOI: `10.1088/1367-2630/14/9/095022`. URL: `https://doi.org/10.1088%2F1367-2630%2F14%2F9%2F095022`.

[32] Christopher Granade, Joshua Combes, and D G Cory. "Practical Bayesian tomography". In: *New Journal of Physics* 18.3 (Mar. 2016), p. 033024. DOI: `10.1088/1367-2630/18/3/033024`. URL: `https://doi.org/10.1088%2F1367-2630%2F18%2F3%2F033024`.

[33] Yong Siah Teo. "Numerical Estimation Schemes for Quantum Tomography". PhD thesis. National University of Singapore, 2012. URL: `https://arxiv.org/abs/1302.3399`.

[34] Carlton M. Caves, Christopher A. Fuchs, and Rüdiger Schack. "Unknown quantum states: The quantum de Finetti representation". In: *Journal of Mathematical Physics* 43.9 (Sept. 2002), pp. 4537–4559. ISSN: 1089-7658. DOI: `10.1063/1.1494475`. URL: `http://dx.doi.org/10.1063/1.1494475`.

[35] David G. Cory, Amr F. Fahmy, and Timothy F. Havel. "Ensemble quantum computing by NMR spectroscopy". In: *Proceedings of the National Academy of Sciences* 94.5 (1997), pp. 1634–1639. ISSN: 0027-8424. DOI: `10.1073/pnas.94.5.1634`. eprint: `https://www.pnas.org/content/94/5/1634.full.pdf`. URL: `https://www.pnas.org/content/94/5/1634`.

[36] Jonathan A. Jones, Michele Mosca, and Rasmus H. Hansen. "Implementation of a quantum search algorithm on a quantum computer". In: *Nature* 393.6683 (1998), pp. 344–346. ISSN: 1476-4687. DOI: `10.1038/30687`. URL: `https://doi.org/10.1038/30687`.

[37] C. Monroe. "Quantum information processing with atoms and photons". In: *Nature* 416.6877 (2002), pp. 238–246. ISSN: 1476-4687. DOI: `10.1038/416238a`. URL: `https://doi.org/10.1038/416238a`.

[38] Ferdinand Schmidt-Kaler et al. "Realization of the Cirac-Zoller controlled-NOT quantum gate". In: *Nature* 422.6930 (2003), pp. 408–411. ISSN: 1476-4687. DOI: `10.1038/nature01494`. URL: `https://doi.org/10.1038/nature01494`.

[39] Max Born. "Zur Quantenmechanik der Stoßvorgänge". In: *Zeitschrift für Physik* 37.12 (Dec. 1926), pp. 863–867. ISSN: 0044-3328. DOI: `10.1007/BF01397477`. URL: `https://doi.org/10.1007/BF01397477`.

[40] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010. DOI: `10.1017/CBO9780511976667`. URL: `https://www.cambridge.org/core/books/quantum-computation-and-quantum-information/01E10196D0A682A6AEFFEA52D53BE9AE#`.

[41] Paulo E. M. F. Mendon ça et al. "Alternative fidelity measure between quantum states". In: *Phys. Rev. A* 78 (5 Nov. 2008), p. 052330. DOI: `10.1103/PhysRevA.78.052330`. URL: `https://link.aps.org/doi/10.1103/PhysRevA.78.052330`.

[42] Stephen Barnett. *Quantum information*. Vol. 16. Oxford University Press, 2009.

[43] P. R. Halmos. "What Does the Spectral Theorem Say?" In: *The American Mathematical Monthly* 70.3 (1963), pp. 241–247. DOI: `10.1080/00029890.1963.11990075`. eprint: `https://doi.org/10.1080/00029890.1963.11990075`. URL: `https://doi.org/10.1080/00029890.1963.11990075`.

[44] Christoph Dankert. "Efficient simulation of random quantum states and operators". In: *arXiv preprint quant-ph/0512217* (2005).

[45] Christian Schwemmer. "Efficient Tomography and Entanglement Detection of Multiphoton States". PhD thesis. Ludwig-Maximilians-Universität, München, 2015. URL: `https://xqp.physik.uni-muenchen.de/publications/files/theses_phd/phd_schwemmer.pdf`.

