

# 8 Characterization of Quantum Devices

Giacomo Mauro D'Ariano and Paoloplacido Lo Presti

QUIT Group of INFN, Dipartimento di Fisica "A. Volta", via Bassi 6, Pavia, Italy, [dariano@unipv.it](mailto:dariano@unipv.it), [lopresti@unipv.it](mailto:lopresti@unipv.it), [www.qubit.it](http://www.qubit.it)

**Abstract.** Using quantum tomography and a single entangled state it is possible to characterize completely a quantum device, a channel, or a measuring apparatus. The method is very robust to imperfections of the tomographers and of the input state (which more generally can be a "faithful" state), and can be made very efficient by max-likelihood methods specially designed for this purpose. Using this method with homodyne detection one can in principle achieve the first absolute characterization of a photocounter.

## 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 transformations 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 interaction with a bath. The class of probabilistic transformations, on the other hand, typically describe 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 that 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 [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 method in liquid nuclear magnetic resonance systems [5–7], 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 [10] and [11] it was shown 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 [12, 13]. 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 [14–16] for optical qubits, and proposed for optical “continuous variables” systems using homodyne tomography [10].

In [17], the two methods—“many inputs” versus “single entangled input”—have been bridged together in a complete classification of all states (and/or all ensembles 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 [15]. In [17], a measure of the “faithfulness” of the state has also been 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 [18] 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.

The present chapter is aimed at a complete and self-contained presentation of the theoretical basis of the methods for imprinting quantum operations on quantum states, also providing concrete examples of experimental setups based on homodyne tomography, to be used for tomography of either quantum processes or detectors. In Sect. 8.1 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 with customary unitary evolutions and projective measurements, and then to analyze the problem of the characterization of a device. This mathematical framework is then employed in Sect. 8.2 for giving the complete classification of the *faithful states*—i.e. the input states that can be used for the characterization of a quantum device—and to address the problem of quantifying their degree of faithfulness. Finally, Sect. 8.3 is devoted to the exposition of a quantum optical setup for performing a device characterization by homodyne tomography using an entangled input state from parametric down-conversion of vacuum. Finally, we report numerical simulations of experimental results that can be obtained with the current technology for the homodyne tomography of an amplitude displacing device and of an On/Off photo-detector, using either the averaging or the maximum-likelihood strategies.

## 8.1 Quantum Operations and Quantum Measurements

Quantum operations (QO), introduced for the first time in [1,2], describe all possible transformations—either deterministic or probabilistic—of the state  $\rho$  of a quantum system. Mathematically, QO's are completely positive (CP) linear maps from the set of (trace-class) operators on  $\mathcal{H}$  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 section 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, we will introduce a one-to-one correspondence between CP maps and positive operators on  $\mathcal{H}^{\otimes 2}$ , which provides the easiest framework to proof most relevant results concerning quantum operations.

At the end, we will review the concept of POVM and its connection with quantum operations, for representing the probability distribution of the outcomes of a quantum measurement.

### 8.1.1 Properties of Quantum Operations

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

$$\rho \mapsto \mathcal{T}(\rho) \tag{8.1}$$

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  $\rho$  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  $\rho$

$$\mathcal{T}(\rho) \in \mathbb{T}(\mathbb{H}), \quad \mathcal{T}(\rho) \geq 0, \quad \text{and} \quad \text{tr} \mathcal{T}(\rho) = 1, \tag{8.2}$$

where  $\mathbb{T}(\mathbb{H})$  denotes trace-class operators on  $\mathbb{H}$ . 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 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

$$p(\rho) = \sum_i p_i p(\rho_i), \tag{8.3}$$

$$\mathcal{T}(\rho) = \frac{\sum_i p_i p(\rho_i) \mathcal{T}(\rho_i)}{\sum_i p_i p(\rho_i)}. \tag{8.4}$$

The first equation implies that  $p(\rho)$  is a *linear* function of  $\rho$ , 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  $\mathcal{E}(\rho) \doteq p(\rho)\mathcal{T}(\rho)$ , (8.4) tells us that  $\mathcal{E}$  is a *linear* function of  $\rho$ . Taking the trace of the above definition of  $\mathcal{E}$ , and remembering that  $\text{tr} \mathcal{T}(\rho) = 1$ , we find  $p(\rho) = \text{tr} \mathcal{E}(\rho)$ , so that the transformation  $\mathcal{T}$

and the probability  $p(\rho)$  can be written in terms of  $\mathcal{E}$  as follows

$$\mathcal{T}(\rho) = \frac{\mathcal{E}(\rho)}{\text{tr } \mathcal{E}(\rho)}, \quad p(\rho) = \text{tr } \mathcal{E}(\rho). \quad (8.5)$$

