Contents

Introduction

The new field of quantum information has opened the way to a new kind of astonishingly efficient information processing achieved by physical transformations. This new kind of processing will be performed by a radically new generation of quantum devices, and this will make the design of characterization tools for such devices of paramount importance, besides being already of foundational interest by themselves, for the obvious possibility of experimental determination of the dynamics of a quantum system.

Quantum devices can perform either deterministic or probabilistic trasformations of a quantum state. The transformations of the deterministic class are generally referred to as "processes" or "channels", and describe the evolution of closed systems or of open systems undergoing an irreversible dynamics, such as due to an in interaction with a bath. The class of probabilistic transformations, on the other hand, typically describes the so called "state reduction" occurring in a quantum measurement. Both types of transformations can be described in the language of quantum operations (QO) [1, 2], and, within this common mathematical structure, both deterministic and non-deterministic transformations can be characterized by the same means.

At the root of the characterization problem, there is the need of finding a way to imprint the description of the QO of the device on a suitable input state which is processed by the device, and is then characterized at its output by some quantum tomographic means. Linearity of QO's is the first key ingredient for solving this "quantum black box" problem. In Refs. [3, 4] it was shown that for a "complete" set of input states, i.e. for a set of generators of the space of states, the "transfer matrix" of the device remains encoded in the input/output correlations in the same way as for any classical linear system, the only difference being that in the quantum case one needs many copies of the outputs to perform their quantum tomography. Quantum process tomography

was achieved by this methods in liquid nuclear magnetic resonance systems [5, 6, 7], and for processes on qubits encoded in the polarization of a radiation mode [8, 9]. Unfortunately, this method needs the preparation of an orthogonal set of input states along with some relative superpositions, and such sets of states are very seldom available in the lab: for example, they are not achievable in quantum optics.

Quantum Mechanics, however, offers a unique opportunity to achieve our goal by using a composite system. In fact, in Refs. [10] we showed that the action of a quantum process on one system of an entangled pair produces a joint output state containing a complete description of the process itself, a result also known as the Jamiolkowsky isomorphism [11, 12]. In simple words, a fixed maximally entangled input state supports the imprinting of any QO, as if it was effectively running all possible input states in parallel, and in this way the determination of the process is achieved by simply performing the tomography of the joint state at the output, with the device acting on one of the two entangled systems only. Experiments of process tomography using entangled input probes have been recently implemented [13, 14, 15] for optical qubits, and proposed for optical "continuous variables" systems using homodyne tomography [10].

In our work of Ref. [16], the two methods—"many inputs" versus "single entangled input"—have been bridged together in a complete classification of all states (and/or all set of states) that support a complete imprinting of a generic QO, thereafter named "faithful states". There, the existence of separable faithful states has been established, thus clarifying that for the "quantum black box" problem the only thing that matters is the use of composite systems (i. e. with the tensor product rule), more than entanglement itself. Among such faithful separable states there are also the Werner states used in the process tomography experiment of Ref. [14]. In Ref. [16], a measure of the "faithfulness" of the state has been also given, which measure in some way the precision of the tomographic characterization, showing that maximally entangled states offer the best performance.

Once the information on the process is encoded on the quantum state, all known techniques of state-tomography and state-discrimination can be applied. Such techniques will allow in the future a precise characterization of any kind of quantum device, from an optical fiber for a quantum communication channel, to an NMR qubit gate, from a parametric amplifier to a

photon-counting detector.

Structure of the Thesis

The present Thesis is aimed at a presentation of the theoretical basis for either a mathematical or an experimental characterization of quantum devices.

In Chapter 1 we introduce the formalism of quantum operations (QO) and positive operator-valued measures (POVM) for describing the state transformations operated by a quantum device and the statistics of the outcomes of a quantum measurement, respectively. The properties of these two mathematical objects are derived as necessary consequences of the definition and interpretation of quantum state, and the definition of the states of composite systems. These properties are the starting point for constructing powerful representations of QO's that will be used first for illustrating the relation between QO's and POVM's with customary unitary evolutions and projective measurements, and then to analyze first the problem of the mathematical characterization of a extremal and covariant devices, and then the problem of the experimental characterization of quantum device.

In Chapter 2 the convex structure of the set of quantum devices and of the set of POVM's is analyzed and its physical meaning is explained, showing how the extremal points of these sets can be considered – in some sense – as affected by intrinsically quantum randomness. A mathematical characterization of the extremal points is found either for the set of POVM's or for the set of quantum devices, along with an algorithm to decompose a non-extremal point into the convex linear combination of extremals [17]. In the second section of the chapter we give a complete mathematical characterization of covariant QO's, which correspond to physical transformations that propagate, from the input to the output, the action of the elements of a unitary representation of a group. The characterization of extremal covariant QO's is given, and finally quantum cloning is interpreted as a quantum channel being covariant with respect to permutations, and the result of the previous section is employed to parametrize cloning transformations, and then to calculate some novel examples of optimal covariant cloning transformations [18].

Chapter 3 is entirely devoted to the classification of *faithful states* — i.e. the input states that can be used for the characterization (intended as measurement) of a quantum device — and to address the problem of quantifying their degree of faithfulness [16].

In Chapter 4 we propose a quantum optical setup for performing a device characterization by homodyne tomography using as input faithful state an entangled state from parametric down-conversion of vacuum. Numerical simulations are reported of the experimental results obtainable with the current technology for the homodyne tomography of an amplitude displacing device [10], of an On/Off photo-detector, and of a photon-counting detector [19], using either the pattern function averaging or the maximum-likelihood strategies for the reconstruction.

In Chapter 5 we analyze the role of entanglement in improving either the precision or the stability of quantum measurements resorting to the discrimination of unitary transformations [20]. For the case of the estimation of unitary transformations belonging to an irreducible representation of a group, to use an entangled probe produces an enhancement of several figures of merit, whereas in other situations, entanglement makes the measurement more robust with respect to noise or miscalibration of the measuring device. Finally, we show how the use of multiparty entanglement would make it possible a perfect discrimination of among a finite number of unitary transformations, when several yet limeted uses of the questioned transformation are at our disposal.

Chapter 1

Quantum operations and quantum measurements

Quantum operations (QO), introduced for the first time in Refs. [1, 2], describe all possible transformations — either deterministic or probabilistic — of the state ρ of a quantum system. Mathematically, QO's are completely positive linear maps from the set of (trace-class) operators on H to itself, and are trace preserving when deterministic and trace-decreasing when probabilistic, with the probability of occurrence given by the output trace.

In this chapter we'll show how these properties for QO's can be traced back to the indistinguishability of different preparations of the same ensemble of systems and to the tensor product structure of composite systems. After these observations, by means of a particularly useful notation for bipartite states we have adopted since Ref. [21], we will introduce two representations of CP maps in terms of operators on $\mathsf{H}^{\otimes 2}$, one evidencing the properties of the map descending from linearity, the other the ones bound to its complete positivity. These representations will provide the easiest framework to proof the most relevant results concerning quantum operations. Then, we will review the concept of POVM for representing the probability distribution of the outcomes of a quantum measurement, and its connection with quantum operations.

1.1 Properties of quantum operations

Let's consider a system in the state ρ, and suppose it enters a device in which a physical transformation described by the map $\mathcal{T}$

$$\rho \mapsto \mathcal{T}(\rho) \tag{1.1}$$

occurs with a probability $p(\rho)$, in such a way that we know whether the transformation has occurred or not. This situation describes a general quantum measurement, in which an "occurrence flag" for the transformation represents the "outcome", and the dependence of $p(\rho)$ on ρ will give us some information on the state of the system. Since $\mathcal{T}(\rho)$ is a quantum state, then the map $\mathcal{T}$ satisfies for all ρ

$$\mathcal{T}(\rho) \in \mathsf{T}(\mathsf{H}) \;, \quad \mathcal{T}(\rho) \geq 0 \;, \quad \text{and} \quad \mathrm{tr}\,\mathcal{T}(\rho) = 1 \;, \tag{1.2}$$

where $\mathsf{T}(\mathsf{H})$ denotes trace-class operators on H. Consider now an ensemble of systems prepared as $\{(p_i, \rho_i)\}$. After the action of the device, the portion of systems having undergone the transformation is $\sum_i p_i\, p(\rho_i)$, and the selection of these systems yields an ensemble described by the state

$$\rho' = \frac{\sum_i p_i\, p(\rho_i)\, \mathcal{T}(\rho_i)}{\sum_i p_i\, p(\rho_i)} \;.$$

On the other hand, the initial ensemble is also represented by the state $\rho = \sum_i p_i\, \rho_i$, so that the final post-selected ensemble will correspond to the state $\mathcal{T}(\rho)$, with a fraction of transformed systems equal to $p(\rho)$. The two descriptions must be consistent, because of the hypothesized indistinguishability of two different preparations of the same ensemble, thus the fraction of transformed systems and the final state must be the same in both cases, namely

$$p(\rho) \;=\; \sum_i p_i\, p(\rho_i) \;, \tag{1.3}$$

$$\mathcal{T}(\rho) \;=\; \frac{\sum_i p_i\, p(\rho_i)\, \mathcal{T}(\rho_i)}{\sum_i p_i\, p(\rho_i)} \;. \tag{1.4}$$

The first equation implies that $p(\rho)$ is a *linear* function of ρ, and, as we shall see later, this holds for any probability distribution of the outcomes of a quantum measurement, and it is unrelated to the details of the state transformation corresponding to each outcome: this will allow us to introduce the concept

of POVM, which gives only the probability distribution of the outcomes as a function of the state. In the present context we are actually describing a "yes/no" measurement, i.e. our transformation $\mathcal{T}$ "has" or "has not" occurred.

If we now introduce the map $\mathcal{E}(\rho) \doteq p(\rho)\mathcal{T}(\rho)$, Eq. (1.4) tells us that $\mathcal{E}$ is a *linear* function of ρ. Taking the trace of the above definition of $\mathcal{E}$, and remembering that $\operatorname{tr}\mathcal{T}(\rho) = 1$, we find $p(\rho) = \operatorname{tr}\mathcal{E}(\rho)$, so that the transformation $\mathcal{T}$ and the probability $p(\rho)$ can be written in terms of $\mathcal{E}$ as follows

$$\mathcal{T}(\rho) = \frac{\mathcal{E}(\rho)}{\operatorname{tr}\mathcal{E}(\rho)} , \quad p(\rho) = \operatorname{tr}\mathcal{E}(\rho) . \tag{1.5}$$

From Eq. (1.2), and from the fact that $p(\rho)$ is a probability, one argues that also the following properties must hold for $\mathcal{E}$

$$\begin{aligned} \mathcal{E}(\rho) \geq 0 \qquad & (\textit{positivity}) , \\ \operatorname{tr}\mathcal{E}(\rho) \leq 1 \qquad & (\textit{trace decreasing or preserving}) . \end{aligned} \tag{1.6}$$

The more stringent property of complete positivity for $\mathcal{E}$ follows from the tensor-product structure of composite systems in Quantum Mechanics. In fact, when $\mathcal{T}$ acts only on a single subsystem of a bipartite quantum system, the joint state R of the system transforms according to

$$\rho \mapsto (\mathcal{E} \otimes \mathcal{I})(R) ,$$

and thus not only $\mathcal{E}$ but also its extension $\mathcal{E} \otimes \mathcal{I}$ must be positive, in such a way that the result of the local transformation is still a quantum state. This must hold for all possible extensions of to larger composite systems. This property is called *complete positivity* and it is not equivalent to positivity, as counterexamples exist. For example, the transposition of the state with respect to a given basis $\rho \mapsto \rho^T$ is a linear, positive, trace preserving map, but generally gives a non positive operator when acting on a system of an entangled pair, whence it is not completely positive and it can't be achieved physically. In the following, we will refer to completely positive linear maps simply as *CP maps*, quantum operations corresponding to the class of trace non-increasing CP maps.

Up to now, we have shown that any transformation of the state of a quantum system is described by a *quantum operation* (QO), namely a linear, completely positive, trace non increasing map $\mathcal{E} : \mathsf{T}(\mathsf{H}) \to \mathsf{T}(\mathsf{H})$, with the state transformation given by $\rho \mapsto \mathcal{E}(\rho)/\operatorname{tr}\mathcal{E}(\rho)$, and occurring with probability

$p(\rho) = \mathrm{tr}\,\mathcal{E}(\rho)$. Trace preserving QO's describe deterministic transformations — also called *quantum processes*, or *channels* — namely with $p(\rho) = 1$, whereas trace decreasing QO's describe the transformation of the state of a system undergoing a quantum measurement in concomitance with a given outcome occurring with a probability $p(\rho) = \mathrm{tr}\,\mathcal{E}(\rho) \leq 1$.

1.2 Representing CP maps

In the following we will suppose $\dim(\mathsf{H}) < \infty$, whence we will generically denote trace-class, Hilbert-Schmidt and bounded operators on H simply as $\mathsf{B}(\mathsf{H})$. CP maps are nothing but a special subset of the set of linear maps from $\mathsf{B}(\mathsf{H})$ to $\mathsf{B}(\mathsf{H})$, and therefore they can be represented by means of their "matrix elements"

$$\mathcal{E}_{ij}^{lm} = \langle i|\ \mathcal{E}\big(|l\rangle\langle m|\big)\ |j\rangle\ , \tag{1.7}$$

so that, once defined $\rho_{lm} = \langle l|\rho|m\rangle$, $\mathcal{E}(\rho)$ can be evaluated as

$$\mathcal{E}(\rho) = \sum_{ijlm} \mathcal{E}_{ij}^{lm}\rho_{lm}|i\rangle\langle j|\ . \tag{1.8}$$

However, to have more insight into the structure of linear maps, it is preferable to reorganize the set of matrix elements $\mathcal{E}_{ij}^{lm}$ into an operator on $\mathsf{H}\otimes\mathsf{H}$, aiming that the properties of the map (being CP, trace decreasing, invertible, etc.) have a simple translation into properties of the associated operator.

The following notation [21] will be useful to simplify calculations by avoiding the use of a lot of indices in our equations, thus making them more insightful. Fixed an orthonormal basis $\{|m\rangle\}$ for the Hilbert space H, we identify any vector $|\Psi\rangle\!\rangle \in \mathsf{H}\otimes\mathsf{H}$,

$$|\Psi\rangle\!\rangle = \sum_{m,n} \Psi_{mn}|m\rangle\otimes|n\rangle\ , \tag{1.9}$$

with the operator $\Psi \in \mathsf{B}(\mathsf{H})$ whose matrix elements on the chosen basis are Ψ_{mn}. For example, the vector $|I\rangle\!\rangle$ represents the maximally entangled unnormalized vector $\sum_m |m\rangle\otimes|m\rangle$. It is easy to check that

$$\begin{aligned} A\otimes B|C\rangle\!\rangle &= |ACB^T\rangle\!\rangle\ , & \langle\!\langle A|B\rangle\!\rangle &= \mathrm{tr}[A^\dagger B]\ , \\ \mathrm{tr}_2[|A\rangle\!\rangle\langle\!\langle B|] &= AB^\dagger\ , & \mathrm{tr}_1[|A\rangle\!\rangle\langle\!\langle B|] &= A^T B^*\ , \end{aligned} \tag{1.10}$$

where O^T and O^* denote respectively the transposition and the complex conjugation of the operator O with respect to the chosen basis.

Focusing our attention on the linearity of $\mathcal{E}$, with the notation introduced in Eq. (1.9), we notice that the vector $|\mathcal{E}(\rho)\rangle\!\rangle$ is a linear transformation of $|\rho\rangle\!\rangle$, and thus the relation between the two vectors can be expressed by means of an operator $\check{S}_{\mathcal{E}} \in \mathsf{B}(\mathsf{H} \otimes \mathsf{H})$ such that

$$|\,\mathcal{E}(\rho)\,\rangle\!\rangle = \check{S}_{\mathcal{E}}|\rho\rangle\!\rangle \;. \tag{1.11}$$

The map is faithfully represented by $\check{S}_{\mathcal{E}}$, since the previous relation defines its action on any state. By substituting in the above equation the definition of $\mathcal{E}_{ij}^{lm}$ given in Eq. (1.7) one finds

$$\check{S}_{\mathcal{E}} = \sum_{ijlm} \mathcal{E}_{ij}^{lm}\, |i\rangle\langle l| \otimes |j\rangle\langle m| \;. \tag{1.12}$$

The power of this representation of linear maps resides in the fact that it translates the composition of two maps into the multiplication of their related operators, as one can easily verify from the following identity

$$|\,\mathcal{E}_1 \circ \mathcal{E}_2(\rho)\,\rangle\!\rangle = \check{S}_{\mathcal{E}_1}|\,\mathcal{E}_2(\rho)\,\rangle\!\rangle = \check{S}_{\mathcal{E}_1}\check{S}_{\mathcal{E}_2}|\rho\rangle\!\rangle \;. \tag{1.13}$$

Moreover, such a representation provides a useful tool to evaluate some properties of the map. For example, the image of the map $\mathcal{E}(\mathsf{B}(\mathsf{H}))$ corresponds to the set the operators A such that $|A\rangle\!\rangle \in \mathsf{Rng}\,\check{S}_{\mathcal{E}}$, where "$\mathsf{Rng}$" denotes the range (i.e. the image) of an operator. Analogously, the kernel of $\mathcal{E}$, i.e. the set of operators A such that $\mathcal{E}(A) = 0$, is exactly the set of operators A such that $|A\rangle\!\rangle \in \mathsf{Ker}\,\check{S}_{\mathcal{E}}$. Finally, by definition, $\mathcal{E}$ is invertible iff $\check{S}_{\mathcal{E}}$ is so, and the two inverses are related through the identity

$$|\,\mathcal{E}^{-1}(\rho)\,\rangle\!\rangle = \check{S}_{\mathcal{E}}^{-1}|\rho\rangle\!\rangle \;, \tag{1.14}$$

so that

$$|\,\mathcal{E}^{-1} \circ \mathcal{E}(\rho)\,\rangle\!\rangle = \check{S}_{\mathcal{E}}^{-1}\check{S}_{\mathcal{E}}|\rho\rangle\!\rangle \;= |\rho\rangle\!\rangle \;.$$

Being too much geared around linearity, unfortunately the above representation of maps tells us nothing about complete positivity. For such reason it is convenient to introduce another operator representation of the map $\mathcal{E}$ in terms of the operator $S_{\mathcal{E}} \in \mathsf{B}(\mathsf{H} \otimes \mathsf{H})$ resulting from the action of the extended map $\mathcal{E} \otimes \mathcal{I}$ on the operator $|I\rangle\!\rangle\langle\!\langle I| \in \mathsf{B}(\mathsf{H} \otimes \mathsf{H})$ [12, 11], namely

$$S_{\mathcal{E}} = (\mathcal{E} \otimes \mathcal{I})\,[\,|I\rangle\!\rangle\langle\!\langle I|\,] = \sum_{ijlm} \mathcal{E}_{ij}^{lm}|i\rangle\langle j| \otimes |l\rangle\langle m| \;. \tag{1.15}$$

The inverse relation of identity (1.15) can be easily checked to be

$$\mathcal{E}(\rho) = \mathrm{tr}_2[(I \otimes \rho^T) S_{\mathcal{E}}\ ] \ . \tag{1.16}$$

A comparison between Eq. (1.12) and Eq. (1.15) shows that $\check{S}_{\mathcal{E}}$ and $S_{\mathcal{E}}$ are connected by a transposition of indices: if the matrix elements of the first are $\mathcal{E}_{ij}^{lm}$, the ones of the second are $\mathcal{E}_{il}^{jm}$, or in other terms

$$\check{S}_{\mathcal{E}} = (\ S_{\mathcal{E}}^{T_2}\ E)^{T_2} = (E\ S_{\mathcal{E}}^{T_1}\ )^{T_1}\ , \tag{1.17}$$

where $E = \sum_{ij} |i\rangle\langle j| \otimes |j\rangle\langle i|$ is the so called *swap* operator, and O^{T_l} denotes the partial transposition of the operator O on the l-th Hilbert space.

One immediately notices that if $\mathcal{E}$ is CP, then $S_{\mathcal{E}}$ is a positive operator, since it results from the application of the extension $\mathcal{E} \otimes \mathcal{I}$ of a CP map to the positive operator $|I\rangle\!\rangle\langle\!\langle I|$. Actually, the converse holds too, namely any map defined through Eq. (1.16) with $S_{\mathcal{E}} \geq 0$ is CP. In fact, given that $S_{\mathcal{E}}$ is positive, it can be decomposed as

$$S_{\mathcal{E}} = \sum_i |A_i\rangle\!\rangle\langle\!\langle A_i| \ , \tag{1.18}$$

so that by substituting the above equation into Eq. (1.16), and applying the rules of Eq. (1.10), one finds that the resulting map can be expressed in the so called *Kraus form* [22]

$$\mathcal{E}(\rho) = \sum_i A_i \rho A_i^\dagger \ . \tag{1.19}$$

Any map of this form is completely positive, in fact the result of the action of its extension $\mathcal{E} \otimes \mathcal{I}$ on a positive operator $R \in \mathsf{B}(\mathsf{H} \otimes \mathsf{K})$ is

$$R_{\mathcal{E}} = \mathcal{E} \otimes \mathcal{I}[\,R\,] = \sum_i (A_i \otimes I)\ R\ (A_i^\dagger \otimes I)\ , \tag{1.20}$$

which is still positive since

$$\langle\!\langle \Psi | R_{\mathcal{E}} | \Psi \rangle\!\rangle = \sum_i \langle\!\langle A_i^\dagger \Psi | R | A_i^\dagger \Psi \rangle\!\rangle \geq 0\ , \quad \forall\ |\Psi\rangle\!\rangle\ . \tag{1.21}$$

Of course, any CP map $\mathcal{E}$ admits a Kraus form that can be found by decomposing $S_{\mathcal{E}}$ as we did in Eq. (1.18). When this decomposition is a diagonalization, i.e. $|A_i\rangle\!\rangle$ are the unnormalized orthogonal eigenvectors, then the related Kraus form is said *canonical*, and it has the minimum required number of operators, corresponding to the eigenvectors of $S_{\mathcal{E}}$, i.e. the cardinality

of the Kraus decomposition is rank $S_{\mathcal{E}}$. Any couple of Kraus decompositions $\{A_i\}$ and $\{B_i\}$ are connected as $B_i = \sum_j v_{ij} A_j$, where v_{ij} is an isometry (i.e. $\sum_j v_{ij} v_{jk} = \delta_{ik}$). In terms of a Kraus decomposition $\{A_i\}$ of the map $\mathcal{E}$ one can also express $\check{S}_{\mathcal{E}}$ as

$$\check{S}_{\mathcal{E}} = \sum_i A_i \otimes A_i^* \ , \tag{1.22}$$

as it easily follows from the definition of $\check{S}_{\mathcal{E}}$ in Eq. (1.11) and the first rule in Eq. (1.10).

Several properties other than complete positivity can be expressed in terms of $S_{\mathcal{E}}$ or equivalently in terms of the elements of a Kraus decomposition $\{A_i\}$, for example, the trace decreasing condition becomes

$$\mathrm{tr}_1 \, S_{\mathcal{E}} \leq I \qquad \text{or equivalently} \qquad \sum_i A_i^\dagger A_i \leq I \ , \tag{1.23}$$

where the equality sign would imply that the map is trace preserving.

If $\mathrm{rank}\, S_{\mathcal{E}} = 1$ then the map is *pure* (i.e. it preserves purity of input states), and its Kraus decomposition has only one element. Unitary evolutions are the only pure trace preserving transformations, and they play a special role since any other deterministic map can be realized as a unitary transformation acting on the system plus an ancilla whose state is then disregarded. In fact, given a Kraus decomposition $\{A_i\}_{i=1\ldots r}$ of the map $\mathcal{E}$ one can define an operator U on the Hilbert space $\mathsf{H} \otimes \mathbb{C}^r$ whose action on the vectors of the basis of the form $|m\rangle\!|0\rangle$ is defined as

$$U|m\rangle\!|0\rangle = \sum_{i=1}^{r} (A_i|m\rangle)\,|i\rangle = |m,0\rangle\!\rangle' \ , \tag{1.24}$$

Since the map is trace preserving, then $\sum_i A_i^\dagger A_i = I$, and this assures that the resulting vectors $|m,0\rangle\!\rangle'$ in Eq. (1.24) are orthonormal: the operator U can be then be easily estended to a unitary operator using a larger orthonormal set by means of the customary Gram-Schmidt procedure. By making the ancilla prepared in the state $|0\rangle$ interact with the system in the state ρ by means of the unitary transformation U, the final "local" state of the system only reads

$$\mathcal{E}(\rho) = \mathrm{tr}_2[\, U\,(\rho \otimes |0\rangle\langle 0|)\, U^\dagger \,] \ . \tag{1.25}$$

Notice that instead of disregarding the ancilla as we did in the previous equation, one could instead perform a measurement on it, for example by measuring

the orthonormal basis $|i\rangle$, thus obtaining the state of the system in correspondence of the outcome i in terms of the pure trace decreasing quantum operation

$$\rho_i = \frac{A_i \, \rho \, A_i^\dagger}{\mathrm{tr}\,[\, A_i \, \rho \, A_i^\dagger \,]} \; . \tag{1.26}$$

If we do not read the result of such a measurement, we still end up with a system in the state $\mathcal{E}(\rho) = \sum_i p(i|\rho)\rho_i$: the emergence of a non-pure quantum operation such as $\mathcal{E}$ can be interpreted as a "measurement without reading the outcome", or else as an information leakage in an environment. This is another way to understand how unitary operators describe the evolution of a closed system, whereas non pure trace preserving CP maps represent the evolution of open systems in interaction with a reservoir.

The procedure used to build U actually accomplishes a *purification* of $\mathcal{E}$ which is analogous to the purification of a mixed state, and it is a sort of purification of the operator $S_{\mathcal{E}}$. It also returns unitaries to their priviliged role at the axiomatic level.

As we argued from Eq. (1.26), it is possible to realize a trace decreasing map by means of a suitable joint unitary evolution of the system coupled with an ancilla, followed by a final projective measurement on the ancilla. Consider for example a measurement leading to N possible results $\mathsf{E} = \{1 \dots N\}$, and such that in relation to the outcome k the state is transformed according to a map $\mathcal{E}^{(e)}$ whose Kraus decomposition is $\{A_i^{(e)}\}_{i=1\dots r_e}$. If we do not read the outcomes of the measurement and we do not separate the reduced systems accordingly, the final ensemble will be described by the state

$$\mathcal{E}(\rho) = \sum_e \, p(e|\rho) \frac{\mathcal{E}^{(e)}(\rho)}{\mathrm{tr}\,\mathcal{E}^{(e)}(\rho)} = \sum_e \mathcal{E}^{(e)}(\rho) \; . \tag{1.27}$$

The map $\mathcal{E}$ is a non-pure deterministic map admitting $\{A_i^{(e)}\}$ as its Kraus decomposition, hence $\sum_{e=1}^{N} \sum_{i=1}^{r_e} A_i^{\dagger(e)} A_i^{(e)} = I$. If we define U on $\mathsf{H} \otimes \mathbb{C}^{r_{\max}} \otimes \mathbb{C}^N$ such that on the elements of the basis of the form $|m\rangle|0\rangle|0\rangle$ it behaves as

$$U|m\rangle|0\rangle|0\rangle = \sum_{e=1}^{N} \sum_{i=1}^{r_e} (A_i^{(e)}|m\rangle) \, |i\rangle|e\rangle \; , \tag{1.28}$$

then U can be completed to a unitary operator on the whole space, since the resulting vectors in the above equation are orthonormal. The original maps can now be realized by evolving with the unitary U the system in the state

ρ jointly with the two additional ancillas prepared in the state $|0\rangle|0\rangle$, and then performing a projective measurement $|e\rangle\langle e|$ on the second ancilla while disregarding the first one with a partial trace, i.e.

$$\mathcal{E}^{(e)}(\rho) = {}_3\langle e|\ \mathrm{tr}_2\left[\, U\left(\rho \otimes |0\rangle\langle 0| \otimes |0\rangle\langle 0|\right) U^\dagger \,\right]\ |e\rangle_3 \,. \tag{1.29}$$

Also in this case, the maps $\mathcal{E}^{(e)}$ are non-pure because some infomation has leaked into the first ancilla, which has been disregarded. If we would measure also the basis of the first ancilla, instead of taking the partial trace, in correspondence with the outcome (i, e) the state of the system would be described by a pure quantum operation

$$\rho_{(i,e)} = \frac{A_i^{(e)}\,\rho\,A_i^{(e)\dagger}}{\mathrm{tr}\left[\,A_i^{(e)}\,\rho\,A_i^{(e)\dagger}\,\right]} \,. \tag{1.30}$$

1.3 Positive operator valued measures (POVM)

When what matters in a quantum measurement is only the probability distribution of outcomes in relation to the state of the system, we don't need the detailed description of the measurement process given in terms of quantum operations. Following reasoning lines similar to those followed in Sec. 1.1, in particular referring to Eq. (1.3), it follows that the probability distribution of the outcomes of any quantum measurement must be linear in the state ρ, and that therefore it is described by the so called Born's rule

$$p(e|\rho) = \mathrm{tr}[\rho\, P_e] \,, \tag{1.31}$$

where e is the outcome and the set $\{P_e\}$ is called *positive operator valued measure* (*POVM*), namely it is a set of operators that must be positive and with $\sum_e P_e = I$, in order to have $p(e|\rho)$ a properly positive and normalized probability distribution [23]. In the next chapter we will denote the POVM of a measuring device as the vector $\mathbf{P} = (P_1, \ldots, P_N)$.