[46] Alexander Niggebaum. "Quantum State Tomography of the 6 qubit photonic symmetric Dicke State". MA thesis. Munich: Ludwig Maximilians Universität, 2011. URL: `https://xqp.physik.uni-muenchen.de/publications/theses_master/master_niggebaum.html`.

[47] H. P. Robertson. "The Uncertainty Principle". In: *Phys. Rev.* 34 (1 July 1929), pp. 163–164. DOI: `10.1103/PhysRev.34.163`. URL: `https://link.aps.org/doi/10.1103/PhysRev.34.163`.

[48] Lars M. Johansen. "Quantum theory of successive projective measurements". In: *Phys. Rev. A* 76 (1 July 2007), p. 012119. DOI: `10.1103/PhysRevA.76.012119`. URL: `https://link.aps.org/doi/10.1103/PhysRevA.76.012119`.

[49] Lane P. Hughston, Richard Jozsa, and William K. Wootters. "A complete classification of quantum ensembles having a given density matrix". In: *Physics Letters A* 183.1 (1993), pp. 14–18. ISSN: 0375-9601. DOI: `https://doi.org/10.1016/0375-9601(93)90880-9`. URL: `http://www.sciencedirect.com/science/article/pii/0375960193908809`.

[50] Ermes Toninelli et al. "Concepts in quantum state tomography and classical implementation with intense light: a tutorial". In: *Adv. Opt. Photon.* 11.1 (Mar. 2019), pp. 67–134. DOI: `10.1364/AOP.11.000067`. URL: `http://aop.osa.org/abstract.cfm?URI=aop-11-1-67`.

[51] U. Fano. "Description of States in Quantum Mechanics by Density Matrix and Operator Techniques". In: *Rev. Mod. Phys.* 29 (1 Jan. 1957), pp. 74–93. DOI: `10.1103/RevModPhys.29.74`. URL: `https://link.aps.org/doi/10.1103/RevModPhys.29.74`.

[52] Roman Schmied. "Quantum state tomography of a single qubit: comparison of methods". In: *Journal of Modern Optics* 63.18 (2016), pp. 1744–1758. DOI: `10.1080/09500340.2016.1142018`. eprint: `https://doi.org/10.1080/09500340.2016.1142018`. URL: `https://doi.org/10.1080/09500340.2016.1142018`.

[53] Joseph B. Altepeter, Daniel F.V. James, and Paul G. Kwiat. "4 Qubit Quantum State Tomography". In: *Quantum State Estimation*. Ed. by Matteo Paris and Jaroslav Řeháček. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 113–145. ISBN: 978-3-540-44481-7. DOI: `10.1007/978-3-540-44481-7_4`. URL: `https://doi.org/10.1007/978-3-540-44481-7_4`.

[54] D. H. Mahler et al. "Adaptive Quantum State Tomography Improves Accuracy Quadratically". In: *Phys. Rev. Lett.* 111 (18 Oct. 2013), p. 183601. DOI: `10.1103/PhysRevLett.111.183601`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.111.183601`.

[55] Robin Blume-Kohout. "Optimal, reliable estimation of quantum states". In: *New Journal of Physics* 12.4 (Apr. 2010), p. 043034. DOI: `10.1088/1367-2630/12/4/043034`. URL: `https://doi.org/10.1088%2F1367-2630%2F12%2F4%2F043034`.

[56] André-Louis Cholesky. "Sur la résolution numérique des systèmes d'équations linéaires". In: *Bulletin de la Sabix. Société des amis de la Bibliothèque et de l'Histoire de l'École polytechnique* 39 (2005), pp. 81–95.

[57] Augustin Cauchy. "Méthode générale pour la résolution des systemes d'équations simultanées". In: *Comp. Rend. Sci. Paris* 25.1847 (1847), pp. 536–538.

[58] Magnus Rudolph Hestenes and Eduard Stiefel. *Methods of conjugate gradients for solving linear systems*. Vol. 49. 1. NBS Washington, DC, 1952.

[59] Lukas Knips. "Reconstruction and analysis ofmulti-photon entangled quantum states". MA thesis. Munich: Ludwig Maximilians Universität, 2013. URL: `https://xqp.physik.uni-muenchen.de/publications/theses_master/master_knips.html`.