From (8.2), and from the fact that  $p(\rho)$  is a probability, one argues that also the following properties must hold for  $\mathcal{E}$

$$\begin{aligned} \mathcal{E}(\rho) &\geq 0 \quad (\textit{positivity}), \\ \text{tr } \mathcal{E}(\rho) &\leq 1 \quad (\textit{trace decreasing or preserving}). \end{aligned} \quad (8.6)$$

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 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  $\mathcal{E} : \mathbb{T}(\mathbb{H}) \rightarrow \mathbb{T}(\mathbb{H})$ , with the state transformation given by  $\rho \mapsto \mathcal{E}(\rho)/\text{tr } \mathcal{E}(\rho)$ , and occurring with probability  $p(\rho) = \text{tr } \mathcal{E}(\rho)$ . 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 for a given outcome occurring with a probability  $p(\rho) = \text{tr } \mathcal{E}(\rho) \leq 1$ .

### 8.1.2 Representing CP Maps

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

of their “matrix elements”

$$\mathcal{E}_{ij}^{lm} = \langle i | \mathcal{E}(|l\rangle\langle m|) |j\rangle, \tag{8.7}$$

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

$$\mathcal{E}(\rho) = \sum_{ijklm} \mathcal{E}_{ij}^{lm} \rho_{lm} |i\rangle\langle j|. \tag{8.8}$$

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  $\mathbb{H} \otimes \mathbb{H}$ , 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 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  $\mathbb{H}$ , we identify any vector  $|\Psi\rangle\rangle \in \mathbb{H} \otimes \mathbb{H}$ ,

$$|\Psi\rangle\rangle = \sum_{m,n} \Psi_{mn} |m\rangle \otimes |n\rangle, \tag{8.9}$$

with the operator  $\Psi \in \mathbb{B}(\mathbb{H})$  whose matrix elements on the chosen basis are  $\Psi_{mn}$ . 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

$$\begin{aligned} A \otimes B |C\rangle\rangle &= |ACB^T\rangle\rangle, & \langle\langle A|B\rangle\rangle &= \text{tr}[A^\dagger B], \\ \text{tr}_2[|A\rangle\rangle\langle\langle B|] &= AB^\dagger, & \text{tr}_1[|A\rangle\rangle\langle\langle B|] &= A^T B^*, \end{aligned} \tag{8.10}$$

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 (8.9), 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  $\check{S}_\mathcal{E} \in \mathbb{B}(\mathbb{H} \otimes \mathbb{H})$  such that

$$|\mathcal{E}(\rho)\rangle\rangle = \check{S}_\mathcal{E} |\rho\rangle\rangle. \tag{8.11}$$

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 (8.7) one finds

$$\check{S}_\mathcal{E} = \sum_{ijklm} \mathcal{E}_{ij}^{lm} |i\rangle\langle l| \otimes |j\rangle\langle m|. \tag{8.12}$$

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

$$|\mathcal{E}_1 \circ \mathcal{E}_2(\rho)\rangle\rangle = \check{S}_{\mathcal{E}_1} |\mathcal{E}_2(\rho)\rangle\rangle = \check{S}_{\mathcal{E}_1} \check{S}_{\mathcal{E}_2} |\rho\rangle\rangle . \tag{8.13}$$

Moreover, such a representation provides a useful tool to evaluate some properties of the map. For example, the image of the map  $\mathcal{E}(\mathbf{B}(\mathbf{H}))$  corresponds to the set the operators  $A$  such that  $|A\rangle\rangle \in \text{Rng } \check{S}_{\mathcal{E}}$ , where ‘‘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  $|A\rangle\rangle \in \text{Ker } \check{S}_{\mathcal{E}}$ . Finally, by definition,  $\mathcal{E}$  is invertible iff  $\check{S}_{\mathcal{E}}$  is invertible, and the two inverses are related through the identity

$$|\mathcal{E}^{-1}(\rho)\rangle\rangle = \check{S}_{\mathcal{E}}^{-1} |\rho\rangle\rangle , \tag{8.14}$$