In the present context we are interested in deriving the connection between QO's and POVM's using the operator representation of maps. Considering a measuring process which for each outcome e is described by the CP maps $\mathcal{E}^{(e)}$, by means of Eq. (1.16), the probability distribution of the outcomes reads

$$p(e|\rho) = \mathrm{tr}[\,\mathcal{E}^{(e)}(\rho)\,] = \mathrm{tr}\left[\,\rho\ \mathrm{tr}_1[S_{\mathcal{E}^{(e)}}]^T\,\right] = \mathrm{tr}[\rho\, P_e] \,, \tag{1.32}$$

and thus the measurement maps $\mathcal{E}^{(e)}$ induce the POVM elements P_e which can also be expressed as

$$P_e = \mathrm{tr}_1[S_{\mathcal{E}^{(e)}}]^T = \sum_i A_i^{(e)\dagger} A_i^{(e)} \ , \tag{1.33}$$

$\{A_i^{(e)}\}$ being a Kraus decomposition of the e-th map. On the contrary, given a POVM $\{P_e\}$ one can always find a set of QO's $\mathcal{E}^{(e)}$ describing a measuring process with the given POVM, for example using $A_e = \sqrt{P_e}$ and $\mathcal{E}^{(e)}(\rho) = A^{(e)}\rho A^{(e)\dagger}$. By "purifying" these maps with the unitary transformation U defined in Eq. (1.28) of the previous section, we see that it is possible to realize any POVM in terms of an indirect measurement scheme in which a projective measurement is performed on an ancilla after a unitary interaction with the system [24, 25, 23, 26, 22, 27].

Chapter 2

Convexity and covariance for quantum devices

This chapter is devoted to the mathematical characterization of devices satisfying particular yet significative mathematical properties.

In first part of this chapter, the convex structure of the sets of QO's and POVM's will be discussed in relation with its physical significance, along with our results of Ref. [17] concerning the characterization of the extremal points of such sets, which turn out to be affected by purely quantum noise, and the construction of a convex decomposition algorithm for both the cases.

The second part of the chapter reports our work of Ref. [18]: we will introduce covariant QO's and their characterization, which we will employ for parametrizing covariant cloning transformations, an essential step for cloning optimization.

2.1 Convex structure of POVM's and quantum devices

Let's consider two quantum devices A and B with the same set of possible outcomes $\mathsf{E} = \{1 \dots N\}$, and with state reduction maps $\mathcal{E}_A^{(e)}$ and $\mathcal{E}_B^{(e)}$, respectively. Now suppose to build a black box in which either the device A or the device B are employed with probability p_A and $p_B = 1 - p_A$, respectively. The black box can be interpreted as a new device with state reduction maps $p_A\mathcal{E}_A^{(e)} + p_B\mathcal{E}_B^{(e)}$ and POVM $p_A\mathbf{P}^A + p_B\mathbf{P}^B$. Exactly as for states, also in the case of quantum devices, the emergence of linear convex combinations reflects

a lack of knowledge about the system, which can be thought as originated by an external source of additional randomness (in our example it is due to the random choice between the devices A and B). This randomness causes the mixing of the state reduction maps and of the POVM's, and it manifests itself as an added noise either in the statistics of the outcomes of the device, or at least in the statistics of a measurement performed later on the reduced states, where the noise is measured as an increased entropy in the statistics of the outcome, which corresponds to a flatter statistics and thus to a less sensitive measurement. In fact, within the above example, the statistics of the outcomes of the black box reads

$$p(e|\rho) = p_A \operatorname{tr}[\rho P_e^A] + p_B \operatorname{tr}[\rho P_e^B] \,, \tag{2.1}$$

so that if $\mathbf{P}^A \neq \mathbf{P}^B$, for the convexity property of the entropy, the statistics resulting from the mixing is noisier than the ones of the devices A and B alone. In the case $\mathbf{P}^A = \mathbf{P}^B$, the statistics of the box is equal to the one of the devices, however a further measurement with POVM $\mathbf{P}'$ carried on the reduced state $\rho_e = p_A \rho_e^A + p_B \rho_e^B$ corresponding to the outcome e of the box leads to the statistics

$$p(i|\rho_e) = p_A \operatorname{tr}[P'_i \rho_e^A] + p_B \operatorname{tr}[P'_i \rho_e^B] \,, \tag{2.2}$$

which is noisier then the ones one would obtain by using one device at once.

Actually, the sets of quantum devices and POVM's are convex — i.e. any linear convex combinations of their elements still belongs to the same set — and the device resulting from a combination can be physically implemented with the black box scheme we have described above, in which each component of the combination is randomly chosen with a probability equal to its weight in the combination. Therefore, each element that can be written as combination of others can be thought as affected by an additional source of classical randomness that manifests itself in noisier outcome statistics. On the other hand, indecomposable elements, i.e. the ones that cannot be realized by mixing others, will exhibit an intrinsic noise of purely quantum nature, and will generate all the other elements of the set through mixing. They are the extremal points of the considered set and, in analogy with states, we will also call them "pure".

Thus, pure devices and POVM's are of fundamental interest either because they generate all the other devices by means of mixing, or because they are free

from added noise. Moreover, they deserve attention also from the practical point of view, when dealing with problems of measurement optimization, that very often can be restricted to extremal devices or POVM's.

Reporting our results of Ref. [17], in what follows we will characterize "pure" devices and POVM's.

2.1.1 Extremal POVM's characterization

Let's denote by $\mathscr{P}_N$ the convex set of POVM's on a finite dimensional Hilbert space H, with a number N of outcomes $\mathsf{E} = \{1, \dots N\}$. We will represent a POVM in the set as the vector $\mathbf{P} = (P_1, \dots, P_N)$ of the N positive operators P_e satisfying the completeness constraint $\sum_e P_e = I$, that assures the probability distribution of the outcomes is normalized. The fact that the set $\mathscr{P}_N$ is convex means that it is closed under convex linear combinations, namely for any $\mathbf{P}', \mathbf{P}'' \in \mathscr{P}_N$ also $\mathbf{P} = p\mathbf{P}' + (1-p)\mathbf{P}'' \in \mathscr{P}_N$ with $0 \leq p \leq 1$ — i.e. p is a probability. Then, $\mathbf{P}$ can be also equivalently achieved by randomly choosing between two different apparatuses corresponding to $\mathbf{P}'$ and $\mathbf{P}''$, respectively, with probability p and $1-p$, since, the overall statistical distribution $p(e|\rho)$ will be the convex combination of the statistics coming from $\mathbf{P}'$ and $\mathbf{P}''$. Notice that $\mathscr{P}_N$ contains also the set of POVM's whose possible outcomes are a subset $\mathsf{E}' \subset \mathsf{E}$: for such POVM's the elements corresponding to outcomes in $\mathsf{E} \setminus \mathsf{E}'$ will be zero, so that they have zero probability of occurrence for all states.

The extremal points of $\mathscr{P}_N$ represent indecomposable measurements, i.e. which cannot be performed by mixing other measurements as above. Besides representing the apparatuses with purely intrinsically quantum noise, such extremal POVM's are also of practical relevance, since in any linear optimization problem of quantum estimation theory [23] the optimal apparatuses will have extremal POVM's.

Let's start with a simple example. Consider the following two-outcome POVM for a qubit

$$\mathbf{P} = \left(\frac{1}{2}|0\rangle\langle 0|,\ \frac{1}{2}|0\rangle\langle 0| + |1\rangle\langle 1| \right). \tag{2.3}$$

By defining $\mathbf{D} = \frac{1}{2}(|0\rangle\langle 0|, -|0\rangle\langle 0|)$, the two vectors $\mathbf{P}_\pm = \mathbf{P} \pm \mathbf{D}$ correspond to the following different POVM's

$$\mathbf{P}_+ = (|0\rangle\langle 0|,\ |1\rangle\langle 1|)\ , \quad \mathbf{P}_- = (0,\ I)\,, \tag{2.4}$$

which, intuitively are extremal, whereas $\mathbf{P}$ is not, since $\mathbf{P} = \frac{1}{2}\mathbf{P}_+ + \frac{1}{2}\mathbf{P}_-$.

How to assess whether a POVM is extremal or not? Looking at the above example, one notices that $\mathbf{P}$ is not extremal if there exists a vector of operators $\mathbf{D} \neq \mathbf{0}$ such that $\mathbf{P}_\pm = \mathbf{P} \pm \mathbf{D}$ are POVM's, because in this case $\mathbf{P}$ could be written as a convex combination of $\mathbf{P}_+$ and $\mathbf{P}_-$. The normalization for $\mathbf{P}_\pm$ requires vanishing sum of the elements of $\mathbf{D}$, $\sum_e D_e = 0$, whereas in order to guarantee positivity of all elements of both $\mathbf{P}_+$ and $\mathbf{P}_-$ it necessary to have D_e Hermitian and $|D_e| \leq P_e$ for all $e \in \mathsf{E}$.

The existence of a non-vanishing $\mathbf{D} = (D_e)$, with $\sum_e D_e = 0$, D_e Hermitian, and $|D_e| \leq P_e$, is therefore a sufficient condition for non extremality of the POVM $\mathbf{P} = (P_e)$. Actually, such condition is also necessary. In fact, for a non extremal $\mathbf{P}$ one has $\mathbf{P} = \sum_i p_i \mathbf{P}_i$, with $0 < p_i < 1$ strictly, and unequal POVM's $\mathbf{P}_i$. Then

$$\mathbf{P} = \mathbf{P}_1 + \left(\sum_{i>1} p_i \mathbf{P}_i - (1 - p_1)\mathbf{P}_1 \right) = \mathbf{P}_1 + \mathbf{D}, \tag{2.5}$$

with $\mathbf{D} = (D_e)$ a non vanishing vector of Hermitian operators D_e with zero sum, due to normalization of both $\mathbf{P}$ and $\mathbf{P}_1$. Also, $D_e = P_e - P_{1e}$, and thus any vector which is not in the kernel of D_e neither is in the kernel of P_e, so that at the end $\mathsf{Supp}(D_e) \subseteq \mathsf{Supp}(P_e)$, where $\mathsf{Supp}(D_e) = \mathsf{Ker}^\perp(D_e)$. Then, by rescaling $\mathbf{D}$, we can fulfill also the bound $|D_e| \leq P_e$.

Therefore, we can state the following:

Theorem 1 *A POVM $\mathbf{P} = (P_e)$ is not extremal iff there exists $\mathbf{D} = (D_e) \neq \mathbf{0}$ with Hermitian D_e and $\sum_e D_e = 0$, such that $|D_e| \leq P_e$.*

This result can be expressed in other useful equivalent forms. In fact the requirements on the D_e's can be independently relaxed: the D_e's can be simply linearly dependent instead of having zero sum, i.e. $\sum_e \lambda_e D_e = 0$ with some non-vanishing λ_e's, and they can satisfy $\mathsf{Supp}(D_e), \mathsf{Rng}(D_e) \subseteq \mathsf{Supp}(P_e)$ instead of being Hermitian and satisfying $|D_e| \leq P_e$. In fact, one can redefine each D_e with the substitution $D_e \leftarrow \lambda_e D_e + \lambda_e^* D_e^\dagger$ and a possible overall rescaling, thus obtaining a set of Hermitian D_e's such that $\sum_e D_e = 0$ and $|D_e| \leq P_e$, namely the conditions of the theorem.

The above considerations show that what really matters in assessing the extremality of a POVM $\mathbf{P} = (P_e)$ is just the condition on supports $\mathsf{Supp}(P_e)$ — not the actual values of P_e — i.e. we have the equivalent statement

Theorem 2 *The extremality of the POVM* $\mathbf{P}$ *is equivalent to the nonexistence of non trivial solutions* $\mathbf{D}$ *for the equation*

$$\sum_e D_e = 0, \quad \mathsf{Supp}(D_e), \mathsf{Rng}(D_e) \subseteq \mathsf{Supp}(P_e). \tag{2.6}$$

This condition can be made more explicit by choosing a basis $\{|v_n^{(e)}\rangle\}$ (not necessarily orthogonal or normalized) for each $\mathsf{Supp}(P_e)$: Eq. (2.6) becomes the linear homogeneous system of equations in the variables $D_{nm}^{(e)}$

$$\sum_{e\in\mathsf{E}} \sum_{nm=1}^{\mathrm{rank}(P_e)} D_{nm}^{(e)} |v_n^{(e)}\rangle\langle v_m^{(e)}| = 0, \tag{2.7}$$

where $D_{nm}^{(e)}$ are actually the components of the operators D_e on the basis $|v_n^{(e)}\rangle\langle v_m^{(e)}|$. This means the following:

Theorem 3 *A POVM* $\mathbf{P}$ *is extremal iff the operators* $|v_n^{(e)}\rangle\langle v_m^{(e)}|$ *are linearly independent.*

By choosing $|v_n^{(e)}\rangle$ as the eigenvectors of P_e, Theorem 3 is the characterization of extremal POVM's derived by Parthasaraty [28], using the correspondence between POVM's and CP-maps from the Abelian C^*-algebra built on $\mathbb{C}^N$ to operators on H, and then deriving a version of Choi's theorem for this kind of CP-maps.

It is immediate to draw from the main result some simple corollaries, such as:

Corollary 1 *If* $\sum_e \dim[\mathsf{Supp}(P_e)]^2 > d^2$*, then the POVM* $\mathbf{P} = (P_e)$ *is not extremal.*

This means that a POVM with too many elements (i. e. $N > d^2$) will be decomposable into several POVM's, each with less than d^2 non-vanishing elements.

Corollary 2 *If some elements have non-disjoint supports, then* $\mathbf{P}$ *is not extremal.*

Corollary 3 *Orthogonal POVM's are extremal.*

As an example of application of Theorem 1, one can check that the set of four projectors (multiplied by $\frac{1}{2}$) at the vertexes of a tetrahedron on the Bloch

sphere provides a four-outcome extremal POVM for a qubit. Notice that, even though this POVM is extremal, it can't generate extremal statistics, i.e. there will be no state corresponding to a sharp probability distribution — i.e. $p(e|\rho) = 1$ for a single e — thus providing an example of un-sharp POVM with purely quantum noise for all states.

2.1.2 Convex decomposition of a POVM

Here we address the problem of how to decompose a POVM $\mathbf{P} \in \mathscr{P}_N$ into extremal POVM's. The answer will be an algorithm that allows us to find those "primitive" measurements that employed randomly with the right probabilities would give the same outcome statistics as $\mathbf{P}$. Starting from $\mathbf{P} = (P_e)$, let's denote by $\mathsf{D}(\mathbf{P})$ the set of solutions $\mathbf{D} = (D_e)$ of

$$\sum_e D_e = 0. \tag{2.8}$$

with D_e Hermitian and $\mathsf{Supp}(D_e) \subseteq \mathsf{Supp}(P_e)$. Since the above equation is linear and homogeneous, $\mathsf{D}(\mathbf{P})$ is a real vector space with strictly positive dimension $r = \dim[\mathsf{D}(\mathbf{P})] > 0$ iff $\mathbf{P}$ is not extremal (see Th. 1). The elements $\mathbf{D}$ of this vector space satisfying $[D_e]^- \leq P_e$ form a convex set $\mathsf{C}(\mathbf{P})$ containing the origin, and each of them is in one-to-one correspondence with a POVM $\mathbf{P}'$ that is $\mathbf{P}$-compatible, in the sense that there is a convex decomposition of $\mathbf{P}$ containing $\mathbf{P}'$. The correspondence is $\mathbf{D} \leftrightarrow \mathbf{P} + \mathbf{D}$, in fact, thanks to the hypotheses on the supports $\mathsf{Supp}[D_e]$, for a small enough $\lambda > 0$ the inequality $\lambda[D_e]^+ \leq P_e$ holds for all $e \in \mathsf{E}$, so that $\mathbf{P} - \lambda\mathbf{D}$ will be a POVM too, and $\mathbf{P} + \mathbf{D}$ will be compatible with $\mathbf{P}$ since

$$\mathbf{P} = \frac{\lambda}{1+\lambda}(\mathbf{P} + \mathbf{D}) + \frac{1}{1+\lambda}(\mathbf{P} - \lambda\mathbf{D}). \tag{2.9}$$

Of course, the $\mathbf{D}$'s corresponding to extremal $\mathbf{P}$–compatible POVM's will belong to the boundary $\partial\mathsf{C}(\mathbf{P})$ of $\mathsf{C}(\mathbf{P})$, whence we will search the elements of extremal convex decompositions of $\mathbf{P}$ right on that boundary only.

Now, let's fix any "norm" on $\mathsf{D}(\mathbf{P})$, and consider a normalized $\hat{\mathbf{D}}$. We can move away from the origin of the space $\mathsf{D}(\mathbf{P})$ towards the two opposite directions $\hat{\mathbf{D}}$ and $-\hat{\mathbf{D}}$, until we reach $\partial\mathsf{C}(\mathbf{P})$ respectively in the two vectors $\mathbf{D}_\pm = \lambda_\pm\hat{\mathbf{D}}$, corresponding to the $\mathbf{P}$–compatible POVM's $\mathbf{P}_\pm = \mathbf{P} + \mathbf{D}_\pm$. The constant λ_+ corresponds to the greatest value of λ for which $\mathbf{P} + \lambda\hat{\mathbf{D}}$ is still a POVM, namely for which $P_e + \lambda\hat{D}_e \geq 0$ for all $e \in \mathsf{E}$. In other words, λ_+ is the

smallest positive value of λ for which at least one of the operators $P_e + \lambda \hat{D}_e$ develops a further null eigenvalue with respect to the ones that P_e previously had. Analogous considerations hold for λ_-, which, on the contrary, is the smallest value of λ such that $\mathbf{P} + \lambda \hat{\mathbf{D}}$ is still a POVM. It is easy to check that $P_e + \lambda \hat{D}_e$ has an additional null eigenvalue iff λ is a non-vanishing eigenvalue of $-P_e \hat{D}_e^\ddagger$, where $\hat{D}_e^\ddagger$ is the Moore-Penrose pseudo-inverse (see note 1 in Sec. 3.3) of $\hat{D}_e$, since $D_e D_e^\ddagger$ is the orthogonal projector on $\mathsf{Supp}(D_e)$, whence λ_+ and λ_- are respectively the smallest positive and the greatest negative numbers among all the non-vanishing eigenvalues of $-P_e \hat{D}_e^\ddagger$ for varying $e \in \mathsf{E}$.

The original POVM $\mathbf{P}$ can now be split into the convex combination $\mathbf{P} = p_+ \mathbf{P}_+ + p_- \mathbf{P}_-$ of the two "children" POVM's $\mathbf{P}_\pm$

$$\mathbf{P} \overset{p_+}{\underset{p_-}{\lessgtr}} \begin{matrix} \mathbf{P}_+ \\ \mathbf{P}_- \end{matrix}$$

with weights $p_\pm = \frac{\mp \lambda_\mp}{\lambda_+ - \lambda_-}$. For the particular choice of $\lambda_\pm$, one has that $\mathsf{D}(\mathbf{P}_\pm) \subseteq \mathsf{D}(\mathbf{P}) \setminus \mathsf{Span}(\hat{\mathbf{D}})$, whence the dimension of the space $\mathsf{D}(\mathbf{P}_\pm)$ of solutions of Eq. (2.8) for the two children is decreased at least by one with respect to $\dim \mathsf{D}(\mathbf{P})$. By applying the same splitting scheme to both children recursively, we obtain a weighted binary tree of POVM's rooted in $\mathbf{P}$, with the property that the POVM $\mathbf{P}'$ at each node can be written as convex combination of its descendants, and that the subtree starting from the node $\mathbf{P}'$ has a depth bounded by $\dim \mathsf{D}(\mathbf{P}')$ which is necessarily decreasing at each splitting. Of course the leaves of the tree are extremal POVM's $\mathbf{P}_i$, and one can combine them to obtain the original POVM $\mathbf{P}$ weighting each leaf $\mathbf{P}_i$ with the product of all weights found along the path from the root $\mathbf{P}$ to the leaf $\mathbf{P}_i$.

Unfortunately, this raw algorithm can produce up to a maximum of 2^r extremal POVM's $\mathbf{P}_i$ in the decompositions $\sum_i p_i \mathbf{P}_i = \mathbf{P}$, each leaf being addressed by the vector $\mathbf{D}_i = \mathbf{P} - \mathbf{P}_i \in \partial \mathsf{C}(\mathbf{P})$, with $\sum_i p_i \mathbf{D}_i = \mathbf{0}$. However, by the Caratheodory's theorem [29], we know that a maximum of $r+1$ elements is enough to decompose $\mathbf{P}$. In fact if the number of $\mathbf{D}_i$'s is larger than $r+1$, then they must be linearly dependent, and thus there exist λ_i's not all vanishing and not all positive such that $\sum_i \lambda_i \mathbf{D}_i = \mathbf{0}$. Since $\sum_i p_i \mathbf{D}_i + \mu(\sum_i \lambda_i \mathbf{D}_i) = \mathbf{0}$, by choosing the greatest μ such that $p_i + \mu \lambda_i \geq 0$ for all the i's, one finds that $\mathbf{0}$ can be written as a convex combination of a subset of the $\mathbf{D}_i$'s. This procedure can be applied recursively to this subset (the p_i's must be also upgraded) until one remains with the $\mathbf{D}_i$'s whose only combination giving $\mathbf{0}$ has positive coefficients: at this point their number is for sure not larger than $r+1$. We

will call this procedure "*Caratheodory reduction*". Therefore, from an intial decomposition of $\mathbf{P}$ into many elements, we end up with a decomposition with less than $r+1$ operators $\mathbf{D}_i$ and probabilities p_i such that $\sum_{i=1}^{r+1} p_i \mathbf{D}_i = \mathbf{0}$, whence

$$\mathbf{P} = \sum_{i=1}^{r+1} p_i(\mathbf{P} + \mathbf{D}_i) = \sum_{i=1}^{r+1} p_i \mathbf{P}_i. \tag{2.10}$$

Notice that the evaluation of $\lambda_\pm$ at each splitting step involves an eigenvalue evaluation, whence the algorithm generally doesn't provide analytical decompositions. Numerically, when splitting along $\hat{\mathbf{D}}$ the numerical roundoff may give $\mathsf{D}(\mathbf{P}_\pm) \subseteq \mathsf{D}(\mathbf{P})$ — instead of $\mathsf{D}(\mathbf{P}_\pm) \subseteq \mathsf{D}(\mathbf{P}) \setminus \mathsf{Span}(\hat{\mathbf{D}})$ — and in the continuation along the binary tree we will obviously end up on leaves that are only approximately extremal, with numerical errors due to those on $\lambda_\pm$ accumulated at each step, which poses a precision problem for large dimension of H. We emphasize that what makes the algorithm feasible, with respect to the case of a generic convex decomposition, is the availability of a border condition in terms of a solvable (eigenvalue) problem.

It is clear that the above decomposition algorithm can be modified by adding a "Caratheodory reduction" after each splitting step, thus keeping the number of leaves in the tree smaller than $r+1$ at each step, with the net result of maintaining the algorithm polynomial in r.

It would be also interesting to investigate the possibility of defining a set of rules for the choice of the splitting direction with the purpose of optimizing the decomposition algorithm, perhaps leading to fewer elements without any post-reduction.

2.1.3 Translating the result to QO's

We want now to briefly show how our result can be also used in the classification and convex decomposition of quantum operations describing the state reduction one has in correspondence of a fixed outcome of a measurement, whose related POVM element is equal to P. Mathematically they are all the QO's whose corresponding operator $S_{\mathcal{E}}$ satisfies the constraint $(\mathrm{tr}_1[S_{\mathcal{E}}])^T = P$, and we will call them P-maps, they form a cone, which we will call P-cone, and we are going to characterize the extremal points of this cone. The non-extremal points correspond to the state reductions that can be performed by mixing other maps, and that thus produce more mixed reduced states, a mixing that would affect the statistics of any subsequent measurement.

Here, in strict analogy with POVM's, the operator $S_{\mathcal{E}} \geq 0$ corresponds to a non extremal P-map when there exists a non vanishing operator D on $\mathsf{H} \otimes \mathsf{H}$ such that both $S_{\pm} = S_{\mathcal{E}} \pm D$ correspond to P-maps, since then $S_{\mathcal{E}} = \frac{1}{2}(S_+ + S_-)$ would be non extremal. Of course D must satisfy $\mathrm{tr}_1[D] = 0$ in order to have both $S_{\pm}$ in the same P-cone, and it must be Hermitian with $|D| \leq S_{\mathcal{E}}$ in order to have $S_{\pm}$ positive. It is clear that with lines analogous to those used for Theorem 1, one proves the condition also to be necessary. Also in this case both Hermiticity of D and the bound $|D| \leq S_{\mathcal{E}}$ can be relaxed to $\mathsf{Supp}(D), \mathsf{Rng}(D) \subseteq \mathsf{Supp}(S_{\mathcal{E}})$. By expanding $D = \sum_{ij} D_{ij} |A_i\rangle\!\rangle\langle\!\langle A_j|$ on vectors which span $\mathsf{Supp}(S_{\mathcal{E}})$ (for example the elements of the canonical Kraus decomposition of the map) the condition $\mathrm{tr}_1[D] = 0$ rewrites as

$$\sum_{ij} D_{ij} A_i^T A_j^* = 0, \tag{2.11}$$

and the theorem then asserts that the P-map $\mathcal{E}$ is extremal iff $A_i^\dagger A_j$ are linearly independent. By choosing the $|A_i\rangle\!\rangle$'s as the eigenvectors of $S_{\mathcal{E}}$ one has exactly the characterization of extremal QO's given by Choi [12].

The parallelism between QO's and POVM's is evident: the operators $S_{\mathcal{E}}$ and D take the place of $\mathbf{P}$ and $\mathbf{D}$ respectively, Eq. (2.11) substitutes Eq. (2.8), the space $\mathsf{D}(S_{\mathcal{E}})$ of the Herimitian operators D satisfying Eq. (2.11) substitutes $\mathsf{D}(\mathbf{P})$. Then, it is clear that also for QO's it is possible to find a decomposition algorithm by cascaded splittings: a direction $\hat{D}$ is chosen in $\mathsf{D}(S_{\mathcal{E}})$, the two constant $\lambda_{\pm}$ are determined as the greatest and smallest values of λ such that $S_{\mathcal{E}} + \lambda D$ is still positive, $S_{\mathcal{E}}$ being the convex combination the two children QO's $S_{\pm} = S_{\mathcal{E}} + \lambda_{\pm} D$. We end up with a weighted binary tree of QO's rooted in $S_{\mathcal{E}}$, with extremal leaves to which we can apply the Caratheodory theorem, exactly as for POVM's. Also in this case, the application of a "Caratheodory reduction" after each splitting step, would keep the decomposition algorithm polynomial.