[60] Eliot Bolduc et al. "Projected gradient descent algorithms for quantum state tomography". In: *npj Quantum Information* 3.1 (2017), p. 44. ISSN: 2056-6387. DOI: `10.1038/s41534-017-0043-1`. URL: `https://doi.org/10.1038/s41534-017-0043-1`.

[61] Jiangwei Shang, Zhengyun Zhang, and Hui Khoon Ng. "Superfast maximum-likelihood reconstruction for quantum tomography". In: *Phys. Rev. A* 95 (6 June 2017), p. 062336. DOI: `10.1103/PhysRevA.95.062336`. URL: `https://link.aps.org/doi/10.1103/PhysRevA.95.062336`.

[62] Christian Schwemmer et al. "Systematic Errors in Current Quantum State Tomography Tools". In: *Phys. Rev. Lett.* 114 (8 Feb. 2015), p. 080403. DOI: 10.1103/PhysRevLett.114.080403. URL: https://link.aps.org/doi/10.1103/PhysRevLett.114.080403.

[63] Robin Blume-Kohout. "Hedged Maximum Likelihood Quantum State Estimation". In: *Phys. Rev. Lett.* 105 (20 Nov. 2010), p. 200504. DOI: 10.1103/PhysRevLett.105.200504. URL: https://link.aps.org/doi/10.1103/PhysRevLett.105.200504.

[64] Christopher Ferrie and Robin Blume-Kohout. "Minimax Quantum Tomography: Estimators and Relative Entropy Bounds". In: *Phys. Rev. Lett.* 116 (9 Mar. 2016), p. 090407. DOI: 10.1103/PhysRevLett.116.090407. URL: https://link.aps.org/doi/10.1103/PhysRevLett.116.090407.

[65] Christopher Ferrie and Robin Blume-Kohout. "Maximum likelihood quantum state tomography is inadmissible". In: *arXiv preprint arXiv:1808.01072* (2018).

[66] B. Efron and R. Tibshirani. "Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy". In: *Statist. Sci.* 1.1 (Feb. 1986), pp. 54–75. DOI: 10.1214/ss/1177013815. URL: https://doi.org/10.1214/ss/1177013815.

[67] Christopher Ferrie. "Self-Guided Quantum Tomography". In: *Phys. Rev. Lett.* 113 (19 Nov. 2014), p. 190404. DOI: 10.1103/PhysRevLett.113.190404. URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.190404.

[68] Steven T. Flammia and Yi-Kai Liu. "Direct Fidelity Estimation from Few Pauli Measurements". In: *Phys. Rev. Lett.* 106 (23 June 2011), p. 230501. DOI: 10.1103/PhysRevLett.106.230501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.106.230501.

[69] J. C. Spall. "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation". In: *IEEE Transactions on Automatic Control* 37.3 (Mar. 1992), pp. 332–341. DOI: 10.1109/9.119632.

[70] G. Tóth et al. "Permutationally Invariant Quantum Tomography". In: *Phys. Rev. Lett.* 105 (25 Dec. 2010), p. 250403. DOI: 10.1103/PhysRevLett.105.250403. URL: https://link.aps.org/doi/10.1103/PhysRevLett.105.250403.

[71] Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger. "Going Beyond Bell's Theorem". In: *Bell's Theorem, Quantum Theory and Conceptions of the Universe.* Ed. by Menas Kafatos. Dordrecht: Springer Netherlands, 1989, pp. 69–72. ISBN: 978-94-017-0849-4. DOI: 10.1007/978-94-017-0849-4_10. URL: https://doi.org/10.1007/978-94-017-0849-4_10.

[72] W. Dür, G. Vidal, and J. I. Cirac. "Three qubits can be entangled in two inequivalent ways". In: *Phys. Rev. A* 62 (6 Nov. 2000), p. 062314. DOI: `10.1103/PhysRevA.62.062314`. URL: `https://link.aps.org/doi/10.1103/PhysRevA.62.062314`.

[73] R. H. Dicke. "Coherence in Spontaneous Radiation Processes". In: *Phys. Rev.* 93 (1 Jan. 1954), pp. 99–110. DOI: `10.1103/PhysRev.93.99`. URL: `https://link.aps.org/doi/10.1103/PhysRev.93.99`.