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. In order to explore this, it is convenient to introduce another operator representation of the map  $\mathcal{E}$  in terms of the operator  $S_{\mathcal{E}} \in \mathbf{B}(\mathbf{H} \otimes \mathbf{H})$  resulting from the action of the extended map  $\mathcal{E} \otimes \mathcal{I}$  on the operator  $|I\rangle\rangle\langle\langle I| \in \mathbf{B}(\mathbf{H} \otimes \mathbf{H})$  [12,13], namely

$$S_{\mathcal{E}} = (\mathcal{E} \otimes \mathcal{I}) [|I\rangle\rangle\langle\langle I|] = \sum_{ijklm} \mathcal{E}_{ij}^{lm} |i\rangle\langle j| \otimes |l\rangle\langle m| . \tag{8.15}$$

The inverse relation of identity (8.15) can be easily checked to be

$$\mathcal{E}(\rho) = \text{tr}_2[(I \otimes \rho^T) S_{\mathcal{E}}] . \tag{8.16}$$

A comparison between (8.12) and (8.15) 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

$$\check{S}_{\mathcal{E}} = (S_{\mathcal{E}}^{T_2} E)^{T_2} = (E S_{\mathcal{E}}^{T_1})^{T_1} , \tag{8.17}$$

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 (8.16) with  $S_{\mathcal{E}} \geq 0$  is CP. In fact, given that  $S_{\mathcal{E}}$  is positive, it can be decomposed as

$$S_{\mathcal{E}} = \sum_i |A_i\rangle\rangle\langle\langle A_i| , \tag{8.18}$$

so that by substituting the above equation into (8.16), and applying the rules of (8.10), one finds that the resulting map can be expressed in the so called *Kraus form* [19]

$$\mathcal{E}(\rho) = \sum_i A_i \rho A_i^\dagger. \tag{8.19}$$

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  $R \in \mathbf{B}(\mathbf{H} \otimes \mathbf{K})$  is

$$R_{\mathcal{E}} = \mathcal{E} \otimes \mathcal{I}[R] = \sum_i (A_i \otimes I) R (A_i^\dagger \otimes I), \tag{8.20}$$

which is still positive since

$$\langle\langle \Psi | R_{\mathcal{E}} | \Psi \rangle\rangle = \sum_i \langle\langle A_i^\dagger \Psi | R | A_i^\dagger \Psi \rangle\rangle \geq 0, \quad \forall |\Psi\rangle. \tag{8.21}$$

Of course, any CP map  $\mathcal{E}$  admits a Kraus form that can be found by decomposing  $S_{\mathcal{E}}$  as we did in (8.18). When this decomposition is a diagonalization, i.e., when  $|A_i\rangle$  are the unnormalized orthogonal eigenvectors, then the related Kraus form is said to be *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  $\text{rank } S_{\mathcal{E}}$ . 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

$$\check{S}_{\mathcal{E}} = \sum_i A_i \otimes A_i^*, \tag{8.22}$$

as it easily follows from the definition of  $\check{S}_{\mathcal{E}}$  in (8.11) and the first rule in (8.10).

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

$$\text{tr}_1 S_{\mathcal{E}} \leq I \quad \text{or equivalently} \quad \sum_i A_i^\dagger A_i \leq I, \tag{8.23}$$

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

If  $\text{rank } S_{\mathcal{E}} = 1$  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  $\{A_i\}_{i=1\dots r}$  of the map  $\mathcal{E}$  one

can define an operator  $U$  on the Hilbert space  $\mathbf{H} \otimes \mathbb{C}^r$  whose action on the vectors of the basis of the form  $|m\rangle|0\rangle$  is defined as

$$U|m\rangle|0\rangle = \sum_{i=1}^r (A_i|m\rangle) |i\rangle = |m, 0\rangle\rangle', \quad (8.24)$$

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 (8.24) are orthonormal: the operator  $U$  can then be easily extended 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  $\rho$  by means of the unitary transformation  $U$ , the final “local” state of the system only reads

$$\mathcal{E}(\rho) = \text{tr}_2[ U (\rho \otimes |0\rangle\langle 0|) U^\dagger ]. \quad (8.25)$$

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

$$\rho_i = \frac{A_i \rho A_i^\dagger}{\text{tr} [ A_i \rho A_i^\dagger ]}. \quad (8.26)$$

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  $\mathcal{E}$  that 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 privileged role at the axiomatic level.