2.1.4 Translating the result to quantum devices

The very same technique can be applied also for characterizing the extremal points of the set of quantum devices having outcomes is $\mathsf{E} = \{1 \dots N\}$. A quantum device can be represented as the vector $\mathbf{S} = (S_e)$ of positive operators $S_e = S_{\mathcal{E}(e)} \in \mathsf{B}(\mathsf{H} \otimes \mathsf{H})$ describing the state reduction for the outcome e, and fulfilling the completeness constraint $\sum_e \mathrm{tr}_1[S_e]^T = I$, which guarantees

the normalization of the probability distribution of the outcomes. In strict analogy to the previous cases, the quantum device is indecomposable iff there is not a non-vanishing vector of operators $\mathbf{D} = (D_e)$ such that $S_e \pm D_e \geq 0$, $\mathsf{Supp}(D_e), \mathsf{Rng}(D_e) \subseteq \mathsf{Supp}(S_e)$, and satisfying the homogeneous version of the completeness constraint, i.e. $\sum_e \mathrm{tr}_1[D_e]^T = 0$. In fact, in that case $\mathbf{S} \pm \mathbf{D}$ would describe two quantum devices that used at random with a probability $1/2$ would lead to the quantum device $\mathbf{S}$, which then would be decomposable.

If the operators $A_i^{(e)}$ give the canonical Krauss decomposition of the map $\mathcal{E}^{(e)}$, then the condition for the extremality of the quantum device is that all the operators $A_i^{(e)\dagger} A_j^{(e)}$ for varying i, j, e must be linearly independent. Also in this case, a decomposition algorithm can be devised along the lines of the ones given for POVM's and QO's.

2.2 Covariant QO's

Many important problems in quantum information theory rely on the implementation of physical transformations being covariant with respect to some group $\mathbf{G}$ of transformations of the input state, quantum cloning being a remarkable example. In this section we will report the our results of Ref. [18] in which the covariant quantum operations where classified in terms of their corresponding operator. These results are also used first to characterize extremal covariant QO's by means of the techniques of the previous section, and the, in the next session, for the construction of optimal covariant cloning transformations.

2.2.1 Elements of group theory

To talk about covariance, we must first give some important definitions (see, for example, Ref. [30]). A unitary (projective) representation T of the group $\mathbf{G}$ on H is an homomorphism associating any element $g \in \mathbf{G}$ to a unitary transformation $T_g \in \mathsf{B}(\mathsf{H})$ in such a way that the composition law of the group is preserved under the correspondence, i.e.

$$T_{g_1} T_{g_2} = \omega(g_1, g_2) T_{g_1 g_2} \ , \tag{2.12}$$

where $|\omega(g_1, g_2)| = 1$ are the so called co-cycles, and they satisfy the following restrictions

$$\omega(g_1 g_2, g_3)\omega(g_1, g_2) = \omega(g_1, g_2 g_3)\omega(g_2, g_3) \ ,$$
$$\omega(g, g^{-1}) = \omega(g, e) = 1 \ . \tag{2.13}$$

A unitary representation is *irreducible* (UIR) if there are not proper subspaces leaved invariant by the action of all its elements. Two $\mathbf{G}$-representations T^1 on H^1 and T^2 on H^2 are equivalent if there exists an isomorphism (a one-to-one and norm-preserving linear correspondence) $I : \mathsf{H}^1 \to \mathsf{H}^2$, such that $I\, T^1_g = T^2_g\, I$ for any $g \in \mathbf{G}$.

The most important result for UIR is the so called Schur's lemma: let T^1 on H^1 and T^2 on H^2 be irreducible unitary $\mathbf{G}$-representation, and let $B \in \mathsf{B}(\mathsf{H}^1, \mathsf{H}^2)$ satisfy

$$B\, T^1_g = T^2_g\, B \qquad \forall g \in \mathbf{G} \ , \tag{2.14}$$

if T^1 and T^2 are equivalent then B is proportional to the isomorphism I connecting them, otherwise B is null.

Group invariant operators

Suppose the $\mathbf{G}$-representation W on H is reducible, i. e. the space can be decomposed into a direct sum of minimal invariant subspaces M_i

$$\mathsf{H} = \bigoplus_{i=1} \mathsf{M}_i \ , \tag{2.15}$$

each M_i supporting a unitary irreducible representation (UIR) T^i of the group. Given this decomposition one can look at any operator O on H as a set of operators O^i_j in $\mathsf{B}(\mathsf{M_j}, \mathsf{M_i})$, so that $O = \sum_{ij} O^i_j$, in a sort of block decomposition. Being the M_i the minimal irreducible invariant subspaces, the blocks of any group representative W_g will satisfy the equation

$$(W_g)^i_j = \delta_{ij}\, T^j_g \ ,$$

where T^j is the UIR supported by M_j. Whenever two UIR T^i and T^j are equivalent, $i \sim j$, there exists an isomorphism $I^i_j \in \mathsf{B}(\mathsf{M_j}, \mathsf{M_i})$ such that $T^j = (I^i_j)^{-1} T^i I^i_j$.

Now suppose we have an operator R on H which is invariant with respect to the reducible unitary representation W of the group $\mathbf{G}$, namely

$$R\, W_g = W_g\, R \quad \forall g \in \mathbf{G}. \tag{2.16}$$

The above equation can be expressed equivalently block-by-block as

$$R^i_j\, T^j_g = T^i_g\, R^i_j \quad \forall g \in \mathbf{G}\ ,$$

so that, by Schur's lemmas, one deduces that the blocks of R must be of the form

$$R^i_j = c_{ij}\, I^i_j\ , \tag{2.17}$$

where if $i \nsim j$ then $c_{ij} = 0$, and, if $i \sim j$, c_{ij} can be different from zero.

Since equivalent unitary representations are related by an isomorphism, in any invariant subspace M_i one can choose the basis $\{|i,l\rangle,\ l = 1 \dots \dim\mathsf{M}_i\}$ so that for $i \sim j$

$$\langle i,l|T^i|i,m\rangle = \langle j,l|T^j|j,m\rangle\ , \tag{2.18}$$

and thus

$$I^i_j = \sum_l |i,l\rangle\langle j,l|\ , \tag{2.19}$$

namely I^i_j is an identity block. Therefore, the $\mathbf{G}$ invariant operator $R = \sum_{ij} c_{ij} I^i_j$, in the basis $\{|i,l\rangle\}$ is represented by a block matrix, whose block structure corresponds to the decomposition of H in (2.15), and whose blocks are proportional to the identity in the case they connect equivalent subspaces, and vanishing everywhere else.

By labeling with μ the inequivalent representations, the invariant operator R can also be written as

$$R = \sum_\mu C^\mu \otimes I^\mu\ , \tag{2.20}$$

where for each term in the above sum the tensorial polarization is different, and the vector $|i\rangle \otimes |l\rangle$ should be interpreted as $|i_\mu, l\rangle$, where i_μ labels one of the subspaces supporting μ. Hence, the first space factor has dimension equal to the multiplicity of the representation μ whereas the second one has the dimension of the subspaces supporting μ. The operator C^μ has matrix elements equal to $c_{i_\mu j_\mu}$, and actually it represents the block of coefficients connecting the equivalent subspaces.

2.2.2 Covariant QO's characterization

Now we will use the tools presented in the previous subsection to deal with covariant QO's. We would like to find a suitable parametrization of all the QO's being covariant with respect to some chosen group.

Let $\mathcal{E}$ be a QO on $\mathsf{B}(\mathsf{H})$, and let $\mathbf{G}$ be a group with unitary representations U and V on H. The map $\mathcal{E}$ is $\mathbf{G}$-covariant with respect to U and V if

$$\mathcal{E}\big(U_g\,\rho\,U_g^\dagger\big) = V_g\,\mathcal{E}(\rho)\,V_g^\dagger\ , \tag{2.21}$$

for any $\rho \in \mathsf{B}(\mathsf{H})$ and $g \in \mathbf{G}$. Physically, this equation translates the requirement that any transformation U_g on the input must be propagated to the output as the transformation V_g. For example, we could be interested in a QO operating on qubits, with the property that a rotation around the z axis of the block sphere on the input is propagated to the output as well, or maybe as a rotation around the y axis.

By means of Eq. (1.16), the covariance condition becomes

$$\begin{aligned}\mathcal{E}(\rho) &= \mathrm{tr}_{\mathsf{H}}\left[\, I\otimes\rho^T\ S_{\mathcal{E}}\,\right] = \\ &\equiv \mathrm{tr}_{\mathsf{H}}\left[\, I\otimes\rho^T\ V_g^\dagger\otimes U_g^T\ S_{\mathcal{E}}\ V_g\otimes U_g^*\,\right]\ ,\end{aligned} \tag{2.22}$$

so that we conclude that $\mathcal{E}$ is $\mathbf{G}$-covariant if and only if

$$S_{\mathcal{E}} = V_g^\dagger\otimes U_g^T\ S_{\mathcal{E}}\ V_g\otimes U_g^*\ , \quad \forall g\in\mathbf{G}\ , \tag{2.23}$$

or equivalently

$$\left[\, S_{\mathcal{E}}\ ,\ V_g\otimes U_g^*\,\right] = 0\ , \quad \forall g\in\mathbf{G}\ . \tag{2.24}$$

Thus, $\mathbf{G}$-covariance of a QO $\mathcal{E}$ is equivalent to $\mathbf{G}$-invariance of the corresponding positive operator $S_{\mathcal{E}}$ with respect to the reducible representation $W_g = V_g\otimes U_g^*$ of the group $\mathbf{G}$ on the space $\mathsf{H}\otimes\mathsf{H}$. By decomposing the space $\mathsf{H}\otimes\mathsf{H}$ into the minimal W-invariant subspaces M_i, and by choosing the right basis in each subspace, as we did in the previous subsection, we find that the operator $S_{\mathcal{E}}$ must be equal to

$$S_{\mathcal{E}} = \sum_{ij} c_{ij} I_j^i, \tag{2.25}$$

where as before $c_{ij} = 0$ if $i \nsim j$, and I_j^i is a block equal to the identity.

Since we are working with QO's, the operator $S_{\mathcal{E}}$ must be positive, and thus it could be interesting to analyze how this requirement fits into the structure of $S_{\mathcal{E}}$ dictated by the covariance condition. One can immediately see that, in order to have a positive $S_{\mathcal{E}}$, the matrix c_{ij} must be positive, since, taking $|\psi\rangle\!\rangle = \sum_i \sum_{l=1}^{dim\mathsf{M}_i} \psi_{il}\,|i,l\rangle$, one has

$$\langle\!\langle\psi|S_{\mathcal{E}}|\psi\rangle\!\rangle = \sum_{ij}\sum_{l=1}^{dim\mathsf{M}_i} \psi_{li}^*\,c_{ij}\,\psi_{lj}\ ,$$

which gives a positive result for any choice of $|\psi\rangle\!\rangle$ only in the case c_{ij} is positive.

Recalling that $c_{ij} = 0$ if $i \nsim j$, it can be convenient to reorder the indices i by grouping the ones corresponding to equivalent representations. In fact, in this way, the matrix c_{ij} assumes a block diagonal form, where different blocks correspond to inequivalent representations, and each block includes all representations equivalent to the same one, and it has a dimension equal to the multiplicity of the representation. This reordering is somewhat analogous to what one does to derive Eq. (2.20). Positivity of $S_{\mathcal{E}}$ implies positivity of each block of matrix c_{ij}, and from the point of view of Eq. (2.20) this is equivalent to the positivity of the operators C^{μ}.

We can conclude by saying that Eq. (2.25), with c_{ij} satisfying the properties prescipted above, parametrizes any **G** covariant QO. This parametrization can be effectively employed when one has to look for a covariant QO which is optimal for some given figure of merit.

2.2.3 Extremal covariant QO's

The set of P-maps (see Par. 2.1.3 for their definition) which are covariant with respect to a given group **G** can easily checked to be convex, and also in this case the very same techniques of Sec. 2.1 allow the characterization of its extremal points. Analogously to what we did in Par. 2.1.3, we find that the covariant P-map — represented by the positive operator $S_{\mathcal{E}}$ satisfying $(\mathrm{tr}_1[S_{\mathcal{E}}])^T = P$ and the covariance equation (2.25) — is not extremal iff there exists a non-vanishing operator D which fulfills the homogeneous version of the linear constraint, i.e. $(\mathrm{tr}_1[D])^T = 0$, is Hermitean with $|D| \leq S_{\mathcal{E}}$, and moreover satisfies the covariance requirement of Eq. (2.25). Also in this case the conditions of D Hermitean and $|D| \leq S_{\mathcal{E}}$ can be replaced by the simpler requirement $\mathsf{Supp}(D), \mathsf{Rng}(D) \subset \mathsf{Supp}(S_{\mathcal{E}})$. Actually, this is almost the condition for non-extremality of a generic P-map that we gave in Par. 2.1.3, but with the additional covariance requirement for D, which reflects the fact that here we are dealing with the subset of covariant P-maps.

An equivalent formulation of non-extremality can be expressed in terms of the customary matrix c_{ij} characterizing $S_{\mathcal{E}}$: the map is not extremal if there exists an hermitean matrix d_{ij} such that its support is contained in the support of c_{ij}, and $\sum_{ij} d_{ij} \,\mathrm{tr}_1[I^i_j] = 0$. This condition is the starting point for adapting the convex decomposition algorithm to the present problem.

As in the previous section, a more explicit condition for non-extremality

can be given by choosing a basis $\{v_i^{(n)}\}$ for $\mathsf{Supp}(c_{ij})$: the covariant P-map is not extremal iff the operators

$$A_{nm} = \sum_{ij} v_i^{(n)} v_j^{(m)*} \operatorname{tr}_1[I_j^i] \tag{2.26}$$

are linealy dependent.

2.3 Quantum cloning as a covariant CP map

The impossibility of perfectly cloning an unknown input state is a typical quantum feature [31, 32], nonetheless, in the laws of quantum mechanics there's enough room either to systematically produce approximate copies [33] or to make perfect copies of orthogonal states [34] or of non-orthogonal ones with a non-unit probability [35]. These possibilities have been studied in several works [36, 37, 38].

Recently, quantum cloning has entered the realm of experimental physics [39, 40]. Moreover it has became interesting from a practical point of view, since it can be used to speed-up some quantum computations [41] or to perform some quantum measurements [42, 43]. All these tasks require a "spreading" of the quantum information contained in a system into a larger system, and quantum cloning is a way to achieve it.

Such a spreading is a physical transformation and still it corresponds to a trace preserving CP map $\mathcal{E}$, the only difference with customary quantum channels being that in this case the output space K is larger than the input space H. Still the correspondence between the CP map $\mathcal{E}$ and the positive operator $S_{\mathcal{E}} = \mathcal{E} \otimes \mathcal{I}(|I\rangle\rangle\langle\langle I|)$ holds, together with all the other results, the only difference being that in this case the corresponding operator is in $\mathsf{B}(\mathsf{K} \otimes \mathsf{H})$, and analogously the elements of a Krauss decomposition live in $\mathsf{B}(\mathsf{H}, \mathsf{K})$. This kind of CP map can be realized by letting the system interact with a big ancillary composite system prepared in some way, and then disregarding only part of it. The net result is an output state living on a bigger Hilbert space given by the space of the original system and the portion of ancilla which has not been disregarded. This output will be a linear fucntion of the initial state of the system, more precisely, it will be the result of a CP map applied to the input state.

2.3.1 Optimal covariant cloning

A cloning map is just a quantum channel $\mathcal{C}$ from $\mathsf{B}(\mathsf{H})$ to $\mathsf{B}(\mathsf{H}^{\otimes \mathsf{N}})$ with the output copies invariant under the permutations of the N output spaces. This is equivalent to covariance of the map $\mathcal{C}$ with respect to a particular representation the group of permutations $\mathbf{S}_N$. Actually, it corresponds to the invariance of the positive operator $S_{\mathcal{C}}$ under the representation $W_\pi = V_\pi \otimes I$, where V is the representation of $\mathbf{S}_N$ permuting the N identical output spaces, and I is the trivial representation of $\mathbf{S}_N$ on the input space. In an equation, one has

$$\mathcal{C}(\rho) = V_\pi \, \mathcal{C}(\rho) \, V_\pi^\dagger \ , \quad \forall \, \pi \in \mathbf{S}_N \ . \tag{2.27}$$

Notice that permutation covariance does not imply that the output state has support in the symmetric subspace of the output space $\mathsf{H}^{\otimes N}$.

As explained in the previous section, Eq. (2.27) determines a peculiar block structure for the operator $S_{\mathcal{C}} \in \mathsf{B}(\mathsf{H}^{\otimes \mathsf{N}} \otimes \mathsf{H})$ associated to the map $\mathcal{C}$. Such a structure is strictly related to the decomposition of $\mathsf{H}^{\otimes N} \otimes \mathsf{H}$ into minimal invariant subspaces for $V_\pi \otimes I$. Any possible cloner is described by an $S_{\mathcal{C}}$ with that structure and satisfying the trace-preserving condition $\mathrm{tr}_{H^{\otimes N}}[S_{\mathcal{C}}] = I$ of Eq. (1.23). In this way, one classifies all possible cloning maps through the decomposition of $\mathsf{H}^{\otimes N}$ into the minimal invariant subspaces of the $\mathbf{S}_N$-representation V.

In addition to permutation invariance, which leads to a cloning map, here we will consider also an additional covariance with respect to a group of transformations $\mathbf{G}$, with representation T on H. This corresponds to the following identity

$$\mathcal{C}(T_g \, \rho \, T_g^\dagger) = T_g^{\otimes N} \, \mathcal{C}(\rho) \, T_g^{\dagger \otimes N} \ . \tag{2.28}$$

This is the definition of a $\mathbf{G}$ covariant cloning map, i.e. a map that propagates the same transformation T_g performed on the input to all the clones. Its corresponding $S_{\mathcal{C}}$ will be covariant with respect to both the permutation group and the additional group $\mathbf{G}$.

The interest in having either the cloning covariance or the additional covariance is dictated by the many situations in which the quality of the map is evaluated by a score function $\Gamma(S_{\mathcal{C}})$ which is invariant as well, namely it fulfills

$$\Gamma(S_{\mathcal{C}}) = \Gamma(T_g^{\otimes N} \otimes T_g^* \, S_{\mathcal{C}} \, T_g^{\dagger \otimes N} \otimes T_g^T) \ . \tag{2.29}$$

A typical example of invariant score function arises when one is interested in cloning a restricted covariant family of states $\rho_g = T_g \, \rho \, T_g^\dagger$ given with a

covariant a priori probability $dp(g)$. In fact, in this case the most appropriate score function is the average fidelity between the input ρ_g and a single copy $\rho'_g = \mathrm{tr}_{H^{\otimes N-1}}[\mathcal{C}(\rho_g)]$. This figure of merit is group invariant being it equal to

$$\begin{aligned}\Gamma(S_{\mathcal{C}}) &= \int_{\mathbf{G}} \mathrm{d}p(g)\,\mathrm{tr}[I^{\otimes N-1}\otimes\rho_g\,\mathcal{C}(\rho_g)] = \\ &= \int_{\mathbf{G}} \mathrm{d}p(g)\,\mathrm{tr}[I^{\otimes N-1}\otimes(T_g\rho T_g^\dagger)\otimes(T_g\rho T_g^\dagger)^T\,S_{\mathcal{C}}] = \\ &= \mathrm{tr}[I^{\otimes N-1}\otimes\rho\otimes\rho^T\,\overline{S_{\mathcal{C}}}] = \Gamma(\,\overline{S_{\mathcal{C}}}\,)\;, \end{aligned} \tag{2.30}$$

where ρ is the seed of the covariant family, and $\overline{S_{\mathcal{C}}}$ is the group-average of $S_{\mathcal{C}}$

$$\overline{S_{\mathcal{C}}} = \int_{\mathbf{G}} dp(g) T_g^{\otimes N}\otimes T_g^*\,S_{\mathcal{C}}\,T_g^{\dagger\otimes N}\otimes T_g^T\;. \tag{2.31}$$

The operator $\overline{S_{\mathcal{C}}}$ corresponds to a covariant cloning map, as it commutes with all the operators $T_g^{\otimes N}\otimes T_g^*$ as one can easily check. Actually it is the covariant counterpart of $S_{\mathcal{C}}$, and it achieves exactly the same score, as Eq. (2.30) indicates. The above equations also indicate that if the score function is invariant for some group, then there always exists an optimal cloning map which is covariant with respect to that group, and this map can be obtained by the group average of any of the other optimal maps.

When one wants to derive the optimal cloner for a given score function $\Gamma(S_{\mathcal{C}})$, one has to find the point of maximum for $S_{\mathcal{C}}$ satisfying

1. invariance under permutation and $\mathbf{G}$,

2. positivity,

3. trace preserving condition.

The constraint 1) implies that $S_{\mathcal{C}}$ must be of the form of Eq. (2.25); constraint 2) can be taken into account by writing the blocks of the representation-wise reordering of the matrix c_{ij} of Eq. (2.25) via a Cholevsky decomposition (see, for example, Ref. [44]); constraint 3) is imposed directly on the resulting parameterization of the operator $S_{\mathcal{C}}$.

In the following we will see how to exploit these parametrization to calculate some examples of optimal cloning maps.

Phase covariant qubit cloning

Here, we consider the problem of cloning a qubit in a $U(1)$-covariant fashion, where the group representation is given by

$$T_\phi = \exp\left[\frac{i}{2}\,\phi\,(I-\sigma_z)\right] \ . \tag{2.32}$$

It is the problem of optimal phase covariant cloning, which turned out to be an interesting eavesdropping strategy. Since the cloning to two copies is already given in Ref. [45], whereas the general case for N copies is very complicated, here for simplicity we will consider the case of $N = 3$ copies. We want to achieve the maximum fidelity between input and clones, when the input is an equatorial qubit

$$|\psi_\phi\rangle = T_\phi \frac{1}{\sqrt{2}}[|0\rangle + |1\rangle] = \frac{1}{\sqrt{2}}[|0\rangle + e^{i\phi}|1\rangle] \ . \tag{2.33}$$

In other terms, we want to maximize the average "equatorial" fidelity

$$F = \int_0^{2\pi} \frac{\mathrm{d}\phi}{2\pi}\ \mathrm{tr}\left[\, I^{\otimes 2} \otimes |\psi_\phi\rangle\langle\psi_\phi|\, \mathcal{C}(|\psi_\phi\rangle\langle\psi_\phi|)\,\right] \ , \tag{2.34}$$

which, by covariance, can be written as

$$F = \mathrm{tr}\left[\, I^{\otimes 2} \otimes |\psi_0\rangle\langle\psi_0| \otimes (|\psi_0\rangle\langle\psi_0|)^T\, S_{\mathcal{C}}\,\right] \ . \tag{2.35}$$

Since the equator is invariant even for spin flipping, here we will require the additional covariance with respect to the group $\mathbb{Z}_2$, with representation $\{I, \sigma_x\}$.

In order to satisfy all the covariance requirements, $S_{\mathcal{C}}$ must be invariant for permutations, phase shift, and spin flip, i. e. for products of any of the following unitary operators

$$V_\pi \otimes I \ , \quad T_\phi^{\otimes 3} \otimes T_\phi^* \ , \quad \sigma_x^{\otimes 3} \otimes \sigma_x^* \ .$$

The Hilbert space $\mathsf{H}^{\otimes 3+1}$ can be decomposed into subspaces which are irreducible with respect to the joint action of $U(1)$ and $\mathbf{S}_3$. In Table 2.1, we list the irreducible subspaces with their basis, reporting in the columns III and IV the kind of representation supported for $U(1)$ and $\mathbf{S}_N$ respectively.

Referring to Table 2.1 and Fig. 2.1, one has to group together the subspaces supporting equivalent representation for $U(1)$ and for $\mathbf{S}_3$. This leads to the peculiar block structure for the matrix c_{ij} that we mentioned in Sec. 2.2. In

Space	Unnormalized Basis	$U(1)$	$\mathbf{S}_3$	Flipped
M_1	$\|0001\rangle$	-1	T	M_5
M_2	$\|0000\rangle$	0	T	M_6
M_3	$\|1001\rangle + \|0101\rangle + \|0011\rangle$	0	T	M_7
M_4	$\|1001\rangle - \|0101\rangle$, $\frac{1}{2}\|1001\rangle + \frac{1}{2}\|0101\rangle - \|0011\rangle$	0	D	M_8
M_5	$\|1110\rangle$	3	T	M_1
M_6	$\|1111\rangle$	2	T	M_2
M_7	$\|0110\rangle + \|1010\rangle + \|1100\rangle$	2	T	M_3
M_8	$\|0110\rangle - \|1010\rangle$, $\frac{1}{2}\|0110\rangle + \frac{1}{2}\|1010\rangle - \|1100\rangle$	2	D	M_4
M_9	$\|1000\rangle + \|0100\rangle + \|0010\rangle$	1	T	M_{10}
M_{10}	$\|0111\rangle + \|1011\rangle + \|1101\rangle$	1	T	M_9
M_{11}	$\|1000\rangle - \|0100\rangle$, $\frac{1}{2}\|1000\rangle + \frac{1}{2}\|0100\rangle - \|0010\rangle$	1	D	M_{12}
M_{12}	$\|0111\rangle - \|1011\rangle$, $\frac{1}{2}\|0111\rangle + \frac{1}{2}\|1011\rangle - \|1101\rangle$	1	D	M_{11}

Table 2.1: $\mathsf{H}^{\otimes 3+1}$ decomposition into $U(1)-\mathbf{S}_3$ irreducibles. $U(1)$ acts on each subspace as a phase shift $e^{in\phi}$, where $n \in \mathbb{Z}$ (column III) labels inequivalent representation. $\mathbf{S}_3$ acts trivially (T) on one-dimensional subspaces, whereas on bidimensional ones it acts as the defining representation (D), which is the one related to the transformations of an equilateral triangle. Spin flipping connects subspaces (column V).

this example, we find that a phase- and flip-covariant cloning map is described though Eq. (2.25) by a matrix c_{ij} having the following positive diagonal blocks:

$$\{1\},\ \{2,3\},\ \{4\},\ \{5\},\ \{6,7\},\ \{8\},\ \{9,10\},\ \{11,12\}.$$

To ensure spin flipping covariance, the elements of c_{ij} connected by a flip must be equal, for example $c_{23} = c_{67}$.

At the end, to fill the blocks of c_{ij} in the right way, we need the parameters $a,\ b,\ c,\ d,\ e,\ f,\ g\ \in \mathbb{R}^+$, $\mathbf{v} \in \mathbb{R}^3$, where $d \geq e$, $f \geq g$, and $c \geq \|\mathbf{v}\|$. Table 2.2 and Fig. 2.1 explain how to employ them.