[74] David Pérez-García et al. "Matrix product state representations". In: *Quantum Information & Computation* 7.5 (2007), pp. 401–430. URL: `http://www.rintonpress.com/xqic7/qic-7-56/401-430.pdf`.

[75] Jiangwei Shang et al. "Monte Carlo sampling from the quantum state space. I". In: *New Journal of Physics* 17.4 (Apr. 2015), p. 043017. DOI: `10.1088/1367-2630/17/4/043017`. URL: `https://doi.org/10.1088%2F1367-2630%2F17%2F4%2F043017`.

[76] Joseph M. Renes et al. "Symmetric informationally complete quantum measurements". In: *Journal of Mathematical Physics* 45.6 (2004), pp. 2171–2180. DOI: `10.1063/1.1737053`. eprint: `https://doi.org/10.1063/1.1737053`. URL: `https://doi.org/10.1063/1.1737053`.

[77] N. Bent et al. "Experimental Realization of Quantum Tomography of Photonic Qudits via Symmetric Informationally Complete Positive Operator-Valued Measures". In: *Phys. Rev. X* 5 (4 Oct. 2015), p. 041006. DOI: `10.1103/PhysRevX.5.041006`. URL: `https://link.aps.org/doi/10.1103/PhysRevX.5.041006`.

[78] Hui Khoon Ng and Berthold-Georg Englert. "A simple minimax estimator for quantum states". In: *International Journal of Quantum Information* 10.04 (2012), p. 1250038. DOI: `10.1142/S0219749912500384`. eprint: `https://doi.org/10.1142/S0219749912500384`. URL: `https://doi.org/10.1142/S0219749912500384`.

[79] Robert E. Kass and Larry Wasserman. "The Selection of Prior Distributions by Formal Rules". In: *Journal of the American Statistical Association* 91.435 (1996), pp. 1343–1370. DOI: `10.1080/01621459.1996.10477003`. eprint: `https://www.tandfonline.com/doi/pdf/10.1080/01621459.1996.10477003`. URL: `https://www.tandfonline.com/doi/abs/10.1080/01621459.1996.10477003`.

[80] Harold Jeffreys. "An invariant form for the prior probability in estimation problems". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 186.1007 (1946), pp. 453–461. DOI: `10.1098/rspa.1946.0056`. URL: `https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1946.0056`.

[81] Jiangwei Shang et al. "Optimal error regions for quantum state estimation". In: *New Journal of Physics* 15.12 (Dec. 2013), p. 123026. DOI: `10.1088/1367-2630/15/12/123026`. URL: `https://doi.org/10.1088%2F1367-2630%2F15%2F12%2F123026`.

[82] Alfred Haar. "Der Massbegriff in der Theorie der Kontinuierlichen Gruppen". In: *Annals of Mathematics* 34.1 (1933), pp. 147–169. ISSN: 0003486X. URL: http://www.jstor.org/stable/1968346.

[83] Jonas Maziero. "Generating Pseudo-Random Discrete Probability Distributions". In: *Brazilian Journal of Physics* 45.4 (Aug. 2015), pp. 377–382. ISSN: 1678-4448. DOI: 10.1007/s13538-015-0337-8. URL: https://doi.org/10.1007/s13538-015-0337-8.

[84] Jonas Maziero. "Random Sampling of Quantum States: a Survey of Methods". In: *Brazilian Journal of Physics* 45.6 (Dec. 2015), pp. 575–583. ISSN: 1678-4448. DOI: 10.1007/s13538-015-0367-2. URL: https://doi.org/10.1007/s13538-015-0367-2.

[85] E. Brüning et al. "Parametrizations of density matrices". In: *Journal of Modern Optics* 59.1 (2012), pp. 1–20. DOI: 10.1080/09500340.2011.632097. eprint: https://doi.org/10.1080/09500340.2011.632097. URL: https://doi.org/10.1080/09500340.2011.632097.

[86] K Zyczkowski and M Kus. "Random unitary matrices". In: *Journal of Physics A: Mathematical and General* 27.12 (June 1994), pp. 4235–4245. DOI: 10.1088/0305-4470/27/12/028. URL: https://doi.org/10.1088%2F0305-4470%2F27%2F12%2F028.