The parameters must satisfy another constraint given by the trace-preserving

Blocks	Content
$\{1\}$, $\{5\}$	a
$\{4\}$, $\{8\}$	b
$\{2,3\}$, $\{6,7\}$	$c\,I + \mathbf{v}\cdot\sigma$
$\{9,10\}$	$d\,I + e\,\sigma_x$
$\{11,12\}$	$f\,I + g\,\sigma_x$

Table 2.2: Content of the blocks of the matrix c_{ij}, chosen in order to have $S_{\mathcal{C}}$ describing the most general CP map from $\mathsf{B}(\mathsf{H})$ to $\mathsf{B}(\mathsf{H}^{\otimes 3})$ which is covariant with respect to permutations, phase shift, and spin flip.

condition defined by $\mathrm{tr}_{\mathsf{H}^N}[S_{\mathcal{C}}] = I$. Within this parameterization it reads

$$a + 2b + 2c + d + 2f = 1\ . \tag{2.36}$$

Substituting this equation into the equatorial fidelity F defined in Eq. (2.35), one has

$$F = \frac{1}{2} + \frac{1}{3}(e-g) + \frac{\sqrt{3}}{3}v_x\ . \tag{2.37}$$

This quantity can be easily maximized by hand, taking into account the constraint given by Eq. (2.36) and the properties of the parameters. The maximum fidelity is $F = \frac{5}{6}$ and is achieved for $d = e = 1$ and all the other parameters equal to zero. The value $F = \frac{5}{6}$ exceeds the bound given in Ref. [45]. The optimal phase covariant cloning is thus described by the operator

$$S_{\mathcal{C}}^{opt} = |\Phi\rangle\!\rangle\langle\!\langle\Phi|\ , \tag{2.38}$$

where

$$\begin{aligned} |\Phi\rangle\!\rangle = \ & \tfrac{1}{\sqrt{3}}[\ \ |1000\rangle + |0100\rangle + |0010\rangle + \\ & + \ \ |0111\rangle + |1011\rangle + |1101\rangle]\ . \end{aligned}$$

The Kraus's decomposition of the optimal cloner is $\mathcal{C}(\rho) = B\,\rho\,B^\dagger$, where

$$\begin{aligned} B = \ & \tfrac{1}{\sqrt{3}}[\ \ |100\rangle\langle 0| + |010\rangle\langle 0| + |001\rangle\langle 0| + \\ & + \ \ |011\rangle\langle 1| + |101\rangle\langle 1| + |110\rangle\langle 1|]\ . \end{aligned} \tag{2.39}$$

The fidelity for the $1 \to 2$ case is $\frac{1}{2} + \sqrt{\frac{1}{8}}$, as demonstrated in Ref. [45], and it is larger than the present $1 \to 3$ value, since the "information" is spread into a smaller number of copies.

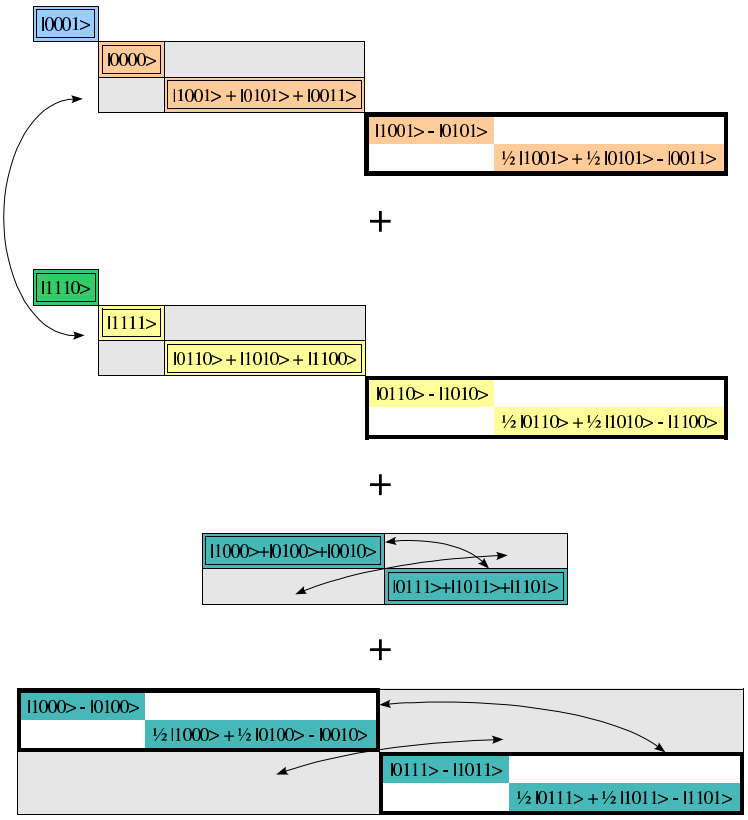

Figure 2.1: With the same color the $U(1)$ equivalent minimal invariant subspaces (EMIS), with the same border the $\mathbf{S}_N$-EMIS, the arrows connect blocks mapped one into each other by flipping and that must be equal one to each other for having flipping covariance, grey blocks connect EMIS for both $U(1)$ and $\mathbf{S}_N$.

Notice that in the general case one could have many cloning maps attaining the same global maximum of a covariant cost-function like F in Eq. (2.35). These maps can be either covariant or, if not, they are mapped one into the other by the covariance group, whence for a continuous group they make a manifold of maps. However, the non-covariant clonings are averaged into a covariant cloning via the integral (2.31). Therefore, for a linear cost-function, every optimal covariant cloning is just a convex combination of maps giving the best fidelity. In this particular example, the optimal covariant cloning is an extremal point of the convex set of all the cloning maps, either covariant or not, and thus we can assert that it is the unique optimal cloning.

Qutrits double-phase covariant cloning

In this section, we show another example of the use of the techniques presented. Here, our target will be the construction of the $1 \to 2$ qutrit cloning which gives the best average fidelity on the set of states of the form

$$|\psi_{\phi\vartheta}\rangle = \frac{1}{\sqrt{3}}[|0\rangle + e^{i\phi}|1\rangle + e^{i\vartheta}|2\rangle] \,. \tag{2.40}$$

According to what we said at the end of the latter example, such an optimal cloning can be found among the ones which are covariant with respect to ϕ and ϑ phase-rotations, and with respect to permutations of the basis $\{\,|0\rangle,\,|1\rangle,\,|2\rangle\,\}$, which are the equivalent of the spin-flip symmetry of the previous example. The above are all the symmetries of the set of input states and of the fidelity.

Our $S_{\mathcal{C}}$ will be a positive operator on $\mathsf{H}^{\otimes 2+1}$ being invariant for products of any of the following unitary transformations

$$V \otimes I \,, \quad T_{\phi\vartheta}^{\otimes 2} \otimes T_{\phi\vartheta}^{*} \,, \quad U_{\pi}^{\otimes 2} \otimes U_{\pi}^{*} \,, \tag{2.41}$$

where V is the permutation of the two clone-spaces (for two copies V is usually called "swap"), and

$$\begin{aligned} T_{\phi\vartheta} &\doteq |0\rangle\langle 0| + e^{i\phi}|1\rangle\langle 1| + e^{i\vartheta}|2\rangle\langle 2| \,, \\ U_{\pi}|i\rangle &\doteq |\pi(i)\rangle \,, \quad \forall \pi \in \mathbf{S_3} \,. \end{aligned} \tag{2.42}$$

Remember that swap-invariance makes $S_{\mathcal{C}}$ a $1 \to 2$ cloning map. The entries of Table 2.3 correspond to all the phase- and swap-invariant subspaces. Since they are all unidimensional, each is labelled by its generating vector.

Let us define the operators O_i, $i = 1...5$, on $\mathsf{H}^{\otimes 2+1}$ having the following matrix elements with respect to the basis reported in Table 2.4

$$\begin{aligned} O_1 &\to \begin{pmatrix} a & d & d \\ d^* & b & c \\ d^* & c & b \end{pmatrix} \,, \qquad O_2 \to \begin{pmatrix} e & f \\ f & e \end{pmatrix} \,, \\ O_3 &\to g \,, \qquad O_4 \to h \,, \qquad O_5 \to i \,. \end{aligned} \tag{2.43}$$

These five operators are clearly invariant with respect to swapping and phase-shifts, as one can see by comparing their expressions with Table 2.3. Sums of operators of the form of the O_i and of the form of the operators obtained by O_i with permutations, i.e. by acting on each O_i with $U_{\pi}^{\otimes 2} \otimes U_{\pi}^{*}$, are

$\otimes$	$\|0\rangle$	$\|1\rangle$	$\|2\rangle$
$\|00\rangle$	S, 0, 0, α	S, -1, 0, β	S, 0, -1, β
$\|11\rangle$	S, 2, 0, β	S, 1, 0, α	S, 2, -1, β
$\|22\rangle$	S, 0, 2, β	S, -1, 2, β	S, 0, 1, α
$\frac{1}{\sqrt{2}}[\|01\rangle + \|10\rangle]$	S, 1, 0, γ	S, 0, 0, γ	S, 1, -1, δ
$\frac{1}{\sqrt{2}}[\|02\rangle + \|20\rangle]$	S, 0, 1, γ	S, -1, 1, δ	S, 0, 0, γ
$\frac{1}{\sqrt{2}}[\|12\rangle + \|21\rangle]$	S, 1, 1, δ	S, 0, 1, γ	S, 1, 0, γ
$\frac{1}{\sqrt{2}}[\|01\rangle - \|10\rangle]$	A, 1, 0, γ	A, 0, 0, γ	A, 1, -1, δ
$\frac{1}{\sqrt{2}}[\|02\rangle - \|20\rangle]$	A, 0, 1, γ	A, -1, 1, δ	A, 0, 0, γ
$\frac{1}{\sqrt{2}}[\|12\rangle - \|21\rangle]$	A, 1, 1, δ	A, 0, 1, γ	A, 1, 0, γ

Table 2.3: $\mathsf{H}^{\otimes 2+1}$ decomposition into unidimensional invariant subspaces. The invariant subspaces are obtained by making the tensor products of any vector from the first column with either $|0\rangle$, $|1\rangle$, or $|2\rangle$: the corresponding cell in the table gives the full symmetry of the subspace. The first letter denotes the kind of action of the swap (Symmetric-Antisymmetric), the two numbers indicate the representation for ϕ and ϑ phase-shifts, respectively. Subspaces having the same greek letter are connected by a permutation $U_\pi^{\otimes 2} \otimes U_\pi^*$ for some $\pi \in \mathbf{S_3}$.

swap- and phase-invariant. One may notice that by permutations O_5 generates 3! different operators, whereas the other ones generate only 3 different operators each, since they are invariant with respect to the transposition $|1\rangle \leftrightarrow |2\rangle$.

Thanks to these observations, one realizes that these five operators give rise to five independent families of covariant cloning maps described by the invariant positive operators

$$S_{\mathcal{C}}^i = \sum_{\pi \in \mathbf{S_3}} U_\pi \, O_i \, U_\pi \ . \tag{2.44}$$

The positivity constraint for any family simply becomes $O_i \geq 0$, while the trace preserving condition leads to $\mathrm{tr}\, O_1 = \mathrm{tr}\, O_2 = \mathrm{tr}\, O_3 = \mathrm{tr}\, O_4 = 1/2$ and $\mathrm{tr}\, O_5 = 1/4$.

Any other covariant cloning map can be written as a convex linear combination of these five kind of maps in a unique way. Since the average fidelity is linear in $S_{\mathcal{C}}$, we can look for the optimal maps among these five families separately. With a little algebra one finds $\max(F_2) = 1/2$ and $F_3 = F_4 = F_5 = 1/3$, while $\max(F_1) = \frac{1}{12}(5+\sqrt{17}) \simeq 0.76$. Thus the optimal covariant cloning map

Operator	Ordered Basis
O_1	$\|000\rangle,\ \frac{1}{\sqrt{2}}[\|011\rangle + \|101\rangle],\ \frac{1}{\sqrt{2}}[\|022\rangle + \|202\rangle]$
O_2	$\frac{1}{\sqrt{2}}[\|011\rangle - \|101\rangle],\ \frac{1}{\sqrt{2}}[\|022\rangle - \|202\rangle]$
O_3	$\frac{1}{\sqrt{2}}[\|210\rangle + \|120\rangle]$
O_4	$\frac{1}{\sqrt{2}}[\|210\rangle - \|120\rangle]$
O_5	$\|001\rangle$

Table 2.4: Vector basis to which the matrix elements of the operators O_i are referred.

belongs to the $S^1_{\mathcal{C}}$ family, in particular it is obtained for the following values of the parameters

$$a = \tfrac{1}{4}(1 - \tfrac{1}{\sqrt{17}})\,, \quad c = b = \tfrac{1}{8}(1 + \tfrac{1}{\sqrt{17}})\,,$$
$$d = \sqrt{\tfrac{ab}{2}} = \tfrac{1}{2\sqrt{17}}\,, \tag{2.45}$$

which have been determined maximizing the quantity

$$F_1 = \frac{2}{3}[a + 2b + c + 2\sqrt{2}\,\mathrm{Re}(d)]\,, \tag{2.46}$$

within the constraints of trace preserving and positivity

$$a + 2b = \tfrac{1}{2}\,,$$
$$a, b \geq 0\,, \qquad |c| \leq b\,, \qquad |d|^2 \leq \frac{a(b+c)}{2}\,. \tag{2.47}$$

Cloning of continuous variables

The parameterization of channels though the related positive operator and its specialization to the covariant case are useful tools for engineering measurements. The idea is to "spread" a quantum state of H on a larger system K with a channel $\mathcal{E}$, and then to perform a measurement on the spread state. The connection between the POVM M_e on the larger space K and the resulting one M'_e on H is given by

$$M'_e = \mathcal{E}^\vee(M_e) \doteq \mathrm{tr}_{\mathsf{K}}\left[\, M_e \otimes I\; S_{\mathcal{E}}^{T_{\mathsf{H}}}\,\right]\,, \tag{2.48}$$

where $\mathcal{E}^\vee$ is the dual map of $\mathcal{E}$, and the symbol T_{H} stands for transposition with respect to H only.

In Ref. [42], the cloning map for continuous variables of Ref. [46] is used to achieve the optimal POVM for the joint measurement of two conjugated quadratures X_0 and $X_{\frac{\pi}{2}}$ of an oscillator mode a (where $X_\phi = \frac{1}{2}[a^\dagger e^{i\phi} + a e^{-i\phi}]$) by measuring them separately on the two clones. Here, we will briefly show how our general method works on this problem.

Denote by H_3 the input space and by H_1, H_2 the two output spaces of the oscillator modes a_3, a_1, a_2 respectively. The cloning is described by

$$S_{\mathcal{C}} = \frac{1}{2} P_{12} \otimes I_3 \; I_1 \otimes (|I\rangle\!\rangle\langle\!\langle I|)_{23} \; P_{12} \otimes I_3 \; , \tag{2.49}$$

where $P = V\,|0\rangle\langle 0| \otimes I\, V^\dagger$, and V is the 50% beam splitter unitary transformation $V = \exp[\frac{\pi}{4}(a_1^\dagger a_2 - a_1 a_2^\dagger)]$.

A simple calculation shows that

$$P = \frac{2}{\pi} \int \mathrm{d}^2\alpha \, |\alpha\rangle\langle\alpha|^{\otimes 2} \; , \tag{2.50}$$

where $|\alpha\rangle = D(\alpha)|0\rangle$, and $D(\alpha) = e^{\alpha a^\dagger - \bar{\alpha} a}$ is the displacement operator generating the *Weyl-Heisenberg* (WH) group. By means of Eq. (2.50), the invariance of $S_{\mathcal{C}}$ defined in Eq. (2.49) with respect to permutations and displacements can be easily verified.

Using the dual cloning map as in Eq. (2.48), we should check that

$$\mathcal{C}^\vee(E_x^0 \otimes E_y^{\pi/2}) = \frac{1}{\pi}|\alpha\rangle\langle\alpha| \; , \quad \alpha = x + iy \; , \tag{2.51}$$

where $E_x^\phi = |x\rangle_\phi \, {}_\phi\langle x|$, and $X_\phi |x\rangle_\phi = x|x\rangle_\phi$. In fact, the last term of Eq. (2.51) is the well-known optimal POVM for the joint measurement of conjugated quadratures, whereas E_x^ϕ is the POVM of the ϕ-quadrature measurement. Hence identity (2.51) guarantees that the cloning achieves the optimal joint measurement of the two conjugated quadrature via commuting measurements on clones.

Noticing that

$$E_x^0 \otimes E_y^{\pi/2} = D(\alpha)^{\otimes 2} E_0^0 \otimes E_0^{\pi/2} D(\alpha)^{\otimes 2\dagger} \; , \tag{2.52}$$

and exploiting the WH covariance, Eq. (2.51) reduces to

$$\mathcal{C}^\vee(E_0^0 \otimes E_0^{\pi/2}) = \frac{1}{\pi}|0\rangle\langle 0| \; . \tag{2.53}$$

Substituting Eq. (2.48) into this last equation, and taking matrix elements $\langle i|\dots|j\rangle$, one finally must check that

$$ {}_0\langle 0|\,{}_{\frac{\pi}{2}}\langle 0|\langle i|S_{\mathcal{C}}|0\rangle_0|0\rangle_{\frac{\pi}{2}}|j\rangle = \frac{1}{\pi}\delta_{i0}\,\delta_{j0}\ . \tag{2.54} $$

Since $V|0\rangle_0|0\rangle_{\frac{\pi}{2}} = \sqrt{\frac{2}{\pi}}|I\rangle\rangle$ and $V|0\rangle|0\rangle = |0\rangle|0\rangle$ (see Ref. [47]), one has that $P|0\rangle_{\frac{\pi}{2}}|0\rangle_0 = \sqrt{\frac{2}{\pi}}|0\rangle|0\rangle$. Thus Eq. (2.54) holds, and the cloning really achieves the wanted POVM.

Universal cloning

Clearly, the universal covariant cloning of Werner [37] is a special case of covariant cloning for the covariance group $U(d)$, $d = \dim\mathsf{H}$, of all unitary operators on H. Here, for sake of comparison to Ref. [37], we consider more generally the cloning from M to $N > M$ copies. Hence the cloning is a CP map $\mathcal{C}$ from $\mathsf{B}(\mathsf{H}^{\otimes\mathsf{M}})$ to $\mathsf{B}(\mathsf{H}^{\otimes\mathsf{N}})$ such that for any $U \in U(d)$ and $\sigma \in \mathsf{B}(\mathsf{H})$

$$ \mathcal{C}(U^{\otimes M}\sigma^{\otimes M}U^{\dagger\otimes M}) = U^{\otimes N}\mathcal{C}(\sigma^{\otimes M})U^{\dagger\otimes N}\ . \tag{2.55} $$

The score function to optimize is the fidelity between clones and input

$$ \Gamma(S_{\mathcal{C}}) = \mathrm{tr}\left[\,\sigma^{\otimes N}\,\mathcal{C}(\sigma^{\otimes M})\,\right]\ , \tag{2.56} $$

where σ is pure. Owing to covariance, the fidelity Γ does not depend on σ, since any pure state lies in the $U(d)$ orbit of any other pure state.

The optimal cloning map of Ref. [37] is given by

$$ \mathcal{C}(\rho) = \frac{d(M)}{d(N)}S_N(\rho\otimes I^{\otimes(N-M)})S_N\ , \tag{2.57} $$

where $\rho \in \mathsf{B}(\mathsf{H}^{\otimes\mathsf{M}})$, S_N is the projector on the symmetric subspace $\mathsf{H}_+^{\otimes N}$, and $d(N) = \dim(\mathsf{H}_+^{\otimes N})$. In our framework, one has

$$ S_{\mathcal{C}} = \frac{d(M)}{d(N)}\,\tilde{S}\,I_{\mathsf{H}^{\otimes(N-M)}}\otimes(|I\rangle\rangle\langle\langle I|)_{\mathsf{H}^{\otimes(M+M)}}\,\tilde{S}\ , \tag{2.58} $$

where $\tilde{S} = S_N\otimes I^{\otimes M}$. It can be easily verified that $S_{\mathcal{C}}$ is both covariant and permutation invariant as it must be.

Chapter 3

Imprinting quantum operations into quantum states

Characterizing a quantum device means to perform measurements providing information about the QO operated by the device. However, quantum measurements can only give information about the state of a system, and that's why we need to devise a way to encode the information about the QO into a quantum state. This will then allow us to use the whole theory of state-discrimination and state-tomography also for discrimination and tomography of QO's.

The way to encode the QO of a device on the state is to let the device act on some systems suitably prepared, so that their final states contain an imprinting of the device. The aim of this chapter is to classify, along the lines of our work in Ref. [16], the input states that support a full imprinting of the QO of the device, i. e. what we call *faithful states*. We will also contextually consider the case in which the information on the QO is carried not by a single state, but by a set of them, and we will correspondingly call this set *faithful*.

After briefly recalling the first proposed methods for quantum process tomography [3, 4], based on the use of many different input states, we shall report our result of Ref. [10] showing how a single pure entangled state can support a full imprinting of the QO. Then we will extend the analysis to mixed states, reporting our results of Ref. [16] showing how entanglement is not strictly needed, and finally giving a complete characterization of faithful states and

sets of states, along with a measure of their "faithfulness". It will become clear that the possibility of characterizing a device with a single fixed input state is a distinctive feature of Quantum Mechanics with no classical analog, and it is rooted in the tensor-product nature of composite quantum systems, instead of the cartesian-product "classical" composite systems. However, the fact that entanglement is not strictly necessary for faithfulness also indicates that the classical input-output correlations are enough to represent the device itself, but using a set of states, whereas the possibility of imprinting a complete description of the device into these correlations for a "single passage" of the device intimately pertains to Quantum Mechanics.

In what follows, we will first restrict the analysis to devices performing quantum processes (i.e. deterministic QO's), and then we will extend the treatment to devices performing non deterministic QO's. Finally we will also present our result of Ref. [19] on the encoding of a POVM on quantum states.

3.1 Sets of input states versus a single entangled state

The first proposed method for quantum process tomography [3, 4] exploited the linearity of the map representing the process, and since a linear operator is defined by its action on a set of vectors spanning the Hilbert space, in the same way, any linear map is completely defined by its action on a set of operators generating the linear space of all operators $\mathsf{B}(\mathsf{H})$. Hence, for encoding a quantum process $\mathcal{E}$ on states, one should look for a set of states ρ_i which span $\mathsf{B}(\mathsf{H})$, since then their respective output states $\mathcal{E}(\rho_i)$ would completely determine $\mathcal{E}$, namely the set of states would be *faithful*. In fact, the action of the map $\mathcal{E}$ on a generic ρ can be recovered by expanding ρ on the generators of the space, $\rho = \sum_i c_i \rho_i$, so that by linearity one obtains the action of the map as $\mathcal{E}(\rho) = \sum_i c_i \mathcal{E}(\rho_i)$.

As an example, consider the set of states given in Ref. [48] for quantum process tomography

$$\left\{ |m\rangle,\, |\phi_{mn}\rangle = \frac{|m\rangle + |n\rangle}{\sqrt{2}},\, |\psi_{mn}\rangle = \frac{|m\rangle + i|n\rangle}{\sqrt{2}} \right\} \tag{3.1}$$

it is a *faithful set of states*, as it is a set of generators for $\mathsf{B}(\mathsf{H})$ because the

elements of the basis $|m\rangle\langle n|$ of $\mathsf{B}(\mathsf{H})$ can be written as

$$|m\rangle\langle n| = |\phi_{mn}\rangle\langle\phi_{mn}| + i|\psi_{mn}\rangle\langle\psi_{mn}| - \frac{1+i}{2}|m\rangle\langle m| - \frac{1+i}{2}|n\rangle\langle n| \,. \qquad (3.2)$$

Quantum process tomography has been realized with this methods in liquid nuclear magnetic resonance systems [5, 6, 7], and for qubits encoded in the polarization of a radiation mode [8, 9], all situations where the dimension of the Hilbert space of the system is small. A method using the eigenstates of the quadrature operator as inputs has also been proposed in Ref. [49], for a phase-space representation of quantum transformations.

The above method has its main drawback in the difficulty – usually impossibility – of preparing the needed number – of the order of $\dim(\mathsf{H})^2$ – of different inputs. As we will see in the following, the method also turns out to be quite inefficient in achieving the information on the channel with a minimal number of measurements (the point is not that the device must be used several times to imprint the information on the channel only once, since quantum tomography even of a single output state would need anyway many measurements).

A viable alternative to the above method of "spanning states", inspired by the operator representation of a channel (1.15), was presented by us in Ref. [10] and by others in Ref. [50], and experimentally implemented for polarization qubits in Refs. [13, 14, 15]. By preparing a bipartite system in the initial state $R = |A\rangle\!\rangle\langle\!\langle A|$ and letting the first subsystem evolve under the map, as depicted in Fig. 3.1, the output state $R_{\mathcal{E}}$ reads

$$R_{\mathcal{E}} = (\mathcal{E}\otimes\mathcal{I})\,[\,|A\rangle\!\rangle\langle\!\langle A|\,] = (I\otimes A^T)\,S_{\mathcal{E}}\,(I\otimes A^*)\,. \qquad (3.3)$$

It is clear that whenever the operator A is invertible (i.e. A is full rank, or equivalently the bipartite system is in a maximal Schmidt's number entangled state) it is possible to recover $S_{\mathcal{E}}$ from $R_{\mathcal{E}}$ by the simple inversion

$$S_{\mathcal{E}} = [I\otimes(A^T)^{-1}]\,R_{\mathcal{E}}\,[I\otimes(A^*)^{-1}]\,, \qquad (3.4)$$

and then the action of the map on a state ρ is found via Eq. (1.16), namely

$$\mathcal{E}(\rho) = \mathrm{tr}_2[\,(I\otimes\rho^T)\,S_{\mathcal{E}}\,]\,. \qquad (3.5)$$

Summarizing, any bipartite state with maximal Schmidt number is faithful, namely by entering a quantum device it gets imprinted of the full information

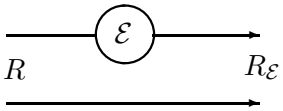

Figure 3.1: Encoding the information about a quantum device on an bipartite state. Two identical quantum systems are prepared in the state R. One of the two systems enters the device and undergoes the map $\mathcal{E}$, whereas the other is left untouched. The joint output state contains information on $\mathcal{E}$. When such information is complete the state R is called *faithful*. A pure input $R = |A\rangle\!\rangle\langle\!\langle A|$ is faithful iff $\operatorname{rank} A = \dim(\mathsf{H})$.

about the device. This method for encoding a channel on a state exploits the quantum parallelism of entanglement, with a fixed bipartite entangled state playing the role of the several input states of the previous method. The information on the device is encoded in a "native" way, which perfectly reflects the nature of the CP map representing the device itself. Moreover, it is encoded with a single use of the device, in contrast to the many uses of the method based on the generating set of states (this feature can be exploited at best in the context of devices discrimination [20], where a single measurement after the imprinting is allowed). Of course, when no prior knowledge of the device is provided, in order to recover the encoded information we have to perform a full quantum tomography of the output state, whence many copies of the imprinted state are still necessary. However, the main advantage of the method based on a single entangled state resides on the fact that a generating set of states is often not available in the lab, whereas we can produce entangled states: this is the case, for example, of quantum optics (in the domain of so-called *continuous variables* in contrast single qubits encoded on polarization of single photons), where a faithful entangled state is provided by a twin-beam from parametric down-conversion of vacuum, whereas photon number states and their superpositions as in Eq. (3.1) will remain an impossible dream for many years. Another relevant advantage of the single-pure-state method versus the generating-set one is a much higher statistical efficiency, i.e. the number of measurements needed to achieve a given statistical error in the reconstruction of the map of the device. In addition, thanks to the "native way" of encoding the transformation — reflecting both completely positivity and trace preserving/decreasing property of the map — the use of the single input state

allows an easy implementation of the maximum likelihood strategies for the characterization of the device.

All the above observations will be analyzed in detail later in this chapter, when a measure of "faithfulness" will be introduced, and also in the next chapter, where some practical applications of this framework for characterizing quantum devices will be exposed.

3.2 Faithful states

In the previous paragraph we showed that a pure entangled bipartite state $|A\rangle\!\rangle$ supports the imprinting of a quantum channel whenever the operator A is invertible. Here we want to extend this result to a generally non pure input state R, in order to characterize all faithful states.

So let's consider a bipartite state R, with spectral decomposition $R = \sum_l |A_l\rangle\!\rangle\langle\!\langle A_l|$. By applying the relation $|A_l\rangle\!\rangle = (I \otimes A_l^T)|I\rangle\!\rangle$, we can rewrite the corresponding output state $R_{\mathcal{E}} = (\mathcal{E} \otimes \mathcal{I})[R]$ as

$$\begin{aligned} R_{\mathcal{E}} = (\mathcal{E} \otimes \mathcal{I})[\, R\,] &= \sum_l (\, I \otimes A_l^T\,)\, (\mathcal{E} \otimes \mathcal{I})[\, |I\rangle\!\rangle\langle\!\langle I|\,]\, (\, I \otimes A_l^*\,) = \\ &= \sum_l (\, I \otimes A_l^T\,)\, S_{\mathcal{E}}\, (\, I \otimes A_l^*\,)\,. \end{aligned} \tag{3.6}$$

If we define the completely positive map $\mathcal{R}$ as

$$\mathcal{R}(\rho) = \sum_l A_l^T \rho A_l^*\,, \tag{3.7}$$

it is immediate to notice that

$$R_{\mathcal{E}} = (\mathcal{I} \otimes \mathcal{R})\,[S_{\mathcal{E}}]\,, \tag{3.8}$$

and therefore whenever the map $\mathcal{R}$ is invertible the output state $R_{\mathcal{E}}$ will be in one-to-one correspondence with $S_{\mathcal{E}}$, and thus with the map $\mathcal{E}$, namely it will contain all the information about the map.

From the above considerations it follows that the input state R is faithful iff it leads to a map $\mathcal{R}$ which is invertible. Recalling what we wrote in Sec. 1.2, and in particular Eq. (1.14), the invertibility of the CP map $\mathcal{R}$ resorts to the invertibility of a customary operator. In fact, by consider the following equation involving vectors in $\mathsf{H} \otimes \mathsf{H}$

$$|\mathcal{R}(\rho)\rangle\!\rangle = |\sum_l A_l^T \rho A_l^*\rangle\!\rangle = (\sum_l A_l^T \otimes A_l^\dagger\,)\,|\rho\rangle\!\rangle \doteq \check{S}_{\mathcal{R}}|\rho\rangle\!\rangle\,, \tag{3.9}$$

one realizes that the map $\mathcal{R}$ is invertible iff the relation between vectors $|\mathcal{R}(\rho)\rangle\!\rangle \leftrightarrow |\rho\rangle\!\rangle$ is invertible, and looking at the above equation it is clear that this happens iff the operator $\check{S}_{\mathcal{R}} \doteq \sum_l A_l^T \otimes A_l^\dagger$ on $\mathsf{H}\otimes\mathsf{H}$ is invertible. As we already noticed in Sec. 1.2, the action of the inverse map $\mathcal{R}^{-1}$ can be defined through the relation

$$|\mathcal{R}^{-1}(\rho)\rangle\!\rangle \doteq \check{S}_{\mathcal{R}}^{-1}|\rho\rangle\!\rangle \;, \tag{3.10}$$

so that $|\mathcal{R}^{-1}(\mathcal{R}(\rho))\rangle\!\rangle = \check{S}_{\mathcal{R}}^{-1}\check{S}_{\mathcal{R}}|\rho\rangle\!\rangle = |\rho\rangle\!\rangle$. The operator $\check{S}_{\mathcal{R}}$ can be expressed directly in terms of R, without having to evaluate its spectral decomposition, as

$$\check{S}_{\mathcal{R}} = (ER)^{T_2}E = (R^{T_2}E)^{T_1} \tag{3.11}$$

where $E = \sum_{ij} |ij\rangle\langle ji|$ is the swap operator, and O^{T_l} denotes the partial transposition of the operator O on the lth Hilbert space.

In summary, we have found that R is faithful iff $\check{S}_{\mathcal{R}}$ is invertible. In this case the relation between the output state $R_{\mathcal{E}} = (\mathcal{E}\otimes\mathcal{I})[R]$ and the operator $S_{\mathcal{E}}$ is one-to-one, with all the information about the CP map $\mathcal{E}$ encoded in $R_{\mathcal{E}}$. The map $\mathcal{E}$ can be recovered from the joint output state $R_{\mathcal{E}}$ as follows

$$\mathcal{E}(\rho) = \mathrm{tr}_2\,\left[\,(I\otimes\rho^T)\,(\mathcal{I}\otimes\mathcal{R}^{-1})[R_{\mathcal{E}}]\,\right] . \tag{3.12}$$

Later we will show some examples of faithful states, and among them there will be also separable states. On first sight this may be surprising, but it becomes obvious if one realizes that the set of faithful states is *dense*, because it is related to the set of invertible operators which is dense too.

As a further generalization, we now discuss the faithfulness of the bipartite state R of two quantum systems described by different Hilbert spaces H and K. We need now to consider vectors in either $\mathsf{H}\otimes\mathsf{K}$, $\mathsf{H}^{\otimes 2}$, or $\mathsf{K}^{\otimes 2}$, and in all cases we will keep our notation $|A\rangle\!\rangle$ for the vectors, with the corresponding operator A in $\mathsf{B}(\mathsf{K},\mathsf{H})$, $\mathsf{B}(\mathsf{H})$, or $\mathsf{B}(\mathsf{K})$ respectively.

Similarly to the previous reasoning lines, in relation to the biparite input state $R = \sum_l |A_l\rangle\!\rangle\langle\!\langle A_l|$ on $\mathsf{H}\otimes\mathsf{K}$, the output reads $R_{\mathcal{E}} = \mathcal{I}\otimes\mathcal{R}\,[S_{\mathcal{E}}]$, where the map $\mathcal{R}(\rho) = \sum_l A_l^T\rho A_l^*$ now is from $\mathsf{B}(\mathsf{H})$ to $\mathsf{B}(\mathsf{K})$. Then, faithfulness of R is still equivalent to the invertibility of the map $\mathcal{R}$, but now it is more generally equivalent to its *left*-invertibility [1]. The operator $\check{S}_{\mathcal{R}} = \sum_l A_l^T\otimes A_l^\dagger$ associated to $\mathcal{R}$ now maps vectors in $\mathsf{H}^{\otimes 2}$ to vectors in $\mathsf{K}^{\otimes 2}$, and it is still such that $\check{S}_{\mathcal{R}}|\rho\rangle\!\rangle = |\mathcal{R}(\rho)\rangle\!\rangle$. Again, faithfulness of R is equivalent to left-invertibility of the operator $\check{S}_{\mathcal{R}}$ from $\mathsf{H}^{\otimes 2}$ to $\mathsf{K}^{\otimes 2}$, that in turn is equivalent to the condition

rank $\check{S}_{\mathcal{R}} = \dim(\mathsf{H})^2$. Among all the possible left-inverses of $\check{S}_{\mathcal{R}}$ one can use the Moore-Penrose pseudo-inverse $\check{S}^{\ddagger}_{\mathcal{R}}$, and thus define the left-inverse of the map $\mathcal{R}$ as

$$|\mathcal{R}^{-1}(\rho)\rangle\!\rangle \doteq \check{S}^{\ddagger}_{\mathcal{R}}|\rho\rangle\!\rangle \,, \tag{3.13}$$

so that one can recover $S_{\mathcal{E}}$ from $R_{\mathcal{E}}$ by the relation $S_{\mathcal{E}} = (\mathcal{I} \otimes \mathcal{R}^{-1})\,[R_{\mathcal{E}}]$.

3.3 A measure of faithfulness

Even though in principle any faithful state can be used for encoding quantum processes on their outputs, the actual choice of the input will be dictated by some figure of merit depending on the particular situation. For example, consider the case in which we want to discriminate bewteen two processes $\mathcal{E}_1$ and $\mathcal{E}_2$. For input state R, their respective outputs will be

$$R_{\mathcal{E}_1} = (\mathcal{I} \otimes \mathcal{R})\,[S_{\mathcal{E}_1}] \quad \text{and} \quad R_{\mathcal{E}_2} = (\mathcal{I} \otimes \mathcal{R})\,[S_{\mathcal{E}_2}] \,, \tag{3.16}$$

and thus we shall tune R in order to improve the distingishability of these two outputs.

More generally, we see that an overall performance indicator for the faithfulness of the state R is a measure of its ability to keep outputs corresponding to different processes as far as possible in average, namely the ability of the map $\mathcal{R}$ in Eq. (3.16) to keep its outputs as far as possible. By considering the singular value decomposition for the operator $\check{S}_{\mathcal{R}}$

$$\check{S}_{\mathcal{R}} = \sum_i \sigma_i |V_i\rangle\!\rangle\langle\!\langle U_i| \,, \tag{3.17}$$

[1] A generic operator $T : \mathsf{H} \to \mathsf{K}$ is left-invertible iff rank $T = \dim(\mathsf{H})$. For having T left-invertible is therefore necessary that $\dim(\mathsf{K}) \geq \dim(\mathsf{H})$, the inverse being unique whenever the equality holds, whereas non-unique in the case of a strict inequality. Among the infinitely many possible left-inverses, the Moore-Penrose pseudo-inverse $T^{\ddagger}$ [51] is the most used one, due to its nice properties. Starting from the *singular values decomposition* (SVD) of T

$$T = \sum_i \sigma_i |v_i\rangle\langle u_i| \,, \tag{3.14}$$

where $\{|v_i\rangle\}$ and $\{|u_i\rangle\}$ are two sets of orthonormal vectors, and σ_i are positive real numbers (the singular values), $T^{\ddagger}$ is defined as

$$T^{\ddagger} = \sum_i \sigma_i^{-1} |u_i\rangle\langle v_i| \,. \tag{3.15}$$

By definition, $Q = T^{\ddagger}T$ is the orthogonal projector on $\mathsf{Supp}(T) \equiv \mathsf{Ker}(T)^{\perp}$, whence $T^{\ddagger}$ inverts T on its support, which for a left-invertible operator coincides with the whole space H.

with $\{|V_i\rangle\rangle\}$ and $\{|U_i\rangle\rangle\}$ sets of orthonormal vectors, and $\sigma_i > 0$, and by remembering that $|\mathcal{R}(\rho)\rangle\rangle = \check{S}_{\mathcal{R}}|\rho\rangle\rangle$, the action of $\mathcal{R}$ on an operator ρ becomes

$$\mathcal{R}(\rho) = \sum_i \sigma_i \, \mathrm{tr}[U_i^\dagger \rho] \, V_i \; , \tag{3.18}$$

whence it is clear that the smaller are the singular values σ_i, the nearer are the outputs of $\mathcal{R}$, since their components on the basis $\{|V_i\}$ will be shrinked. Therefore, in summary, the larger are the singular values of $\check{S}_{\mathcal{R}}$ the better is the chosen input state R.

Thus, a synthetic measure of faithfulness could be for example

$$F(R) = \sum_i \sigma_i^2 = \mathrm{tr}[\check{S}_{\mathcal{R}}^\dagger \check{S}_{\mathcal{R}}] \; . \tag{3.19}$$

This quantity can be expressed in a more meaningful form by observing that if we use the spectral decomposition $R = \sum_i |A_i\rangle\rangle\langle\langle A_i|$, with the vectors $|A_i\rangle\rangle$ being an orthogonal basis, namely $\langle\langle A_i|A_j\rangle\rangle = \mathrm{tr}[A_i^\dagger A_j] \propto \delta_{ij}$, then $\check{S}_{\mathcal{R}} = \sum_i A_i^* \otimes A_i^\dagger$, and thus the following equations hold

$$\begin{aligned} \mathrm{tr}[\check{S}_{\mathcal{R}}^\dagger \check{S}_{\mathcal{R}}] &= \sum_{ij} \mathrm{tr}[A_i^T A_j^*] \, \mathrm{tr}[A_i A_j^\dagger] = \sum_i \mathrm{tr}[A_i^T A_i^*] \, \mathrm{tr}[A_i A_i^\dagger] = \\ &= \sum_i (\langle\langle A_i|A_i\rangle\rangle)^2 = \mathrm{tr}[R^\dagger R] \; . \end{aligned} \tag{3.20}$$

Therefore, from Eq. (3.19) one obtains

$$F(R) = \mathrm{tr}[R^\dagger R] \; , \tag{3.21}$$

so that faithfulness of a state turns out to be exactly its purity. This result implies that *faithful pure states* are the optimal faithful states, and that they yield outputs states encoding the maps which are the most far apart.

The definition of $F(R)$ can be also interpreted in another way. Imagine to implement quantum process tomography using a finite number of copies of R as input states, and then to reconstruct the output $R_{\mathcal{E}}$. The measured $R_{\mathcal{E}}$ will be affected by experimental errors which will be mostly independent on $R_{\mathcal{E}}$ itself, and these errors will be propagated to the experimental estimation of $S_{\mathcal{E}}$ by the inversion map $\mathcal{R}^{-1}$. Since, in practice, the inversion map involves multiplications by σ_i^{-1}, then the smaller are the singular values of $\check{S}_{\mathcal{R}}$ the higher will be the amplification of experimental errors on the measured $S_{\mathcal{E}}$.

For an unfaithful state R, $\check{S}_{\mathcal{R}}$ ha at least one null singular value, yet $F(R)$ is different from zero. Actually, as we shall see, the state can still be used to

recover the action of a device on some inputs only. Moreover, on such inputs it can achieve an even better reconstruction resolution than a faithful state, since its faithfulness is focused on a smaller subspace.

3.4 Faithful sets of states

Now we will consider the case in which not a single bipartite state, but a certain set of them $\{R^{(n)}\}_{n=1}^{N}$ on $\mathsf{H}\otimes\mathsf{K}$ is used, and we want to understand whether or not it is faithful, namely if it supports a complete imprinting of the information about a quantum process. In other words, we will discuss when the set of outputs $\{R_{\mathcal{E}}^{(n)}\}$, with $R_{\mathcal{E}}^{(n)} = (\mathcal{E}\otimes\mathcal{I})\,[R^{(n)}]$, is a perfect encoding of a generic channel $\mathcal{E}$. This analysis will bridge the scenario with the set of generating states and the one of single bipartite faithful state.

Mathematically, it is evident that the state R_{set} on $\mathsf{H}\otimes\mathsf{K}\otimes\mathbb{C}^N$ defined as

$$R_{\text{set}} = \sum_{n=1}^{N} p_n\, R^{(n)}\otimes|n\rangle\langle n|\,, \tag{3.22}$$

where p_n are fixed non vanishing probabilities, is in 1-to-1 correspondence with the set of states $\{R^{(n)}\}$. The same correspondence holds between the output state

$$R_{\text{set}\,\mathcal{E}} = (\mathcal{E}\otimes\mathcal{I}\otimes\mathcal{I})\,[R_{\text{set}}] = \sum_{n=1}^{N} p_n\, R_{\mathcal{E}}^{(n)}\otimes|n\rangle\langle n| \tag{3.23}$$

and the set of outputs $\{R_{\mathcal{E}}^{(n)}\}$. Hence, if the state $R_{\text{set}\,\mathcal{E}}$ contains all the information about the map, then the same holds also for the set of outputs $\{R_{\mathcal{E}}^{(n)}\}$, or, equivalently, if R_{set} is faithful, then the set $\{R^{(n)}\}$ is faithful too.

Briefly, faithfulness for the set of states $\{R^{(n)}\}$ is translated into faithfulness for the single state R_{set}. The latter can be evaluated with the techniques exposed in the previous paragraph for bipartite states, by simply considering R_{set} as a bipartite state of H and $\mathsf{K}\otimes\mathbb{C}^N$.

The nature of the state R_{set} can be interpreted from two subtly different points of view. On one hand, to use R_{set} is equivalent to running all the states $\{R^{(n)}\}$ in parallel, while keeping track of each of them thanks to the tensoring with the basis $|n\rangle\langle n|$ of $\mathbb{C}^N$. On the other hand, R_{set} represents the situation in which the states $\{R^{(n)}\}$ are employed in the characterization each with a frequency equal to p_n. In fact, when the inital state is R_{set}, to measure the basis $|n\rangle\langle n|$ on $\mathbb{C}^N$ (either before or after the action of the device) is equivalent

to preparing the input $R^{(n)}$ with a probability p_n, where n is the outcome of the measurement.

For this reasons, any quantity (e.g. the faithfulness F) being defined for faithful states can be extended consistently to sets of states simply by evaluating it on the corresponding $R_{\rm set}$. For example, the faithfulness of a set of generating states ρ_n (employed with the same frequency) is equivalent to the faithfulness of the bipartite state $R_{\rm set} = \sum_n \frac{1}{n}\rho_n \otimes |n\rangle\langle n|$, and since the latter is a mixed state, it will lead to a non-optimal faithfulness. This shows why the encoding on an entangled state is theoretically better than the encoding on a set of generating states: while in the first case faithfulness is 1, in the second one it scales as $\mathcal{O}[1/\dim(\mathsf{H})]$.

3.5 Patching sets of unfaithful states

An unfaithful state R can still be useful in encoding only some quantum channels, or at least in encoding a partial information about them, which can then be used to evaluate their action on some particular states. In fact, even if the map $\mathcal{R}$ is not invertible (it maps to zero any state ρ such that $|\rho\rangle\!\rangle \in \mathsf{Ker}(\check{S}_{\mathcal{R}})$), one can still employ its pseudo-inverse $\mathcal{R}^{\ddagger}$ defined as

$$|\mathcal{R}^{\ddagger}(\rho)\rangle\!\rangle \doteq \check{S}^{\ddagger}_{\mathcal{R}}|\rho\rangle\!\rangle \ . \tag{3.24}$$

This map is such that $\mathcal{R}^{\ddagger}\mathcal{R} = \mathcal{Q}$, where $\mathcal{Q}$ is the projection map on the support of the map $\mathcal{R}$, and which is also defined by

$$|\mathcal{Q}(\rho)\rangle\!\rangle = \check{S}^{\ddagger}_{\mathcal{R}}\check{S}_{\mathcal{R}}|\rho\rangle\!\rangle = \check{S}_{\mathcal{Q}}|\rho\rangle\!\rangle \ , \tag{3.25}$$

the operator $\check{S}_{\mathcal{Q}}$ being the projector on $\mathsf{Supp}(\check{S}_{\mathcal{R}}) = \mathsf{Ker}(\check{S}_{\mathcal{R}})^{\perp}$.

It is clear that such pseudo-inversion, instead of returning the full operator $S_{\mathcal{E}}$, gives its projection

$$\tilde{S}_{\mathcal{E}} = (\mathcal{I}\otimes\mathcal{R}^{\ddagger})[R_{\mathcal{E}}] = (\mathcal{I}\otimes\mathcal{Q})[S_{\mathcal{E}}] \tag{3.26}$$

which represents a partial encoding of $\mathcal{E}$. The partially recovered map $\tilde{\mathcal{E}}(\rho) = \mathrm{tr}_2[(I\otimes\rho^T)\,\tilde{S}_{\mathcal{E}}]$ could have also been written as $\tilde{\mathcal{E}} = \mathcal{E}\mathcal{Q}^*$, $\mathcal{Q}^*$ being the projection map corresponding to the operator $\check{S}^*_{\mathcal{Q}}$. Clearly $\tilde{\mathcal{E}}$ coincides with $\mathcal{E}$ for any ρ such that $\check{S}^*_{\mathcal{Q}}|\rho\rangle\!\rangle = |\rho\rangle\!\rangle$.

For any bipartite R one can define a *number of faithfulness* φ as $\varphi(R) = \mathrm{rank}(\check{S}_{\mathcal{R}})$, i.e. as the dimension of the space of input states R for which

the action of the map $\mathcal{E}$ is described faithfully. Clearly, a state is faithful iff $\varphi(R) = \dim(\mathsf{H})^2$. Notice that for $\varphi(R) < \dim(\mathsf{H})^2$ one can have the situation in which $\mathsf{Ker}^{\perp}(\check{S}_{\mathcal{R}}) = \mathsf{Span}\{|\rho\rangle\rangle, \rho \in \mathcal{A}\}$, with $\mathcal{A}$ abelian algebra, in which case the state R allows to reconstruct completely only "classical" channels, with the input restricted to commuting states.

The introduction of pseudo-inversion provides an alternative yet equivalent way for studying the faithfulness of a set of states $\{R^{(n)}\}$. Suppose they lead to the projection maps $\{\mathcal{Q}^{(n)}\}$, then the set will be faithful iff we can recover any operator ρ from its projections $\mathcal{Q}^{(n)}(\rho)$, and this is possible iff, given a basis $\{B_i\}$ for $\mathsf{B}(\mathsf{H})$, one has $\mathsf{Span}\{\mathcal{Q}^{(n)}(B_i)\}_{i,n} = \mathsf{B}(\mathsf{H})$. In such circumstances, any element of the basis can be expressed as a linear combination of the $\mathcal{Q}^{(n)}(B_i)$, i.e. $B_i = \sum_{jn} \lambda^n_{ij} \mathcal{Q}^{(n)}(B_j)$, and therefore it is possible to recover $\rho \equiv \sum_i \mathrm{tr}[B_i^\dagger \rho] B_i$ by "patching" the projections $\mathcal{Q}^{(n)}(\rho)$ as

$$\rho = \sum_{ijn} \lambda^{n*}_{ij} \ \mathrm{tr}[B_j^\dagger \ \mathcal{Q}^{(n)}(\rho)] \, B_i \ . \tag{3.27}$$

Analogously, by patching the partial encodings $\{\tilde{S}^{(n)}_{\mathcal{E}}\}$ (see Eq. (3.26)) we get $S_{\mathcal{E}}$ as

$$S_{\mathcal{E}} = \sum_{ijn} \lambda^{n*}_{ij} \, \mathrm{tr}_2[(I \otimes B_j^\dagger) \, \tilde{S}^{(n)}_{\mathcal{E}}] \otimes B_i \ . \tag{3.28}$$

Of course this patching procedure can also be used with an unfaithful set of states, to obtain a more complete yet still partial encoding of the channel.

3.6 Generalization to QO's and POVM's

Suppose we have a quantum device performing the measurement described by the CP maps $\mathcal{E}^{(e)}$, $e = 1 \dots N$ being the outcomes, is it possible to encode all the maps or else their corresponding POVM? If we use a bipartite input state R and we let the device act on the first subsystem, the output state corresponding to the outcome e will be

$$R_{\mathcal{E}^{(e)}} = \frac{(\mathcal{I} \otimes \mathcal{R}) \, [S_{\mathcal{E}^{(e)}}]}{\mathrm{tr}[\, (\mathcal{I} \otimes \mathcal{R}) \, [S_{\mathcal{E}^{(e)}}] \,]} \ , \tag{3.29}$$

where the denominator is equal to the probability of occurrence for the outcome e. In the case of R faithful, from this output it is possible to recover $S_{\mathcal{E}^{(e)}}$ up to a normalization factor by means of the inverse map $\mathcal{R}^{-1}$.

After preparing an ensemble of systems described by a faithful state R, we let the measuring device act on them, and then we separate them according to the outcome e, thus obtaining N different ensembles, each labeled by the corresponding e, and described by the states $R_{\mathcal{E}^{(e)}}$. The denominator of Eq. (3.29) can be evaluated as the fraction of systems of the original ensemble that have been transformed into the e-th state, therefore an exact reconstruction of all the $S_{\mathcal{E}^{(e)}}$ is possible, being equivalent to the full reconstruction of the measuring device. Notice that, in contrast to what happens for a deterministic device, in the case of a probabilistic QO a single use is not enough to imprint the whole information about it, due to of the need of evaluating the normalization factor.

In many practical situations, e.g. in a photodetector, the measuring device destroys the measured system. Here, however, with the same setup with a bipartite faithful R, the reduced state ρ_e on the unmeasured system is still available, an it reads

$$\rho_e = \mathrm{tr}_1\ R_{\mathcal{E}^{(e)}} = \frac{\mathcal{R}\left[\mathrm{tr}_1\ S_{\mathcal{E}^{(e)}}\right]}{\mathrm{tr}[\mathcal{R}\left[\mathrm{tr}_1\ S_{\mathcal{E}^{(e)}}\right]]} = \frac{\mathcal{R}\left[P_e^T\right]}{\mathrm{tr}[\mathcal{R}\left[P_e^T\right]]}\,, \tag{3.30}$$

where P_e is the POVM of the measurement relative to the outcome e. Hence, by performing a quantum tomography on the above reduced output states, one can recover the POVM of the apparatus by inverting the map $\mathcal{R}$, while evaluating the denominator of the previous equation as the probability of occurrence of e.

3.7 Faithfulness and separability

Since, as we have seen, faithfulness is equivalent to an invertibility condition, the set of faithful states R is *dense* within the set of all bipartite states. Therefore, there must be faithful states among mixed separable ones, which means that classical correlations in mixed bipartite states are sufficient to support the imprinting of any quantum channel. Lets' see some examples of separable faithful states.

The Werner's states for dimension d

$$R_f = \frac{1}{d(d^2-1)}[(d-f)I + (df-1)E], \quad -1 \leq f \leq 1, \tag{3.31}$$

are separable for $f \geq 0$, however, they are faithful for all $f \neq \frac{1}{d}$. In fact, one

has

$$(ER_f)^{T_2} = \frac{1}{d(d^2-1)}[(d-f)|I\rangle\!\rangle\langle\!\langle I| + (df-1)] \,, \tag{3.32}$$

hence the singular values of $\check{S}_{\mathcal{R}_f}$ are $\frac{df-1}{d(d^2-1)}$ with multiplicity d^2-1 and $\frac{1}{d}$ with multiplicity 1. In Ref. [14] an experiment employing these states for quantum process tomography was presented.

Similarly, the "isotropic" states

$$R_f = \tfrac{f}{d}|I\rangle\!\rangle\langle\!\langle I| + \tfrac{1-f}{d^2-1}(I - \tfrac{1}{d}|I\rangle\!\rangle\langle\!\langle I|), \tag{3.33}$$

are faithful for $f \neq \frac{1}{d^2}$ and separable for $f \leq \frac{1}{d}$, the singular values of $\check{S}_{\mathcal{R}_f}$ being $\frac{d^2 f-1}{d(d^2-1)}$ and $\frac{f}{d}$.

3.8 Faithfulness in infinite dimensions

For infinite dimensions (the so-called "continuous variables" in quantum optics), one needs to restrict $\mathsf{B}(\mathsf{H})$ to the Hilbert space of Hilbert-Schmidt operators on H, and this leads to the problem that the inverse map $\mathcal{R}^{-1}$ is unbounded. The result is that we will recover the channel $\mathcal{E}$ from the measured $R_{\mathcal{E}}$, however, with unbounded amplification of statistical errors, depending on the chosen complete set of operators $\mathsf{B} = \{B_j\}$ in $\mathsf{B}(\mathsf{H})$ used for representing the channel map. As an example, let's consider a twin beam from parametric down-conversion of vacuum

$$|\Psi\rangle\!\rangle = \Psi \otimes I|I\rangle\!\rangle, \quad \Psi = (1-|\xi|^2)^{\frac{1}{2}}\xi^{a^\dagger a}, \qquad |\xi| < 1 \tag{3.34}$$

for a fixed ξ, $a^\dagger$ and a, with $[a, a^\dagger] = 1$, denoting the creation and annihilation operators of the harmonic oscillator describing the field mode corresponding to the first Hilbert space in the tensor product (in the following we will denote by $b^\dagger$ and b the creation and annihilation operators of the other field mode). The state is faithful, but the operator Ψ^{-1} is unbounded, whence the inverse map $\mathcal{R}^{-1}$ is also unbounded. In a photon number representation $\mathsf{B} = \{|n\rangle\langle m|\}$, the effect will be an amplification of errors for increasing numbers n, m of photons.