[87] Robert McGill, John W. Tukey, and Wayne A. Larsen. "Variations of Box Plots". In: *The American Statistician* 32.1 (1978), pp. 12–16. ISSN: 00031305. URL: http://www.jstor.org/stable/2683468.

[88] Paul F. Velleman. *Applications, Basics, and Computing of Exploratory Data Analysis*. Wadsworth Pub Co, May 1984. ISBN: 087150409X. URL: https://www.xarg.org/ref/a/087150409X/.

[89] P. Kurzyński et al. "Correlation Complementarity Yields Bell Monogamy Relations". In: *Phys. Rev. Lett.* 106 (18 May 2011), p. 180402. DOI: 10.1103/PhysRevLett.106.180402. URL: https://link.aps.org/doi/10.1103/PhysRevLett.106.180402.

[90] William K. Wootters. "Quantum mechanics without probability amplitudes". In: *Foundations of Physics* 16.4 (Apr. 1986), pp. 391–405. ISSN: 1572-9516. DOI: 10.1007/BF01882696. URL: https://doi.org/10.1007/BF01882696.

[91] I D Ivonovic. "Geometrical description of quantal state determination". In: *Journal of Physics A: Mathematical and General* 14.12 (Dec. 1981), pp. 3241–3245. DOI: 10.1088/0305-4470/14/12/019. URL: https://doi.org/10.1088%2F0305-4470%2F14%2F12%2F019.

[92] Jay Lawrence, Caslav Brukner, and Anton Zeilinger. "Mutually unbiased binary observable sets on N qubits". In: *Phys. Rev. A* 65 (3 Feb. 2002), p. 032320. DOI: 10.1103/PhysRevA.65.032320. URL: https://link.aps.org/doi/10.1103/PhysRevA.65.032320.

[93] Prabha Mandayam et al. "Unextendible mutually unbiased bases from Pauli classes". In: *arXiv preprint arXiv:1302.3709* (2013).

[94]   A Garcia, J L Romero, and A B Klimov. "Generation of bases with def-
        inite factorization for ann-qubit system and mutually unbiased sets con-
        struction". In: *Journal of Physics A: Mathematical and Theoretical* 43.38
        (Aug. 2010), p. 385301. DOI: `10.1088/1751-8113/43/38/385301`. URL:
        `https://doi.org/10.1088%2F1751-8113%2F43%2F38%2F385301`.

[95]   J. L. Romero et al. "Structure of the sets of mutually unbiased bases for
        N qubits". In: *Phys. Rev. A* 72 (6 Dec. 2005), p. 062310. DOI: `10.1103/
        PhysRevA.72.062310`. URL: `https://link.aps.org/doi/10.1103/
        PhysRevA.72.062310`.

[96]   Lukas Knips et al. "Multipartite Entanglement Detection with Minimal
        Effort". In: *Phys. Rev. Lett.* 117 (21 Nov. 2016), p. 210504. DOI: `10.1103/
        PhysRevLett.117.210504`. URL: `https://link.aps.org/doi/10.1103/
        PhysRevLett.117.210504`.

[97]   Harald Weinfurter. "Experimental Bell-State Analysis". In: *Europhysics
        Letters (EPL)* 25.8 (Mar. 1994), pp. 559–564. DOI: `10.1209/0295-5075/
        25/8/001`. URL: `https://doi.org/10.1209%2F0295-5075%2F25%2F8%
        2F001`.

[98]   Tzu-Chieh Wei, Julio T. Barreiro, and Paul G. Kwiat. "Hyperentangled
        Bell-state analysis". In: *Phys. Rev. A* 75 (6 June 2007), p. 060305. DOI:
        `10.1103/PhysRevA.75.060305`. URL: `https://link.aps.org/doi/10.
        1103/PhysRevA.75.060305`.

[99]   G. Lima et al. "Experimental quantum tomography of photonic qudits via
        mutually unbiased basis". In: *Opt. Express* 19.4 (Feb. 2011), pp. 3542–
        3552. DOI: `10.1364/OE.19.003542`. URL: `http://www.opticsexpress.
        org/abstract.cfm?URI=oe-19-4-3542`.