As an example, consider the quantum channel describing the *Gaussian displacement noise* [52]

$$\mathcal{N}_\nu(\rho) = \int_{\mathbb{C}} \frac{\mathrm{d}\alpha}{\pi\nu} \exp[-|\alpha|^2/\nu] D(\alpha)\rho D^\dagger(\alpha), \tag{3.35}$$

where $D(\alpha) = \exp(\alpha a^\dagger - \alpha^* a)$ denotes the usual displacement operator on the phase space. The Gaussian noise is in a sense the analogous of the depolarizing channel for infinite dimension. The maps $\mathcal{N}_\nu$ for varying ν satisfy the multiplication rule $\mathcal{N}_\nu \mathcal{N}_\mu = \mathcal{N}_{\nu+\mu}$, thus the inverse map is formally given by $\mathcal{N}_\nu^{-1} \equiv \mathcal{N}_{-\nu}$. Notice that, since the map $\mathcal{N}_\nu$ is compact, the inverse map $\mathcal{N}_\nu^{-1}$ is necessarily unbounded. As a faithful state consider now the mixed state given by the twin-beam, with one beam spoiled by the Gaussian noise, namely

$$R = \mathcal{I} \otimes \mathcal{N}_\nu(|\Psi\rangle\rangle\langle\langle\Psi|). \tag{3.36}$$

Since the (unnormalizable) vector $|D(z)\rangle\rangle = [D(z) \otimes I]|I\rangle\rangle$ is a eigenvector of the operator $Z = a - b^\dagger$, with eigenvalue z, one can easily find that

$$R = \frac{1}{\nu}(\Psi \otimes I) \exp[-(a - b^\dagger)(a^\dagger - b)/\nu](\Psi^\dagger \otimes I), \tag{3.37}$$

thus its partial transposed on the second space reads

$$R^{T_2} = (\nu + 1)^{-1}(\Psi \otimes I) \left(\frac{\nu - 1}{\nu + 1}\right)^{\frac{1}{2}(a-b)^\dagger(a-b)} (\Psi^\dagger \otimes I), \tag{3.38}$$

where transposition is defined with respect to the basis of eigenvectors of $a^\dagger a$ and $b^\dagger b$. Since our state R is Gaussian, it is separable iff its partial transposition is a positive operator [53], therefore, for $\nu > 1$, R is separable (see also Ref. [54]), yet it is *formally* faithful, since the operator Ψ and the map N_ν are both invertible. Notice that unboundedness of the inversion map can even wash out completely the information on the channel in some particular chosen representation $\mathsf{B} = \{B_j\}$, e. g. when all operators B_j are out of the boundedness domain of $\mathcal{R}^{-1}$. This is the case, for example, of the (overcomplete) representation $\mathsf{B} = \{|\alpha\rangle\langle\beta|\}$, with $|\alpha\rangle$ and $|\beta\rangle$ coherent states, since from the identity

$$\mathcal{N}_\nu(|\alpha\rangle\langle\alpha|) = \frac{1}{\nu + 1} D(\alpha) \left(\frac{\nu}{\nu + 1}\right)^{a^\dagger a} D^\dagger(\alpha), \tag{3.39}$$

one obtains

$$\mathcal{N}_\nu^{-1}(|\alpha\rangle\langle\alpha|) = \frac{1}{1 - \nu} D(\alpha) \left(1 - \nu^{-1}\right)^{-a^\dagger a} D^\dagger(\alpha), \tag{3.40}$$

which has convergence radius $\nu \leq \frac{1}{2}$, which is the well known bound for Gaussian noise for the quantum tomographic reconstruction for coherent-state and Fock representations [55]. Therefore, we say that the state is *formally* faithful, however, we are constrained to representations which are analytical for the inverse map $\mathcal{R}^{-1}$.

Chapter 4

Homodyne tomography of channels and POVM's

Once the information about a device is encoded into quantum states, all the techniques of quantum tomography can be applied to determine the channel or, more generally, the quantum operation describing the device. To date, several experiments of quantum process tomography have been implemented for qubits either in NMR systems [5, 6, 7] or in quantum optics [8, 13, 14, 9]. However, no experiments in the realm of continuous variable optical systems have been realized yet. Here, with the help of Monte Carlo simulations, we analyze the feasibility of some experiments in such context, using as a faithful state a twin-beam emerging from parametric down-conversion of vacuum, and performing a joint homodyne tomography on both the modes of radiation at the output. The actual experimental feasibility of the technique is partly proved by the experiment of Ref. [56], in which quantum homodyne tomography of the (joint number probability distribution of) a twin-beam was achieved using the setup depicted in Fig. 4.1. After a brief introduction on homodyne tomography, we report as an example of quantum process tomography the result we presented in Ref. [10] for the tomography of a displacement unitary transformation. Then we address the problem of the feasibility of the homodyne tomography of a POVM for an ON/OFF photo-detector, and a photon-counting detector. For the tomography of the unitary transformation the tomographic reconstruction will be performed by the method of pattern function averaging. For the tomography of the photo-detector, on the other hand, we will also consider maximum likelihood methods, to show how they

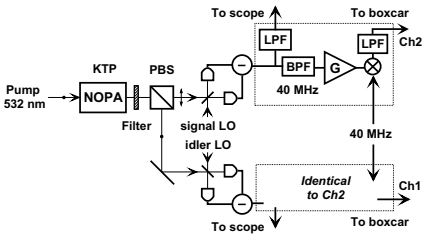

Figure 4.1: A nondegenerate optical parametric amplifier (a KTP crystal) is pumped by the second harmonic of a Q-switched mode-locked Nd:YAG laser, which produces a 100-MHz train of 120-ps duration pulses at 1064 nm. The orthogonally polarized twin-beams emitted by the KTP crystal are separately detected by two balanced homodyne setups that use two independent local oscillators derived from the same laser. The output of the apparatus is a measure of the quadrature amplitudes $X_{\phi'} \otimes X_{\phi''}$ for random phases ϕ' and ϕ'' with respect to the local oscillators. (From Ref. [56])

can give a huge boost to the precision of the characterization, at the sole expense of greater computational complexity.

Overall, homodyne tomography of processes and detectors will become a major diagnostic tool in quantum optics, opening new perspectives for the calibration of measuring apparatuses and the characterization of the dynamics of optical devices.

4.1 Homodyne tomography

A balanced homodyne detector in the strong oscillator limit ideally measures the field quadrature observable

$$X_\phi = \frac{a^\dagger e^{i\phi} + a\, e^{-i\phi}}{2} \,, \tag{4.1}$$

where a and $a^\dagger$ are the annihilation and the creation operators of the mode of interest (set by the local oscillator), for a chosen value of the phase ϕ. In the Fock basis $|n\rangle$ the (unnormalizable) eigenstate $|x\rangle_\phi$ of the quadrature X_ϕ is given by

$$|x\rangle_\phi = \sum_{n=0}^{\infty} \left(\frac{2}{\pi}\right)^{\frac{1}{4}} \frac{1}{\sqrt{2^n n!}} \exp(-x^2) H_n(\sqrt{2}x) e^{in\phi} |n\rangle \,, \tag{4.2}$$

$H_n(x)$ denoting Hermite polynomials. Once fixed the phase ϕ, the ideal measurement realizes the POVM $\mathrm{Hom}(x;\phi) = |x\rangle_{\phi\,\phi}\langle x|$ for the "continuous variable" x, with a probability density distribution of the outcomes given by

$$p(x;\phi) = \mathrm{tr}\,[\,\rho\,\mathrm{Hom}(x;\phi)\,]\ , \tag{4.3}$$

ρ being the state of the system. In the non-ideal situation with non-unit quantum efficiency, the POVM, and in turn the probability distribution of outcomes, becomes Gaussian convoluted with variance $\Delta_\eta^2 = \frac{1-\eta}{4\eta}$, the parameter η denoting the *quantum efficiency* of photo-detectors used in the homodyne.

Homodyne tomography is a method for estimating the state ρ from a finite sample of homodyne data, i.e distributed according to $p(x;\phi)$ in Eq. (4.3). The easiest strategy estimates the ensemble average of any operator O by averaging bounded *pattern function* $\mathcal{P}_\eta[O](x,\phi)$ of homodyne data. This means that one has

$$\langle O\rangle = \mathrm{tr}[\,\rho\,O\,] = \int_0^\pi \frac{\mathrm{d}\phi}{\pi}\int_{-\infty}^{+\infty}\mathrm{d}x\,p_\eta(x;\phi)\,\mathcal{P}_\eta[O](x,\phi)\ , \tag{4.4}$$

and the expectation value is achieved by averaging the pattern function on the homodyne data $\{(x_n,\phi_n)\}$ in the limit of infinitely many data

$$\frac{1}{N}\sum_{n=0}^{N}\mathcal{P}_\eta[O](x_n,\phi_n)\ \overset{N\to\infty}{\longrightarrow}\ \langle O\rangle\quad\text{(with probability 1)}\ . \tag{4.5}$$

By averaging the pattern functions of the form $\mathcal{P}_\eta[|j\rangle\langle i|]$, the matrix elements $\langle i|\rho|j\rangle$ of the state of the system are estimated. These pattern functions can be found in Ref. [57].

Here we are interested in the homodyne tomography of the joint state of two modes of radiation, which can be experimentally separately measured, so that their quadratures X_ϕ and $X'_{\phi'}$ are jointly and independently measured, yielding the set of outcomes $\{(x_n,\phi_n,x'_n,\phi'_n)\}$. It is easy to show that the pattern function of the tensor product of two operators factorizes, namely

$$\mathcal{P}[O_1\otimes O_2](x_n,\phi_n,x'_n,\phi'_n) = \mathcal{P}[O_1](x_n,\phi_n)\mathcal{P}[O_2](x'_n,\phi'_n)\ , \tag{4.6}$$

whence the matrix elements of a bipartite state R can be estimated as

$$\frac{1}{N}\sum_{n=0}^{N}\mathcal{P}_\eta[|j\rangle\langle i|](x_n,\phi_n)\,\mathcal{P}_\eta[|m\rangle\langle l|](x'_n,\phi'_n)\to\langle i|\langle l|R\,|j\rangle|m\rangle\ . \tag{4.7}$$

Another estimation strategy for homodyne tomography is the maximum likelihood one, in which the "true" state $\hat\rho$ is estimated from homodyne data

$\{(x_n, \phi_n)\}$ as the one which most likely has generated the observed data, namely the one that maximizes the likelihood functional

$$\mathcal{L}[\rho] = \sum_n \ln \mathrm{tr}[\, \rho \, \mathrm{Hom}_\eta(x_n; \phi_n)\,] \; . \tag{4.8}$$

Obviously, for finite samples the estimated state will differ from the true one, and an estimation of errors (statistical and systematic) is in order.

The maximum likelihood (ML) method is an effective method for solving more generally LININPOS (i.e. positive linear inverse) problems [58], and the present case of state estimation from homodyne data is just an example. Of course, the ML approach extends straightforwardly to the case of bipartite systems.

4.2 Homodyne tomography of a field displacement

In this first example, the input state $|\Psi\rangle\!\rangle = (1-|\xi|^2)^{\frac{1}{2}} \sum_{n=0}^{\infty} \xi^n |n\rangle|n\rangle$ is generated by parametric downconversion of the vacuum, with $\xi = [\bar{n}/(\bar{n}+1)]^{\frac{1}{2}}$, $\bar{n}$ being the average number of photons in each mode. A displacement unitary transformation $D(z) = \exp(za^\dagger - z^* a)$ is then applied to one of the two beams, thus yielding the output state

$$R_z = [D(z)\otimes I]|\Psi\rangle\!\rangle\langle\!\langle\Psi|[D^\dagger(z)\otimes I] = (1-|\xi|^2)\,|D(z)\,\xi^{a^\dagger a}\rangle\!\rangle\langle\!\langle D(z)\,\xi^{a^\dagger a}\,| \; , \tag{4.9}$$

which is then measured with two balanced homodyne setups, one each mode.

In Fig. 4.2 some results of the Monte Carlo simulation of the proposed experiment are reported. To show how this technique is effective, the matrix elements $\langle n|\langle n|R_z|0\rangle|0\rangle$ are estimated by pattern function averaging, and then an estimate of diagonal elements of the operator $D(z)$ is calculated as

$$A_{nn} = \langle n|D(z)|n\rangle = (1-|\xi|^2)^{-1/2}\xi^{-n}\frac{\langle n|\langle n|R_z|0\rangle|0\rangle}{\sqrt{\langle 0|\langle 0|R_z|0\rangle|0\rangle}} \; , \tag{4.10}$$

and compared with the theoretical value. As one can see, a meaningful reconstruction of the matrix elements of $D(z)$ can be achieved in the range $n = 0 \div 7$ with $10^6 \div 10^7$ data, with approximately $\bar{n} = 3$ thermal photons, and with quantum efficiency as low as $\eta = 0.7$. These experimental parameters correspond to those of the experiment of Ref. [56]. Improving quantum efficiency and increasing the amplifier gain (toward a maximally entangled state)

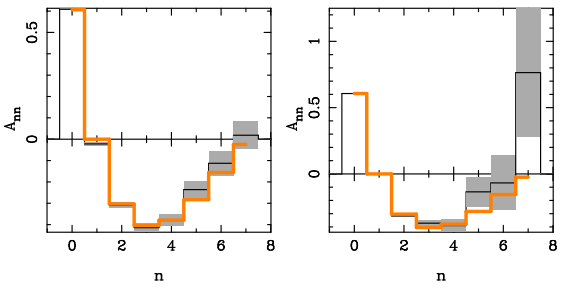

Figure 4.2: From Ref. [10]. Homodyne tomography of the displacement of one mode of the radiation field. The estimated diagonal elements A_{nn} of the displacement operator (shown by thin solid line on an extended abscissa range, with their respective error bars in gray shade) are compared to the theoretical values $\langle n|D(z)|n\rangle$ (thick solid line). Similar results are obtained for the remaining matrix elements. The reconstruction has been achieved using an entangled state $|\Psi\rangle\!\rangle$ at the input corresponding to parametric downconversion of vacuum with mean thermal photon $\bar{n}$ and quantum efficiency at homodyne detectors η. Left: $z = 1$, $\bar{n} = 5$, $\eta = 0.9$, and 1.5×10^6 data have been used. Right: $z = 1$, $\bar{n} = 3$, $\eta = 0.7$, and 6×10^7 data have been used. The last plot corresponds to the same parameters of the experiment in Ref.[56].

have both the effect of making statistical errors smaller and more uniform versus the photon labels n and m of the matrix A_{nm}.

In the experiment of Ref. [56], the relative phases between the local oscillators of the two homodyne detectors and the pump of the twin-beam were completely random and uncontrolled, and this allowed to measure only the diagonal matrix elements $\langle n|\langle m|R|n\rangle|m\rangle$ of the two mode state R, since the corresponding pattern functions are the only ones not depending on the phases. This experimental limitation is difficult but not impossible to overcome.

4.2.1 Comments on the maximum-likelihood strategy

The reconstruction can be made much more efficient by ML methods [59, 60, 61, 62, 63, 64, 65], with a reduction of the needed number of data up to a factor $100 - 1000$. Within our experimental scheme, the action of a generic quantum process $\mathcal{E}$ on one mode of the twin-beam generates the output state $R_{\mathcal{E}} = (I \otimes \Psi^T) S_{\mathcal{E}} (I \otimes \Psi^*)$ (cfr. Eq. (3.3)), $S_{\mathcal{E}}$ being the operator corresponding to the quantum process under analysis, which is positive and

satisfies $\mathrm{tr}_1 S_{\mathcal{E}} = I$. The probability distribution of the result (x, ϕ, x', ϕ') of a double homodyne detection on the two modes becomes

$$\begin{aligned} \Pr(x, \phi, x', \phi'; S_{\mathcal{E}}) &= \mathrm{tr}\left[\mathrm{Hom}_\eta(x;\phi) \otimes \mathrm{Hom}_\eta(x';\phi')\ R_{\mathcal{E}}\right] , \\ &= \mathrm{tr}\left[\mathrm{Hom}_\eta(x;\phi) \otimes (\Psi^* \mathrm{Hom}_\eta(x';\phi')\Psi^T)\ S_{\mathcal{E}}\right] \end{aligned} \quad (4.11)$$

Given a set of N double homodyne data $\{(x_n, \phi_n, x'_n, \phi'_n)\}$, the investigated quantum process can be estimated as the one whose corresponding operator $S_{\hat{\mathcal{E}}}$ maximizes the likelihood functional

$$\mathcal{L}[S_{\mathcal{E}}] = \sum_{n=0}^{N} \ln\left[\ \Pr(x_n, \phi_n, x'_n, \phi'_n; S_{\mathcal{E}})\ \right] , \quad (4.12)$$

within the simplex defined by the constraints $S_{\mathcal{E}} \geq 0$ and $\mathrm{tr}_1 S_{\mathcal{E}} = I$. If some prior knowledge about the process is available (for example, one could already know that the device performs a unitary transformation) then the maximization can be further restricted to a smaller set of candidates, thus improving further the efficiency of the estimation. In contrast to what happens with pattern averaging, here by construction the estimated map is automatically CP and trace preserving, and can fulfill any desired additional requirement.

The only downside of the ML approach is the difficulty involved in the maximization of the non linear functional in Eq. (4.12), which can be tackled either with standard techniques of numerical constrained maximization or with suitable modifications [69] of the iterative algorithms of the kind *expectation-maximization* (EM) for maximum likelihood [70, 58]. In practice, several technical problems may arise, as we will discuss concretely for the examples of the next sections.

Cramer-Rao bound, and its constrained version

As regards statistical efficiency, for the ML estimator we can assert that it is the most efficient with the following reasoning. Given a generic family of probability distributions $\Pr(x;\boldsymbol{\theta})$ depending on the independent and unconstrained parameters $\boldsymbol{\theta} \in \mathbb{R}^d$, one defines the Fisher information matrix as

$$F(\boldsymbol{\theta})_{mn} = \left\langle \frac{\partial \ln \Pr(x;\boldsymbol{\theta})}{\partial \theta_m} \frac{\partial \ln \Pr(x;\boldsymbol{\theta})}{\partial \theta_n} \right\rangle_x , \quad (4.13)$$

and for any unbiased estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$, defined on samples of N data drawn from $\Pr(x;\boldsymbol{\theta})$, one defines also the covariance matrix

$$\Sigma_{mn} = \left\langle (\hat{\theta}_m - \theta_m)(\hat{\theta}_n - \theta_n) \right\rangle_{x_1 \ldots x_N} . \quad (4.14)$$

The two matrices satisfy the Cramer-Rao bound

$$\Sigma \geq \frac{1}{N} F(\boldsymbol{\theta})^{-1} \,, \tag{4.15}$$

which puts a limit to the efficiency of the estimation which is independent on the estimator. It is possible to prove that if there exist an estimator that achieves the bound, then it coincides with the ML estimator, and that ML saturates the bound asymptotically, for increasing sample size N, when the ML estimator becomes approximately Gaussian distributed around $\boldsymbol{\theta}$, with a covariance matrix given by the so called CR matrix $F^{-1}(\boldsymbol{\theta})/N$.

When the parameters $\boldsymbol{\theta} \in \mathbb{R}^d$ are constrained to a subset $\Theta \subset \mathbb{R}^d$, the problem should be reparametrized, at least in a neighborhood of the true value $\boldsymbol{\theta}$, and the new set of independent unconstrained parameters should be then used to calculate a new Fisher information and the related CR matrix. However, this procedure is in general inelegant and difficult to use. In Ref. [66], a much more convenient way to compute the constrained CR bound was presented, based on the distinction between regular points of Θ (i.e. the points in the closure of the set of interior points of Θ) and non-regular points. As an example, for Θ defined by the constraints $0 \leq \theta_i \leq 1$, all the points are regular, whereas for Θ equal to a lower dimensional manifold embedded in $\mathbb{R}^d$ (e. g. a surface defined by some equality constraints) all points are non regular. The result is that if $\boldsymbol{\theta}$ is a regular point, then the CR matrix is unalterated, whereas if $\boldsymbol{\theta}$ is not a regular point then the CR matrix must be corrected by subtracting a positive matrix depending on $\boldsymbol{\theta}$ that makes the CR matrix smaller and singular. The singularity of the CR matrix reflects the fact that some parameters could be actually evaluated as functions of others, and thus do not have an independent associated error. A very simple derivation of equality constrained CR bound can be found in Ref. [67], along with a proof that also for constrained problems, if the bound is achieved by an estimator, then the estimate is a stationary point for the problem of maximizing the likelihood function subject to the constraints. For the problem of $k < d$ equality constraints $f_j(\boldsymbol{\theta}) = 0$, the corrected constrained CR bound becomes

$$\Sigma \geq \frac{1}{N} \left[F^{-1} - F^{-1} G (G^T F^{-1} G)^{-1} G^T F^{-1} \right] \,, \tag{4.16}$$

where G denotes the $d \times k$ matrix of the gradient of the constraints $G_{ij} = \frac{\partial f_j(\boldsymbol{\theta})}{\partial \theta_i}$.

Sieved ML for the infinite dimensional case

For the general problem of quantum process tomography, the likelihood functional $\mathcal{L}[S_{\mathcal{E}}]$ of Eq. (4.12) is defined for a parameter $S_{\mathcal{E}}$ living in an infinite dimensional Hilbert space. The maximum of the likelihood is not achieved over the whole space and it is more appropriate to restrict the attention to a subspace $\mathsf{Q}(N)$ on which the maximum exists, and to let its dimension grow with the number N of data, so to cover the whole parameter space in the limit of infinite size sample. This method—called *sieved maximum likelihood*—has been analyzed in Ref. [68] for homodyne tomography of a quantum state, with the sieves as the span of the Fock states $|0\rangle \dots |d(N)\rangle$, and the function $d(N)$ chosen in order to guarantee the consistency of the estimator, i.e. the convergence of the estimated state to the true value in the limit of infinite N.

For the particular problem at hand, because of the exponentially decreasing twin-beam components on the Fock basis, the choice of a suitable cut-off dimension will not introduce any significant bias in the estimation, and the action of the quantum channel will be reconstructed only on a finite dimensional subspace, consitently with the fact that the faithfulness of the input state rapidly vanishes for larger photon numbers.

4.3 Homodyne tomography of an On/Off detector

In what follows we exploit the ideas of Sec. (3.6) for realizing the tomography of the POVM of a measuring apparatus. One of the beams in the twin-beam state $|\Psi\rangle\!\rangle$ generated by parametric down-conversion of the vacuum (same setup as before) is now measured by an ON/OFF photo-detector. This is described by a two-value POVM, with elements $\Pi^{(0)}$ and $\Pi^{(1)} = I - \Pi^{(0)}$. As discussed in Sec. (3.6), looking at Eq. (3.30), the reduced states of the remaining beam after the measurement will be

$$\rho^{(i)} = \frac{\Psi^T \Pi^{(i)T} \Psi^*}{\mathrm{tr}[\, \Psi^T \, \Pi^{(i)T} \, \Psi^* \,]} \,, \tag{4.17}$$

i being the measurement outcome, with the denominator of the previous expression giving its probability. Homodyne tomography is then performed on each reduced state in order to recover the POVM elements.

As a model of ON/OFF detector with non unit quantum efficiency and dark current, we will use an ideal ON/OFF photodetector preceeded by a beamsplitter of transmissivity τ with one port entered by the mode of interest and

with the other port fed by a thermal radiation state with mean photon number μ [71]. The POVM element for the OFF outcome reads

$$\Pi^{(0)} = \frac{1}{\nu+1}\sum_{n=0}^{\infty}\left(1-\frac{\tau}{\nu+1}\right)^{n}|n\rangle\langle n|\,, \tag{4.18}$$

$\nu = \mu(1-\tau)$ being the resulting mean photon number of the background noise, whereas the POVM element for the ON outcome is $\Pi^{(1)} = I - \Pi^{(0)}$.

4.3.1 Reconstruction using pattern-function averaging

The graphs in Fig. 4.3 show that a meaningful reconstruction can be obtained within the same range of values for the parameters used in the tomography of the displacement. As usual, in order to achieve the reconstruction of the off-diagonal terms of the POVM, the phase-control for the local oscillator of the balanced homodyne detector relative to the pump of the down-converter is required. The presence of non-vanishing off-diagonal terms in the POVM would allow the detector to reveal some form of coherence in the input state, and in our model it could be simulated by having some coherence for the thermal radiation injected in the beam-splitter. Of course, if one already knows that the detector is perfectly phase-insensitive (as for a customary photo-detector, for its intrinsic detection mechanism), one can focus the attention only on the diagonal part of the state, without the need of phase-control for the local oscillators.

It is important to notice that when the only diagonal part of the POVM of the measuring apparatus is under examination, it is not necessary to have the input state R faithful, but is sufficient to have the matrix $R_{mn} = \langle m|\langle n|R|m\rangle|n\rangle$ invertible, with more easily experimentally available input states. In fact, in correspondence with the measurement outcome i, the diagonal matrix elements of reduced state $\rho^{(i)}$ of the auxiliary system are given by

$$\rho^{(i)}_{nn} = \frac{\sum_m R_{mn}\Pi^{(i)}_{mm}}{\mathrm{tr}[\sum_{mn} R_{mn}\Pi^{(i)}_{mm}]}\,, \tag{4.19}$$

so that, once measured $\rho^{(i)}_{nn}$, it is possible to recover $\Pi^{(i)}_{mm}$ given that R_{mn} is invertible. In summary, a "diagonally faithful" state and homodyne tomography (i.e. without phase control) is enough for the reconstruction of a diagonal POVM.

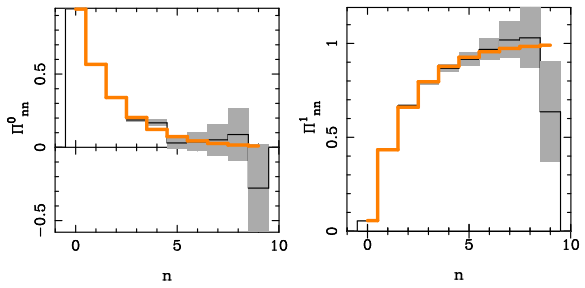

Figure 4.3: Homodyne tomography of an On/Off photo-detector having transmittivity $\tau = 0.4$ and number of thermal noise photons $\nu = 0.1$. Only the diagonal matrix elements of the POVM elements $\Pi^{(0)}$ and $\Pi^{(1)}$ are reported (the off-diagonal ones are zero, and have similar error bars). The reconstruction is obtained by pattern-function averaging of $1.5 \cdot 10^6$ data, for quantum efficiency $\eta = 0.9$ and $\bar{n} = 3$, and presents error bars of the same magnitude as the ones for the displacement reconstruction reported in Fig. 4.2

4.3.2 Reconstruction using maximum-likelihood strategy.

Now, we will analyze the data from the same experimental scheme using the maximum likelihood strategy, assuming, for simplicity, that we already know the POVM is diagonal in the Fock basis for its intrinsic detection mechanisms, such that a bipartite diagonally faithful state R and homodyne tomography without phase-control will suffice for the porpoise of reconstructing the POVM.

Non-ideal homodyne detection can be modeled as the action of the loss map followed by ideal homodyne detection, with a suitable rescaling of outcomes, such that the POVM can be written as follows

$$\mathrm{Hom}_\eta(x,\phi) = \sqrt{\eta}\sum_{j=0}^{\infty} V_j^\dagger e^{i\phi a^\dagger a}|\sqrt{\eta}x\rangle\langle\sqrt{\eta}x|e^{-i\phi a^\dagger a}V_j\ , \tag{4.20}$$

where $V_j = (\eta^{-1}-1)^{\frac{j}{2}} a^j \eta^{\frac{1}{2}a^\dagger a}/\sqrt{j!}$ are the elements of the Kraus decomposition of the loss map, and η denotes the quantum efficiency of the detectors (this scheme is equivalent to have an ideal homodyne detector preceded by a beam-splitter with transmissivity η and its second port fed with the vacuum state). If the phase is out of control and uniformly random, then the POVM corresponding to the measurement is the average over the phase of Eq. (4.20), which yields a diagonal POVM $\mathrm{Hom}_\eta(x)$ (this also makes it clear why without phase control it is impossible to reconstruct the off-diagonal matrix elements).

The probability of getting the outcome (i,x) with the photon-counter signaling outcome i and the homodyne measuring x, is given by

$$\begin{aligned}\Pr(i,x;\Pi) &= \mathrm{tr}[\, R\,\Pi^{(i)}\otimes \mathrm{Hom}_\eta(x)\,] = \sum_{mn} R_{mn}\Pi^{(i)}_{mm}\mathrm{Hom}_{\eta mn}(x) = \\ &= \sum_m \Pi^{(i)}_{mm} A_m(x)\ , \end{aligned} \tag{4.21}$$

where R is denotes input state (here the twin-beam), $R_{mn}=\langle m|\langle n|R|m\rangle|n\rangle$, and the positive coefficents $A_m(x)$ are defined as

$$A_m(x) = \sqrt{\eta}\sum_{n\geq h} R_{mn}\binom{n}{h}\eta^h(1-\eta)^{n-h}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\frac{e^{-2\eta x^2}}{2^h h!}H_h^2(\sqrt{2\eta}x) > 0\ . \tag{4.22}$$

For a given set of experimental data $\{(i_l,x_l)\}$, the maximum likelihood estimate $\hat{\Pi}$ is the one maximizing the functional

$$\mathcal{L}[\Pi] = \sum_l \ln\left[\sum_m \Pi^{(i_l)}_{mm}\cdot A_m(x_l)\right]\ , \tag{4.23}$$

with Π restricted to the simplex of diagonal POVM's.

Comments on the numerical aspects of the likelihood optimization

First, one must choose the dimension of the subspace on which performing the maximization of the likelihood and thus the estimation of the POVM elements. In such a finite dimensional subspace, the ML estimate is well defined, being the point attaining the unique maximum of a convex functional restricted to a simplex. In principle this restriction introduces a bias in the estimation, however, in our case the exponentially decreasing components of the twin-beam state in the Fock basis in practice makes the bias negligible, by making the components $\langle n|\rho^{(i)}|n\rangle$ of the reduced state after the measurement rapidly vanishing for large n.

The maximization of the functional $\mathcal{L}[\Pi]$ is a non-linear convex programming problem, and can be faced with several different kind of algorithms as the *simplex method* (see review of D'Ariano *et al.* [57]), or the methods of *sequential quadratic programming* (SQP), or the methods of *expectation-maximization* (EM) type [70, 58], whose elementary step in this particular example, for the data $\{(i_l,x_l)\}$, reduces to the upgrades

$$\Pi^{(i)}_m \leftarrow \Pi^{(i)}_m \sum_{x_l|i_l=i}\frac{A_m(x_l)}{\Pr(i,x_l;\Pi)} \tag{4.24}$$

followed by a normalization of the resulting Π to a POVM. For all methods convergence is assured, since the functional to be maximized is convex and differentiable over the simplex of diagonal POVM's. However, when applying any of these methods, the convergence speed and the reliability of the result at a given iteration step are two major concerns. In fact, the derivatives of $\mathcal{L}[\Pi]$ with respect to some of the parameters $\boldsymbol{\theta}$ defining Π can be very small, so that very different values of the parameters will give almost the same likelihood, thus making hard to judge whether the point reached at a given iteration step is a good approximation of the point corresponding to the maximum: in few words, the problem becomes numerically ill conditioned, with an extremely low convergence rate.

Notice that the Fisher information matrix (Eq. (4.13)) for the probability distribution $\Pr(i, x; \Pi)$ can be expressed in terms of the expectation value of the derivatives of the likelihood with respect to the independent parameters θ_m defining Π

$$F(\Pi)_{mn} = \frac{1}{N} \left\langle \frac{\partial \mathcal{L}[\Pi]}{\partial \theta_m} \frac{\partial \mathcal{L}[\Pi]}{\partial \theta_n} \right\rangle_{(i_1,x_1)\ldots(i_N,x_N)} , \tag{4.25}$$

so that the derivatives of the likelihood not only affect the numerical stability of the maximization, but also limit the theoretical precision of the estimation via Cramer-Rao lower bound. This bound, in turn, can be used to check whether or not the estimation is good, depending on how much the variance of the estimator is bigger than the lower bound of Eq. (4.15). This, however, needs the calculation of the Fisher information matrix in correspondence of the unknown true value of Π, and this can be approximated by the Fisher information at the estimated value, which is a reasonably good approximation provided the estimated value doesn't deviate too much from the true one.

As already mentioned, in the limit of large size samples, the ML estimator is Gaussian distributed around the true value with covariance matrix equal to $(NF)^{-1}$. Therefore for large samples the confidence levels can be assumed to be Gaussian, with variances calculated from the Fisher information, which can be evaluated on the estimated parameter for not too large errors. However, this is an asymptotic property, so that for finite size samples sometimes there is the problem of establishing the errors and the confidence levels for the estimation. When working with Monte Carlo simulation, the virtual homodyne experiment can be repeated several times, in order to evaluate the distribution of the ML estimator, and thus its confidence levels. Clearly, this approach is not

satisfying for an experimentalist, who would need to collect a lot of data only to evaluate the statistical errors for a small subset of them. A valid alternative is then provided by the method of bootstrap [72], which is based on the simple idea that when some data are drawn from an unknown probability distribution, then the distribution of those data is the best approximation we have of the real probability distribution. Thus, once the experimental data are collected, we can perform the ML estimation on artificial samples repeatedly generated by random sampling the experimental data: the distribution of such estimates approximates well the one that we would have from the real experiment, and can be used to evaluate the confidence levels for the estimator.

Here we will perform the numerical maximization by means the routine `donlp2` [73] which implements an SQP algorithm, and then the self-consistency of the solution will be checked by means of a few EM type iterations. Of course, it would be much easier to implement the only EM algorithm, being a recursive application of the easily implementable step of Eq. (4.24). However, this algorithm has an extremely low convergence speed, which also could make the iteration stop too early, leading to (statistically wrong!) results — which may even fit too well the theoretical POVM when this is a particularly smooth function of the photon number.

Results for the On/Off detector

Back to our problem of On/Off detector tomography, we have produced a Monte Carlo simulation of the joint homodyne and on/off data distributed according to Eq. (4.21), for the POVM model presented in Eq. (4.18), with the same parameters as Fig. 4.3, and various values of the quantum efficiency η. The detector POVM has been estimated with maximum likelihood method, with the only hypotesis of diagonal POVM, and putting the dimesnional cut-off at the first 15 elements of the Fock basis (for a number of photons in the twin-beam equal to $\bar{n} = 3$ this introduces almost no bias, with an actual suppression of a factor 100 between the first diagonal element of the POVM and the first excluded element). The results are reported in Fig. 4.4 for different sample sizes and quantum efficiencies, where the only "Off" element of the POVM is reported, since the "On" element is simply its complement with respect to the identity.

A direct comparison with Fig. 4.3 evidences the much higher efficency of maximum likelihood reconstruction. The graph on the left shows how the

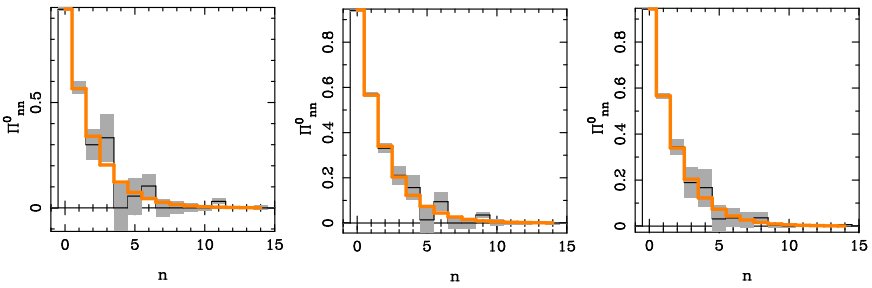

Figure 4.4: Homodyne tomography of an On/Off photodetector with transmittivity $\tau = 0.4$ and number of thermal noise photons $\nu = 0.1$, with $\bar{n} = 3$ photons in the twin-beam. The ML estimation of the diagonal of the only Off POVM element are reported for different values of sample size N and quantum efficiency η. Left: $N = 10^5$, $\eta = 0.7$; Middle: $N = 10^4$, $\eta = 0.9$; Right: $N = 10^6$, $\eta = 0.7$.

same magnitude of errors is achieved on a larger subspace with less than one tenth of the data (10^5 vs. 10^6) and with a much lower quantum efficiency (0.7 vs. 0.9). For the same quantum efficiency $\eta = 0.9$, here the results are much better even with as few as 1% of the data (graph in the middle), analogously, for the same amount of data ($N = 10^6$), here the results are much better even for a quantum efficiency as low as 0.7 (graph on the right).

The distribution of the estimator in each bin, which is necessary for giving proper confidence levels for the result, has been evaluated by repeated Monte Carlo experiments, which is equivalent to the bootstrapping techniques for truly experimental data. As a result the estimator in each bin is not Gaussian distributed, a sign of the fact that the number of data used is not enough to reach the asymptotic Gaussian distribution of the ML estimator. In the plot, the only variances are reported for each bin, showing that the errors are distributed with respect to n differently than for pattern averaging.

4.4 Homodyne tomography of a photon-counter

Now we will implement maximum likelihood techniques for the reconstruction of a photon-counting detector modeled as an ideal photo-detector preceeded by a beam splitter of transmittivity τ fed with the measured mode and with a thermal state, with mean photon number ν, simulating the thermal noise of

the apparatus. The POVM of such an apparatus reads

$$\Pi^{(i)} = \sum_{n,m=0}^{\infty} \sum_{k=\max(0,i-m)}^{\min(n,i)} \binom{n}{k} \binom{m}{i-k} \tau^{m-i+2k} (1-\tau)^{n+i-2k} \frac{\lambda^m}{1-\lambda} |n\rangle\langle n| \; . \tag{4.26}$$

Being it impossible to achieve a meaningful reconstruction for any i, neither on a small subspace, one should focus the attention only on the first I elements, and sum up all the remaining ones into a last fictious POVM element. The likelihood functional has the form in Eq. (4.23), and it can be maximized with the same technique of the preceding example.

In Fig. 4.5 the results obtained by Monte Carlo simulation are reported. The first four elements of the POVM have been considered, and all the others have been "concentrated" into the fifth element. The model has been simulated for transmittivity $\tau = 0.9$ and number of thermal noise photons $\nu = 0.05$, with $\bar{n} = 3$ and $\eta = 0.7$. The plots correspond to a reconstruction obtained with $N = 10^5$ data. Here the number of data necessary for a very good reconstruction is much lower than the values of the On/Off photodetector example, because of the very small overlap of the elements $\Pi^{(i)}$ of the POVM that leads to a much larger Fisher information.

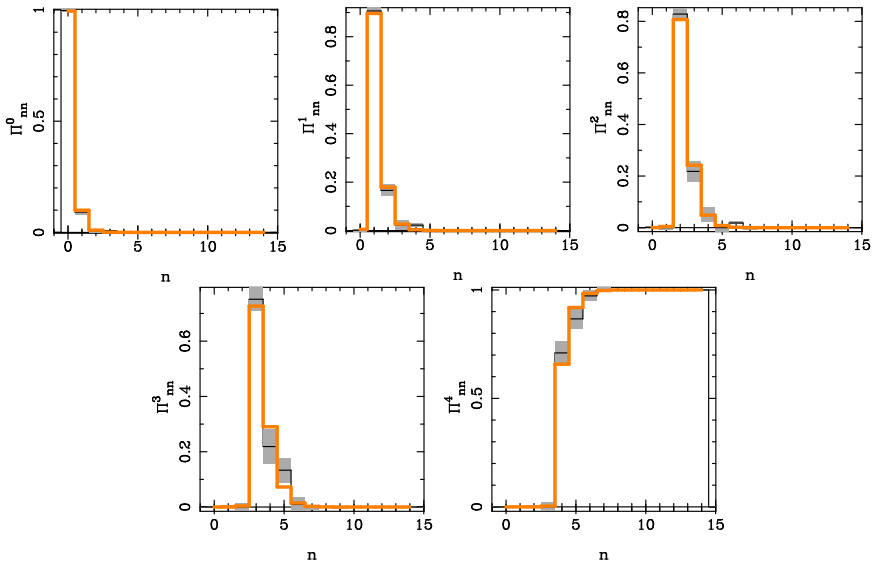

Figure 4.5: Homodyne tomography of an photon-counting detector with transmittivity $\tau = 0.9$ and number of thermal noise photons $\nu = 0.05$, with $\bar{n} = 3$ and $\eta = 0.7$, obtained with $N = 10^5$ data.

Chapter 5

Measurements improved by entanglement

Entanglement is certainly the most distinctive feature of quantum mechanics. The quantum nonlocality due to entanglement, which has puzzled generations of theoreticians since the work of Einstein, Podolsky, and Rosen [74], in the last decade eventually has been harnessed for practical use in the new quantum information technology [75, 48]. Entanglement has become the essential resource for quantum computing, quantum teleportation, and secure cryptographic protocols [48]. Recently, entanglement has been proved as a valuable resource for improving optical resolution [76], spectroscopy [77], and has been shown to be a crucial ingredient for making the tomography of a quantum device [10], with a single input entangled state playing the role of all possible states at the input of the device — another manifestation of the *quantum parallelism*, the feature of entanglement that is the core of quantum computing algorithms [78, 79].

In this chapter we will report our results in Ref. [20] showing how in general entanglement can be used to improve quantum measurements, for either precision or stability. The measurement scheme will be considered in the general framework of quantum estimation theory [80], in which one needs to estimate the parameter θ of the density operator ρ_θ on the Hilbert space H being the result of a unitary transformation $\rho \to \rho_\theta = U_\theta \rho U_\theta^\dagger$. More generally a quantum operation Q_θ could be considered, with $\rho_\theta = \mathrm{Q}_\theta(\rho)$, θ corresponding to a parameter of any physical (amplifying, measuring, etc.) device. In other words, the measurements we are going to consider are aimed at the

discrimination of a family of quantum devices parametrized by θ.

This situation for known input state ρ is very common in practice, e. g. in interferometry [81], and more generally whenever the measurement is *indirect*, resorting to the detection of a change in an ancillary part of the measuring apparatus. In this scenario we will consider the use of an entangled input state R in place of ρ, with the unknown transformation U_θ acting locally only on one side of the entangled state. In tensor notation: $R \to R_\theta = (U_\theta \otimes I)\, R\, (U_\theta^\dagger \otimes I)$. The situation is depicted in Fig. 5.1. As we will see in this section, the entangled configuration is better than the conventional one, for either precision or stability of the measurement. This is due to the fact that, in some sense, the input entangled state is equivalent to many input states in "quantum parallel". In the following we will examine different measurement situations separately, and we will draw general conclusions at the end.

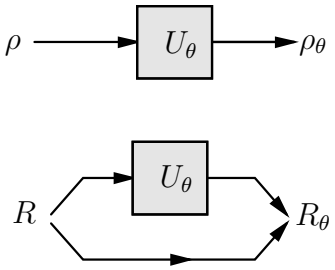

Figure 5.1: Measurement schemes considered in the present chapter. The parameter θ of the density operator ρ_θ is estimated as the result of a unitary transformation $\rho \to \rho_\theta = U_\theta \rho U_\theta^\dagger$ (top). In this scenario the use of an entangled input R in place of ρ is considered, with the unknown transformation U_θ acting locally on one Hilbert space only (bottom).

5.1 Covariant measurements

In a covariant measurement the parameter θ is the element $g \in \mathbf{G}$ of a group $\mathbf{G}$ of transformations. This kind of measurement has been thoroughly analyzed in Ref. [23].

Let us first illustrate the mechanism of entanglement on a simple example. We want to discriminate among the four unitary transformations represented by the Pauli matrices $\sigma_0 \equiv I\,, \sigma_1 \equiv \sigma_x\,, \sigma_2 \equiv \sigma_y\,, \sigma_3 \equiv \sigma_z$. As well known,

they form a unitary discrete group, as they form a projective nonabelian irreducible representation of the (abelian) dihedral group D_2 of π rotations in three dimensions. By applying the four transformations to any single-qubit input state $|\psi\rangle \in \mathbb{C}^2$ we always obtain four linearly dependent states, which makes the conventional scheme in Fig. 5.1 useless for a reliable discrimination. On the contrary, if we apply the four matrices to the maximally entangled input state $\frac{1}{\sqrt{2}}|I\rangle\!\rangle$ we obtain the four Bell states $\sigma_j \otimes I \frac{1}{\sqrt{2}}|I\rangle\!\rangle \equiv \frac{1}{\sqrt{2}}|\sigma_j\rangle\!\rangle$, which are mutually orthogonal.

This simple example is very instructive: the discrimination among the four Pauli transformations σ_j, which is impossible with a single qubit input state, becomes possible and exact when applying σ_j to a maximally entangled state. The mechanism is clear: using an entangled state instead of a single qubit, doubles the dimension of the Hilbert space H_{out} spanned by the output states, allowing perfect discrimination of the four σ_j.

The above example can be generalized easily to any dimension d, when discriminating among the d^2 unitary transformations

$$U(m,n) = \sum_{k=0}^{d-1} e^{2\pi i k m/d} |k\rangle\langle k \oplus n| \;, \tag{5.1}$$

n and m ranging in $0 \div d-1$, and $\oplus$ denoting addition modulo d. The unitary operators $U(m,n)$ form a projective nonabelian irreducible representation of the (abelian) group $\mathbb{Z}_d \times \mathbb{Z}_d$ describing translations on a two-dimensional lattice embedded in a torus. The dihedral group D_2 corresponds to the particular case $d = 2$. Now, using the maximally entagled state $\frac{1}{\sqrt{d}}|I\rangle\!\rangle$ at the input will produce the d^2 orthogonal output states $U(m,n) \otimes \frac{1}{\sqrt{d}}|I\rangle\!\rangle$, which allows perfect discrimination among all $U(m,n)$, whereas a non entangled input $|\psi\rangle \in \mathsf{H}$ would output d^2 linearly dependent states in the d-dimensional H.

More generally, let us consider a set of unitary transformations $\{U_g\}$, $g \in \mathbf{G}$ that form a (projective) representation of the group $\mathbf{G}$. For simplicity let us consider the case of an irreducible representation (the reducible case is technically more complicate, and needs the knowledge of all irreducible components on invariant subspaces). For every operator O on H, from the Schur's lemma one has the trace identity

$$[U_g \, O \, U_g^\dagger]_{\mathbf{G}} = \mathrm{tr}[O] I \;, \tag{5.2}$$

where $[f(g)]_{\mathbf{G}}$ denotes the group averaging $[f(g)]_{\mathbf{G}} \doteq \sum_{g \in \mathbf{G}} \mu(g) f(g)$ with $\mu(g) = \frac{d}{|\mathbf{G}|}$, $|\mathbf{G}|$ the cardinality of $\mathbf{G}$, and $d = \dim \mathsf{H}$. Eq. (5.2) generalizes

to the continuous case for group averaging defined as $[f(g)]_{\mathbf{G}} \doteq \int_{\mathbf{G}} \mu(\mathrm{d}g) f(g)$, $\mu(\mathrm{d}g)$ being a (normalized) invariant measure on $\mathbf{G}$.

In order to show that entanglement is of help in improving the discrimination, and to quantify this improvement, we now consider several parameters, by supposing the input is a generic bipartite state $|E\rangle\!\rangle$ that can be either entangled or not. First of all, as in the first two examples, one can see that the dimension of the Hilbert space H_{out} spanned by the output states is larger for an entangled input than for factorized states. In fact, $\dim(\mathsf{H}_{out})$ can be calculated as the rank of the operator

$$O = \Big[|\Psi_g\rangle\!\rangle\langle\!\langle \Psi_g| \Big]_{\mathbf{G}} = \Big[U_g \otimes I |E\rangle\!\rangle\langle\!\langle E| U_g^\dagger \otimes I \Big]_{\mathbf{G}} \,, \tag{5.3}$$

where $\Psi_g = U_g E$. By means of Eq. (5.2) one has $O = I \otimes \mathrm{Tr}_1[|E\rangle\!\rangle\langle\!\langle E|] = I \otimes (E^\dagger E)^T$, so that

$$\dim(\mathsf{H}_{out}) = d \times \mathrm{rank}(E^\dagger E) \,, \tag{5.4}$$

i.e. the output space is enlarged by a factor equal to the Schmidt number [48] of the input state. Indeed, since probing the operation with a bipartite entangled system gives access to a larger Hilbert space we have, literally, more room for improvement. In the following, we refine these concepts, and give conditions under which and entangled scheme is convenient.

The Schmidt number is only a coarse measure of the amount of entanglement stored in $|E\rangle\!\rangle$, and the dimension of the output space is only indirectly connected to the distinguishability of the outputs. A more refined goodness criterion is given by the Holevo's information χ of the set of output states, all taken with the same probability $p(g) = 1/|\mathbf{G}|$ (or $p(\mathrm{d}g) = \mu(\mathrm{d}g)/\mu(\mathbf{G})$ in the continuous case), this quantity is an upper bound for the accessible information [48]. Denoting by $S(\rho) = -\operatorname{tr} \rho \log \rho$, the von Neumann entropy of ρ, the Holevo's information χ reads

$$\begin{aligned} \chi &= S\left(\frac{1}{\mu(\mathbf{G})} \Big[|\Psi_g\rangle\!\rangle\langle\!\langle \Psi_g| \Big]_{\mathbf{G}} \right) - \frac{1}{\mu(\mathbf{G})} \Big[S(|\Psi_g\rangle\!\rangle\langle\!\langle \Psi_g|) \Big]_{\mathbf{G}} = \\ &= S\left(\frac{1}{\mu(\mathbf{G})} I \otimes E^T E^* \right) = \\ &= \frac{d}{\mu(\mathbf{G})} \log \mu(\mathbf{G}) + \frac{d}{\mu(\mathbf{G})} S(E^T E^*) \,, \end{aligned} \tag{5.5}$$

and thus the bound is increased by an amount proportional to the degree of

entanglement $S(E^T E^*)$[1] of the input state $|E\rangle\!\rangle$ (recall that for discrete groups $\mu(\mathbf{G}) = d$).

Facing the problem with a maximum likelihood strategy, the optimal covariant POVM that discriminates among the $\{|\Psi_g\rangle\!\rangle\}$ takes the form [23]

$$\Pi_g = \mu(g)(U_g \otimes I)P(U_g^\dagger \otimes I) \;, \tag{5.6}$$

with $P \geq 0$ a positive operator on $\mathsf{H} \otimes \mathsf{H}$ normalized as $\mathrm{Tr}_1[P] = I$. By covariance, the likelihood – i.e. the probability of getting an outcome g when the state is $|\Psi_g\rangle\!\rangle$ – is proportional to $\langle\!\langle E|P|E\rangle\!\rangle \leq d$, where the bound comes from the normalization condition on P, which limits the largest possible eigenvalue of P below d. Again, the optimality (saturation of the bound) is reached for a maximally entangled input state, i.e. for $E = d^{-\frac{1}{2}}U$, with U unitary, and $P = |U\rangle\!\rangle\langle\!\langle U|$. The optimality of a maximally entangled input state for the estimation of unitaries in $SU(d)$ has also been noticed in Ref. [82].

Since the overlap of two states is the only parameter that determines their distinguishability, we will consider the average overlap $\Omega(E)$ of all the couples of states in $\{|\Psi_g\rangle\!\rangle\}$: the lower is $\Omega(E)$ the better will be the overall distinguishability. One has

$$\begin{aligned}
\Omega(E) &= \frac{1}{2\mu(\mathbf{G})^2}\Big[|\langle\!\langle \Psi_g|\Psi_{g'}\rangle\!\rangle|^2\Big]_{\mathbf{G}\times\mathbf{G}} = \frac{1}{2\mu(\mathbf{G})}\Big[\langle\!\langle E|\Psi_g\rangle\!\rangle\langle\!\langle \Psi_g|E\rangle\!\rangle\Big]_{\mathbf{G}} = \\
&= \frac{1}{2\mu(\mathbf{G})}\langle\!\langle E|I \otimes (E^T E^*)|E\rangle\!\rangle = \frac{1}{2\mu(\mathbf{G})}\langle\!\langle E|EE^\dagger E\rangle\!\rangle = \\
&= \frac{1}{2\mu(\mathbf{G})}\,\mathrm{tr}[(E^\dagger E)^2] \;.
\end{aligned} \tag{5.7}$$

In order to analyze the properties of $\Omega(E)$, we have to briefly recall the definition of the "majorization" relation between entangled pure states and its physical meaning. Given two states $|A\rangle\!\rangle$ and $|B\rangle\!\rangle$ in $\mathsf{H} \otimes \mathsf{H}$, let $\lambda_A^\downarrow$ and $\lambda_B^\downarrow$ be the vectors of eigenvalues of $A^\dagger A$ and $B^\dagger B$ respectively, sorted in descending order. We say that $|A\rangle\!\rangle \prec |B\rangle\!\rangle$ iff

$$\sum_{j=1}^{k}(\lambda_A^\downarrow)_j \leq \sum_{j=1}^{k}(\lambda_B^\downarrow)_j \;, \quad \text{for each } k \leq d \;. \tag{5.8}$$

The physical meaning of this partial ordering relation has been clarified in Ref. [83]: $|A\rangle\!\rangle$ can be transformed into $|B\rangle\!\rangle$ by local operations and classical communication if and only if $|A\rangle\!\rangle \prec |B\rangle\!\rangle$.

[1] $S(E^T E^*)$ represents the entropy of the partial traces of $|E\rangle\!\rangle$, which indeed is the measure of entanglement for pure states.

Our average overlap $\Omega(E)$ is a so called "Schur convex function" of the eigenvalues of $E^\dagger E$, namely if $|A\rangle\!\rangle \prec |B\rangle\!\rangle$ then $\Omega(A) \leq \Omega(B)$. Since any maximally entangled state is majorized by any other state, it is clear that the minimum overlap is found in correspondence with $|E\rangle\!\rangle$ maximally entangled, and any manipulation of such a state can only increase $\Omega(E)$, thus reducing the distinguishability, and, as a consequence, the sensitivity of the measurement.

As an example in infinite dimensions, consider the problem of estimating the displacement of a harmonic oscillator in the phase space, i. e. the parameter $\alpha \in \mathbb{C}$ of the transformation $\rho \to \rho_\alpha = D(\alpha)\rho D^\dagger(\alpha)$, where $D(\alpha) = \exp(\alpha a^\dagger - \overline{\alpha} a)$ is the displacement operator for annihilation and creation operators a and $a^\dagger$ respectively (in this case $\mathbf{G}$ is the Weyl-Heisenberg group). For unentangled ρ, an estimation of α isotropic on $\mathbb{C}$ is equivalent to a optimal joint measurement of position and momentum, which, as well known, is affected by a unavoidable minimum noise of 3dB [84]. Here, the optimal state (for fixed minimum energy) is the vacuum, and the corresponding conditional probability of measuring z given α is $p(z|\alpha) = \pi^{-1} \exp[-|z-\alpha|^2]$. Now, consider the case in which the estimation is made with $D(\alpha)$ acting on the entangled state

$$|E\rangle\!\rangle = \sqrt{1-|x|^2} \sum_{n=0}^{\infty} x^n |n\rangle|n\rangle \,, \tag{5.9}$$

with $|x| \leq 1$ (the state (5.9) can be achieved by parametric downconversion of vacuum). Here, we can use the orthonormal resolution of the identity $|D(z)\rangle\!\rangle\langle\!\langle D(z)|$ of eigenvectors $|D(z)\rangle\!\rangle$ of $Z = a \otimes I - I \otimes a^\dagger$ with eigenvalue z (this is just a heterodyne measurement [85]), now achieving $p(z|\alpha) = (\pi\Delta^2)^{-1} \exp[-\Delta^{-2}|z-\alpha|^2]$, with variance $\Delta^2 = \frac{1-|x|}{1+|x|}$ that, in principle, can be decreased at will with the state (5.9) approaching a state an eigenstate of Z (by increasing the gain of the downconverter).

5.2 Measurement in the presence of noise

What happens if the estimation is performed in the presence of noise, namely the channel before and after the unknown transformation is affected by noise? Here it is instructive to reconsider the problem of estimating the displacement of a harmonic oscillator in the phase space in the presence of Gaussian

displacement noise, which maps states as follows

$$\rho \to \Gamma_{\overline{n}}(\rho) \doteq \int_{\mathbb{C}} \frac{\mathrm{d}^2\gamma}{\pi\overline{n}} \exp[-|\gamma|^2/\overline{n}] D(\gamma)\rho D^\dagger(\gamma) \ . \tag{5.10}$$

The variance $\overline{n}$ of the noise is usually referred to as "mean thermal photon number". The case of Gaussian displacement noise is particularly simple, since one has the composition law $\Gamma_{\overline{n}} \circ \Gamma_{\overline{m}} = \Gamma_{\overline{n}+\overline{m}}$, and, moreover $\Gamma_{\overline{n}}[D(\alpha)\rho D^\dagger(\alpha)] = D(\alpha)\Gamma_{\overline{n}}(\rho)D^\dagger(\alpha)$. Therefore, if the measurement is made on the entangled state (5.9), one can easily derive a a Gaussian conditional probability distribution with variance $\delta^2 = \Delta^2 + 2\overline{n}_T$, where $\overline{n}_T$ is the total Gaussian displacement noise before and after the displacement $D(\alpha)$, and the noise is doubled since it is supposed equal on the two entangled Hilbert spaces. On the other hand, in the measurement scheme with unentangled input (remind that the optimal is the vacuum), one has $\delta^2 = 1 + \overline{n}_T$. One concludes that the entangled input is no longer convenient above one thermal photon $\overline{n}_T = 1$ of noise. This is exactly the threshold of noise above which the entanglement is totally degraded to a separable state [53], and therefore the quantum capacity of the noisy channel vanishes [86].

5.3 Discrimination between two unitaries

Let us suppose that we have to distinguish among two unitaries U_1 and U_2. Given an input state $|\psi\rangle$, one optimizes over the possible measurements, and the minimum error probability in discriminating $U_1|\psi\rangle$ and $U_1|\psi\rangle$ [80] is given by

$$P_E = \frac{1}{2}\left[1 - \sqrt{1 - |\langle\psi|U_2^\dagger U_1|\psi\rangle|^2}\right] \ , \tag{5.11}$$

so that one has to minimize the overlap $|\langle\psi|U_2^\dagger U_1|\psi\rangle|$ with a suitable choice of $|\psi\rangle$. Chosing as a basis the eigenvectors $\{|j\rangle\}$ of $U_2^\dagger U_1$, and writing $|\psi\rangle = \sum_j \psi_j|j\rangle$, we define

$$z_\psi \doteq \langle\psi|U_2^\dagger U_1|\psi\rangle = \sum_j |\psi_j|^2 e^{i\gamma_j} \ , \tag{5.12}$$

where $e^{i\gamma_j}$ are the eigenvalues of $U_2^\dagger U_1$. The normalization condition for $|\psi\rangle$ is $\sum_j |\psi_j|^2 = 1$, so that the subset $K(U_2^\dagger U_1) \subset \mathbb{C}$ described by z_ψ for varying

$|\psi\rangle$ is the convex polygon having the points $e^{i\gamma_j}$ as vertices. The minimum overlap

$$r(U_2^\dagger U_1) \doteq \min_{||\psi||=1} |\langle\psi|U_2^\dagger U_1|\psi\rangle| \tag{5.13}$$

is the distance of $K(U_2^\dagger U_1)$ from $z = 0$. This geometrical picture indicates in a simple way what is the best one can do in discriminating U_1 and U_2: if K contains the origin then the two unitaries can be exactly discriminated, otherwise one has to find the point of K nearest to the origin, and the minimum probability of error is related to its distance from the origin. Once the optimal point in K is found, the optimal states ψ are those corresponding that point through Eq. (5.12).

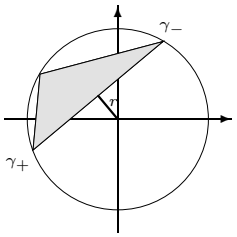

Figure 5.2: r is the minimum distance between the origin and the polygon K

If $\Delta(U_2^\dagger U_1)$ is the angular spread of the eigenvalues of $U_2^\dagger U_1$ (referring to Fig. 5.2, it is $\Delta = \gamma_+ - \gamma_-$), from Eq. (5.11) for $\Delta < \pi$ one has

$$P_E = \frac{1}{2} - \frac{1}{2}\sqrt{1 - \cos^4\frac{\Delta}{2}}\,, \tag{5.14}$$

whereas for $\Delta \geq \pi$ one has $P_E = 0$ and the discrimination is exact.

Given U_1 and U_2 non exactly discriminable, one is interested in understanding wheter or not an entangled input state could be of some use. The answer is negative, in fact using entanglement translates the problem into the one of distinguishing between $U_1 \otimes I$ and $U_2 \otimes I$, thus one has to analyze of the polygon $K(U_2^\dagger U_1 \otimes I)$. Since $U_2^\dagger U_1 \otimes I$ has the same eigenvalues as $U_2^\dagger U_1$, the polygons $K(U_2^\dagger U_1 \otimes I)$ and $K(U_2^\dagger U_1)$ are exactly the same, so that they lead to the same minimum probability of errror.

The situation changes dramatically if N copies of the unitary transformation are used, as depicted in Fig. 5.3: here one has to compare the "performance" of $K(U_2^\dagger U_1)$ to the one of $K((U_2^\dagger U_1)^{\otimes N})$. Since $\Delta((U_2^\dagger U_1)^{\otimes N}) = \min\{N \times \Delta(U_2^\dagger U_1), 2\pi\}$, it is clear that there will be an $\bar{N}$ such that $U_1^{\otimes N}$ and

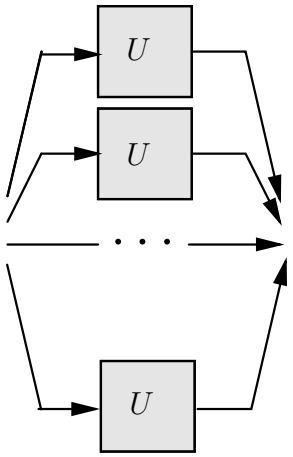

Figure 5.3: When distinguishing between two unitaries $U = U_{1,2}$ it is possible to achieve perfect discrimination even for nonorthogonal U_1 and U_2 for sufficiently large number N of copies of the unitary transformation, using a N-partite entangled state as in figure (see text).

$U_2^{\otimes N}$ will be exactly discriminable. This same result has been demonstrated in Ref. [87] starting from a different approach.

5.4 Improving the stability of the measurement

In the istances in which the optimal discrimination between transformations is already optimized by a unentangled input, an entangled state can still be better in achieving a more stable sensitivity. We have seen that a unentangled input is already optimal in the discrimination of (one use of) two unitaries. A unentangled input is also optimal in the covariant measurement for abelian $\mathbb{G}$, since the irreducible representations are one dimensional. Consider, for example, the problem of distinguishing among displacements on a fixed direction of the phase space, say $D(x)$, with $x \in \mathbb{R}$. In this case one could use a squeezed state $|x_0\rangle_s \doteq \exp[\frac{s}{2}((a^\dagger)^2 - a^2)]D(x_0)|0\rangle$, with $s > 0$, i. e. squeezed in the direction of the "quadrature" $X = \frac{1}{2}(a^\dagger + a)$. Then, a conditional Gaussian probability with variance $\langle \Delta X^2 \rangle = \frac{1}{4}e^{-2s}$ is obtained, which can be narrowed at will by using $n_s = \sinh^2 s$ squeezing photons. However, if the phase of the quadrature is slightly mismatched, and the quadrature $X_\phi = \frac{1}{2}(a^\dagger e^{i\phi} + a e^{-i\phi})$ is measured instead, then the variance becomes $\langle \Delta X_\phi^2 \rangle = \frac{1}{4}(e^{2s}\sin^2\phi + e^{-2s}\cos^2\phi)$,

and the sensitivity is exponentially unstable. Using the entangled input in Eq. (5.9), instead, gives the same Gaussian noise $\Delta^2 = \frac{1-|x|}{1+|x|}$, independently on ϕ, by using $n = 2|x|^2/(1-|x|^2)$ downconverted photons.

5.5 Further generalizations and remarks

Up to now we have focused our analysis only on discrimination among unitaries, however, we could have considered more generally nonunitary quantum operations, to see that entanglement is still a useful resource for improving the measurement. For the case of two operations Q_1 and Q_2 the distinguishability is related to the completely bounded (cb) norm [86] $||p_1\mathrm{Q}_1 - p_2\mathrm{Q}_2||_{cb}$ which is the supremum over all possible entangled input states of the trace-distance between the output states. Since the cb-norm is equivalent to the usual trace-norm for completely positive maps, it follows that a unentangled state already achieves optimality in the special case that the difference $p_1\mathrm{Q}_1 - p_2\mathrm{Q}_2$ is completely positive.

In conclusion, we have seen that entanglement is a useful resource for upgrading the quantum measurements which are based on the estimation of a quantum transformation. It is always of benefit, in improving either precision or stability. In many cases the measurement precision becomes in principle unbounded, even when the conventional measurement is noise limited. The upgrading is effective in the presence of noise, below the threshold of total entanglement degradation.

Conclusions

In this thesis, we have exploited the representation of quantum devices as suitably chosen operators, which has proven to be a powerful tool providing the necessary insight to derive either exquisitely theoretical results, or practical ones, which could have a technological impact in the future for the characterization and calibration of quantum devices.

In Chapter 1 we introduced some known results in an original way, showing how the transformations operated by a quantum devices (quantum operations, QO's) are completely positive maps (CP maps), and presenting some ways to represent these maps. We used the representation of a CP map as a positive operator to easily derive its Krauss decomposition, and its realization as a customary unitary transformation on an enlarged system, followed by a projective measurement. Then we reviewed the concept of positive operator valued measure (POVM) for describing the statistics of the outcomes of a quantum measurement, and the connection between the elements of the POVM and the operators describing the state reduction transformation was reported.

In Chapter 2 we characterized those POVM's whose statistics of the outcomes is affected by an intrinsic noise of purely quantum nature, they are the extremal points of the convex set of POVM's, and we showed how an element of that set can be decomposed as a mixing of pure ones. This result was extended to quantum devices, for which the added noise affects also the statistics of the outcomes of any subsequent measurement carried on the reduced states. Then we classified covariant quantum operations, i.e. physical transformations propagating the action of a unitary representation of a group from the input to the output: the covariance requirement on the transformation becomes an invariance requirement for the corresponding positive operator, and this last property leads – through Schur's lemma – to a peculiar block

structure for such operator. In this way, all the quantum operations which are covariant with respect to a given group can be parametrized. The techniques of extremal point characterization previously employed were also applied to covariant QO's. Quantum cloning was interpreted as a permutation covariant QO, and the parametrization previously explained was used to calculate some examples of optimal phase-covariant cloning.

Chapter 3 was devoted to the classification of all the sets of input states which support the imprinting of the information about a quantum device — i.e. faithful sets of states. This class of states represents all the possible inputs that can be used for the experimental characterization of quantum devices. We showed that a fixed entangled probe can be used for a full reconstruction of a quantum device, a fact with no classical analog and with important practical consequences. Then we developed the theory of faithful states, showing how faithfulness of a state is equivalent to the invertibility of a CP map related to that state. As a consequence, we proved the existence of separable faithful states, thus clarifying that entanglement is not strictly necessary. A measure of faithfulness was proposed as an indicator of the performance of the employed state in providing a sensitive reconstruction of quantum operations, and it turned out that this measure coincides with the purity of the state, so that maximally entangled states are actually the best faithful states. The concept of faithfulness was then extended to set of states, and a procedure for "patching" unfaithful states into faithful ones was presented, thus achieving a complete classification of the input which are suitable for devices characterization.

In Chapter 4 we presented an experimental setup for homodyne tomography of devices in the realm of quantum optics, where our technique is the only feasible one. We analyzed the sensitivity of the method for several examples of devices — a field displacement, an on/off photodetector, and a photon-counter — showing that by means of a maximum likelihood strategy a very good reconstruction of the device can be obtained from homodyne data even with a low quantum efficiency and a small number of data, with all the parameters in the range given by a realistic experimental situation. Our results attest this technique as a promising tool for devices characterization and above all for photo-detectors calibration.

In the last Chapter, we exploited entanglement in improving either the precision or the stability of those quantum measurements which resort to the discrimination of quantum operations. First, we analyzed the case of dis-

crimination of the elements of a unitary irreducible representation of a group, showing how entanglement improves several figures of merit — i.e. the dimension of the space spanned by the outputs, the Holevo information, the average likelihood, the inverse of the average overlap of the outputs. Then we showed by two examples how entanglement helps in presence of noise or in the case of miscalibration of the measuring apparatus. As a final remarkable result, we showed how an unknown transformation picked up from a finite set of known unitaries can be perfectly determined with a single measurement, by means of a multiparty entangled probe and a finite number of uses the unknown transformation.

Bibliography

[1] W. F. Stinespring, Positive functions on C^*-algebras, Proc. Am. Math. Soc. **6**, 211–216 (1955).

[2] K. Kraus, General State Changes in Quantum Theory, Ann. Phys. **64**, 311–335 (1971).

[3] I. L. Chuang and M. A. Nielsen, Prescription for experimental determination of the dynamics of a quantum black box, J. Mod. Opt. **44**, 2455–2467 (1997).

[4] J. F. Poyatos, J. I. Cirac, and P. Zoller, Complete Characterization of a Quantum Process: The Two-Bit Quantum Gate, Phys. Rev. Lett. **78**, 390–393 (1997).

[5] M. A. Nielsen, E. Knill, and R. Laflamme, Complete quantum teleportation using nuclear magnetic resonance, Nature **396**, 52–55 (1998).

[6] A. M. Childs, I. L. Chuang, and D. W. Leung, Realization of quantum process tomography in NMR, Phys. Rev. A **64**, 012314 (2001).

[7] N. Boulant, T. F. Havel, M. A. Pravia, and D. G. Cory, Robust method for estimating the Lindblad operators of a dissipative quantum process from measurements of the density operator at multiple time points, Phys. Rev. A **67**, 042322 (2003).

[8] P. Kwiat, J. Altepeter, D. Branning, E. Jeffrey, N. Peters, and T. Weh, Taming Entanglement, in *Proceedings of the 6th International Conference on Quantum Communications, Measurement and Computing*, edited by J. Shapiro and O. Hirota, page 117, Princeton, NJ, 2003, Rinton Press.

[9] M. W. Mitchell, C. W. Ellenor, S. Schneider, and A. M. Steinberg, Diagnosis, Prescription, and Prognosis of a Bell-State Filter by Quantum Process Tomography, Phys. Rev. Lett. **91**, 120402 (2003).

[10] G. M. D'Ariano and P. Lo Presti, Quantum tomography for measuring experimentally the matrix elements of an arbitrary quantum operation, Phys. Rev. Lett. **86**, 4195 (2001).

[11] A. Jamiolkowski, Linear transformations which preserve trace and positive semidefiteness of operators, Rep. Math. Phys. **3**, 275 (1972).

[12] M.-D. Choi, Completely positive linear maps on complex matrices, Linear Algebra Appl. **10**, 285–290 (1975).

[13] F. D. Martini, A. Mazzei, M. Ricci, and G. M. D'Ariano, Exploiting quantum parallelism of entanglement for a complete experimental quantum characterization of a single-qubit device, Phys. Rev. A **67**, 062307 (2003).

[14] J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, P. G. Kwiat, R. T. Thew, J. L. O'Brien, M. A. Nielsen, and A. G. White, Ancilla-Assisted Quantum Process Tomography, Phys. Rev. Lett. **90**, 193601 (2003).

[15] A. G. White, A. Gilchrist, G. J. Pryde, J. L. O'Brien, M. J. Bremner, and N. K. Langford, Measuring Controlled-NOT and two-qubit gate operation, LANL arXive eprint **0308115**, 1–10 (2003).

[16] G. M. D'Ariano and P. Lo Presti, Imprinting Complete Information about a Quantum Channel on its Output State, Phys. Rev. Lett. **91**, 047902 (2003).

[17] G. M. D'Ariano and P. L. Presti, Convex decomposition of measurements and devices, (2003), accepted for publication.

[18] G. M. D'Ariano and P. Lo Presti, Optimal Non-Universally Covariant Cloning, Phys. Rev. A **64**, 042308 (2001).

[19] G. M. D'Ariano and P. Lo Presti, Quantum homodyne tomography of photodetectors, page unpublished (2003).

[20] G. M. D'Ariano, P. Lo Presti, and M. G. A. Paris, Using entanglement improves precision of quantum measurements, Phys. Rev. Lett. **87**, 270404 (2001).

[21] G. M. D'Ariano, P. Lo Presti, and M. Sacchi, Bell Measurements and observables, Phys. Lett. A **272**, 32 (2000).

[22] K. Kraus, *States, Effects, and Operations*, Springer, Berlin, 1983.

[23] A. S. Holevo, *Probabilistic and statistical aspects of quantum theory*, volume 1 of *Series in Statistics and Probability*, North-Holland, Amsterdam, New York, Oxford, 1982.

[24] M. A. Naimark, Iza. Akad. Nauk USSR, Ser. Mat. **4**, 277 (1940).

[25] M. Ozawa, Quantum measuring processes of continuous observables, J. Math. Phys. **25**, 79 (1984).

[26] E. B. Davies, *Quantum Theory of Open Systems*, Academic Press, London, 1976.

[27] P. Busch, P. J. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement*, volume 2 of *Lecture Notes in Physics*, Springer, Berlin, 1991.

[28] K. R. Parthasaraty, Extremal decision rules in quantum hypothesis testing, Inf. Dim. Anal. **2**, 557 (1999).

[29] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1972.

[30] H. F. Jones, *Group representations and physics*, Institute of Physics Publ., Bristol and Philadelphia, 1999.

[31] D. Dieks, Communication by EPR devices, Phys. Lett. A **92**, 271 (1982).

[32] W. K. Wootters and W. H. Zurek, A single quantum cannot be cloned, Nature **299**, 802 (1981).

[33] V. Bužek and M. Hillery, Quantum copying: Beyond the no-cloning theorem, Phys. Rev. A **54**, 1844 (1996).

[34] H. P. Yuen, Amplification of quantum states and noiseless photon amplifiers, Phys. Lett. A **113**, 405–407 (1986).

[35] L.-M. Duan and G.-C. Guo, Probabilistic cloning and identification of linearly independent quantum states, Phys. Rev. Lett. **80**, 4999 (1998).

[36] N. Gisin and S. Massar, Optimal Quantum Cloning Machines, Phys. Rev. Lett. **79**, 2153 (1997).

[37] R. Werner, Optimal clonining of pure states, Phys. Rev. A **58**, 1827 (1998).

[38] D. Bruß, A. Ekert, and C. Macchiavello, Optimal Universal Quantum Cloning and State Estimation, Phys. Rev. Lett. **81**, 2598 (1998).

[39] C. Simon, G. Weihs, and A. Zeilinger, Optimal Quantum Cloning via Stimulated Emission, Phys. Rev. Lett. **84**, 2993 (2000).

[40] G. M. D'Ariano, F. De Martini, and M. F. Sacchi, Continuous variable cloning via network of parametric gates, Phys. Rev. Lett. **86**, 914 (2001).

[41] E. Galvao and L. Hardy, Cloning and quantum computation, Phys. Rev. A **62**, 022301 (2000).

[42] G. M. D'Ariano, C. Macchiavello, and M. F. Sacchi, Joint measurements via quantum cloning, J. Opt. B-Quantum Semiclass. Opt. **3**, 44 (2001).

[43] D. Bruß, J. Calsamiglia, and N. L utkenhaus, Quantum Cloning and Distributed Measurements, Phys. Rev. A **63**, 042308 (2001).

[44] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.

[45] D. Bruß, M. Cinchetti, G. M. D'Ariano, and C. Macchiavello, Phase covariant quantum cloning, Phys. Rev. A **62**, 012302–1–7 (2000).

[46] N. J. Cerf, A. Ipe, and X. Rottenberg, Universal cloning of continuous quantum variables, Phys. Rev. Lett. **85**, 1754 (2000).

[47] G. M. D'Ariano and M. F. Sacchi, Equivalence between squeezed-state and twin-beam communication channels, Mod. Phys. Lett. B **11**, 1263 (1997).

[48] I. L. Chuang and M. A. Nielsen, *Quantum Information and Quantum Computation*, Cambridge University Press, Cambridge UK, 2000.

[49] A. Luis, Quantum tomography of input-output processes, Phys. Rev. A **62**, 054302 (2000).

[50] D. Leung, *Towords robust quantum computation*, PhD thesis, Stanford University, 2001.

[51] R. B. Bapat, *Linear algebra and linear models*, Springer, Berlin, 2000.

[52] M. J. W. Hall, Gaussian noise and quantum optical communication, Phys. Rev. A **50**, 3295–3303 (1994).

[53] R. Simon, Peres-Horodecki separability criterion for continuous variable systems, Phys. Rev. Lett. **84**, 2726–2729 (2000).

[54] M. G. A. Paris, Evolution of twin-beam in active optical media, J. Opt. B: Quantum Semiclass. Opt. **4**, 442 (2002).

[55] G. M. D'Ariano and N. Sterpi, Robustness of homodyne tomography to phase-insensitive noise, J. Mod. Opt. **44**, 2227 (1997).

[56] M. Vasilyev, S.-K. Choi, P. Kumar, and G. M. D'Ariano, Tomographic measurement of joint photon statistics of the twin-beam quantum state, Phys. Rev. Lett. **84**, 2354 (2000).

[57] G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, Quantum Tomography, Advances in Imaging and Electron Physics **128**, 205–308 (2003).

[58] Y. Vardi and D. Lee, From image deblurring to optimal investments: maximum likelihood solutions for positive linear inverse problems, J. R. Statist. Soc. B **55**, 569 (1993).

[59] Z. Hradil, Quantum-state estimation, PRA **55**, R1561–R1564 (1997).

[60] Z. Hradil, J. Summhammer, G. Badurek, and H. Rauch, Reconstruction of the spin state, Phys. Rev. A **62**, 014101 (2000).

[61] J. Fiurášěk and Z. Hradil, Maximum-likelihood estimation of quantum processes, Phys. Rev. A **63**, 020101 (2001).

[62] K. Banaszek, G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, Maximum-likelihood estimation of the density matrix, Phys. Rev. A **61**, 010304(R) (2000).

[63] M. G. A. Paris, G. M. D'Ariano, and M. F. Sacchi, Maximum-likelihood method in quantum estimation, in *Bayesian inference and maximum entropy methods in science and engineering*, volume 568 of *AIP Conf. Proc.*, page 456, 2001, quant-ph/0101071.

[64] M. F. Sacchi, Characterizing a universal cloning machine by maximum-likelihood estimation, Phys. Rev. A **64**, 022106 (2001).

[65] M. F. Sacchi, Maximum-likelihood reconstruction of completely positive maps, Phys. Rev. A **63**, 054104 (2001).

[66] J. D. Gorman and A. O. Hero, Lower bounds for parametric estimation with constraints, IEEE Trans. Inf. Th. **26**, 1285–1301 (1990).

[67] T. Marzetta, A simple derivation of the constrained multiple parameter Cramer-Rao bound, IEEE Trans. on Signal Processing **41**, 2247–2249 (1993).

[68] R. Gill and M. I. Guţă, An invitation to quantum tomography, preprint, quant–ph/0303020 (2003).

[69] J. Rehacek, Z. Hradil, and M. Jezek, Iterative algorithm for reconstruction of entangled states, Phys. Rev. A **63**, R040303 (2001).

[70] A. P. Dempster, N. M. Laird, and D. B. Rubin, Maximum likelihood from incomplete data via the EM Algorithm, J. R. Statist. Soc. B **39**, 1–38 (1977).

[71] M. G. A. Paris, A robust verification of the quantum nature of light, Phys. Lett. A **489**, 167–171 (2001).

[72] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, New York, 1994.

[73] P. Spellucci, An SQP method for general nonlinear programs using only equality constrained subproblems, Math. Prog. **82**, 413–448 (1998).

[74] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete?, Phys. Rev. **47**, 777–780 (1935).

[75] H.-K. Lo, S. Popescu, and T. S. Eds., *Introduction to Quantum Computation and Information*, World Scientific, Singapore, 1998.

[76] M. I. Kolobov and C. Fabre, Quantum Limits on Optical Resolution, Phys. Rev. Lett. **85**, 3789 (2000).

[77] B. E. A. Saleh, B. M. Jost, H. Fei, and M. C. Teich, Entangled-photon virtual-state spectroscopy, Phys. Rev. Lett. **80**, 3483 (1998).

[78] P. W. Shor, Algorithms for Quantum Computation, Discrete Logarithms and Factoring, in *Proceedings of the 35th Annual Symposium of the Foundations of Computer Science*, page 124, Los Alamitos, CA, 1994, IEEE Computer Society Press.

[79] L. K. Grover, Quantum Mechanics Helps in Searching for a Needle in a Haystack, Phys. Rev. Lett. **79**, 325 (1997).

[80] C. W. Helstrom, *Quantum detection and estimation theory*, volume 123 of *mathematics in science and engineering*, Academic Press, New York, San Francisco, London, 1976.

[81] J. H. Shapiro and S. R. Shepard, Quantum Phase Measurement: A System-Theory Perspective, Phys. Rev. A **43**, 3795 (1991).

[82] A. Acín, E. Jané, and G. Vidal, Optimal estimation of quantum dynamics, Phys. Rev. A **64**, 050302 (2001).

[83] M. A. Nielsen, Conditions for a class of entanglement transformations, Phys. Rev. Lett. **83**, 436–439 (1999).

[84] E. Arthus and M. S. Goodman, Quantum Correlations: A Generalized Heisenberg Uncertainty Relation, Phys. Rev. Lett. **60**, 2447 (1988).

[85] H. P. Yuen and J. H. Shapiro, Quantum Measurements Realizable with Photoemissive Detectors, IEEE Trans. Inf. Theory **26**, 78 (1980).

[86] A. S. Holevo and R. F. Werner, Evaluating capacities of Bosonic Gaussian channels, Phys. Rev. A **63**, 032312 (2001).

[87] A. Acin, Statistical Distinguishability between Unitary Operations, Phys. Rev. Lett. **87**, 177901 (2001).

List of publications

This thesis is based on the following refereed publications

- G. M. D'Ariano, P. Lo Presti, M. F. Sacchi, *Bell measurements and observables*, Physics Letters A **272**, 32 (2000)
- G. M. D'Ariano and P. Lo Presti, *Optimal nonuniversally covariant cloning*, Physical Review A **64**, 042308 (2001)
- G. M. D Ariano and P. Lo Presti, *Quantum Tomography for Measuring Experimentally the Matrix Elements of an Arbitrary Quantum Operation*, Physical Review Letters **86**, 19 (2001)
- G. M. D'Ariano and P. Lo Presti, *Experimental characterization of the transfer matrix of a quantum device*, INFM Highlights 2000/2001
- G. M. D'Ariano, P. Lo Presti, and M. G. A. Paris, *Using Entanglement Improves the Precision of Quantum Measurements*, Physical Review Letters **87**, 27 (2001)
- G. M. D'Ariano, P. Lo Presti, M. F. Sacchi, *A quantum measurement of the spin direction*, Physics Letters A **292**, 233 (2002)
- G. M. D'Ariano, P. Lo Presti, M. G. A. Paris, *Improved discrimination of unitary transformations by entangled probes*, J. Opt. B: Quantum Semiclass. Opt. **4**, S273 (2002)
- M. G. A. Paris, G. M. D'Ariano, P. Lo Presti, P. Perinotti, *About the use of entanglement in the optical implementation of quantum information processing*, Fortschr. Phys. **51** No. 4-5, 449 (2003)
- G. M. D'Ariano, P. Lo Presti, *Imprinting a complete information about a quantum channel on its output state*, Physical Review Letters **91**, 047902 (2003)

- G. M. D'Ariano, P. Lo Presti, *Classical and quantum noise in measurements and transformations*, to appear in Physics Letters A (2003)
- G. M. D'Ariano, P. Lo Presti, *Characterization of Quantum Devices*, to appear as a chapter of a Lecture Notes in Physics, published by Springer-Verlag (2003)