As we argued from (8.26), 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  $\{1 \dots N\}$ , and such that in relation to the outcome  $k$  the state is transformed according to a map  $\mathcal{E}^{(k)}$  whose Kraus decomposition is  $\{A_i^{(k)}\}_{i=1 \dots r_k}$ . 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

$$\mathcal{E}(\rho) = \sum_k p(k|\rho) \frac{\mathcal{E}^{(k)}(\rho)}{\text{tr} \mathcal{E}^{(k)}(\rho)} = \sum_k \mathcal{E}^{(k)}(\rho). \quad (8.27)$$

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

$$U|m\rangle|0\rangle|0\rangle = \sum_{k=1}^N \sum_{i=1}^{r_k} (A_i^{(k)}|m\rangle) |i\rangle|k\rangle, \tag{8.28}$$

then  $U$  can be completed to a unitary operator on the whole space by the Gram-Schmidt procedure, since the resulting vectors in the above equation are orthonormal. The original maps can now be realized by evolving the system in the state  $\rho$  jointly with the two additional ancillas prepared in the state  $|0\rangle|0\rangle$  with the unitary  $U$ , and then performing a projective measurement  $|k\rangle\langle k|$  on the second ancilla while disregarding the first one with a partial trace, i.e.

$$\mathcal{E}^{(k)}(\rho) = {}_3\langle k| \operatorname{tr}_2 [ U (\rho \otimes |0\rangle\langle 0| \otimes |0\rangle\langle 0|) U^\dagger ] |k\rangle_3. \tag{8.29}$$

Also in this case, the maps  $\mathcal{E}^{(k)}$  are non-pure because some information 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  $(i, k)$  the state of the system would be described by a pure quantum operation

$$\rho_{(i,k)} = \frac{A_i^{(k)} \rho A_i^{(k)\dagger}}{\operatorname{tr} [ A_i^{(k)} \rho A_i^{(k)\dagger} ]}. \tag{8.30}$$

### 8.1.3 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 Sect. 8.1.1, in particular from (8.3), it follows that the probability distribution of the outcomes of any quantum measurement must be linear in the state  $\rho$ , and that therefore it is described by the so called Born’s rule

$$p(k|\rho) = \operatorname{tr}[\rho P_k], \tag{8.31}$$

where  $k$  is the outcome and the set  $\{P_k\}$  is called *positive operator valued measure (POVM)*, namely it is a set of operators that must be positive and with  $\sum_k P_k = I$ , in order to have  $p(k|\rho)$  a properly positive and normalized probability distribution [20].

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 for which each outcome  $k$  is described by the CP

maps  $\mathcal{E}^{(k)}$ . By means of (8.16) the probability distribution of outcomes reads

$$p(k|\rho) = \text{tr}[\mathcal{E}^{(k)}(\rho)] = \text{tr}[\rho \text{tr}_1[S_{\mathcal{E}^{(k)}}]^T] = \text{tr}[\rho P_k], \quad (8.32)$$

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

$$P_k = \text{tr}_1[S_{\mathcal{E}^{(k)}}]^T = \sum_i A_i^{(k)\dagger} A_i^{(k)}, \quad (8.33)$$

$\{A_i^{(k)}\}$  being a Kraus decomposition of the  $k$ -th map. On the contrary, given a POVM  $\{P_k\}$  one can always find a set of QO's  $\mathcal{E}^{(k)}$  describing a measuring process with the given POVM, for example using  $A_k = \sqrt{P_k}$  and  $\mathcal{E}^{(k)}(\rho) = A^{(k)}\rho A^{(k)\dagger}$ . By “purifying” these maps with the unitary transformation  $U$  defined in (8.28) 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.

## 8.2 Imprinting Quantum Operations into Quantum States

Characterizing a quantum device means performing a measurement that provides information about the QO performed 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 state contains an imprinting of the device. The aim of this section is to classify 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 an ensemble of them, and we will correspondingly call the ensemble *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 see how a single pure entangled state can support a full imprinting of the QO [10]. Then we will extend the analysis to mixed states, showing how entanglement is not strictly needed [17], and finally giving a complete characterization of faithful states and ensembles, 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 an ensemble of state, 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 extend the treatment to devices performing non deterministic QO’s. Finally we will also present how to encode a POVM on quantum states [21].

### 8.2.1 Ensembles of Input States Versus a Single Entangled State

The first proposed methods 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 is completely defined by its action on a set of operators generating the linear space of all operators  $\mathbf{B}(\mathbf{H})$ . Hence, for encoding a quantum process  $\mathcal{E}$  on states, one should look for a set of states  $\rho_i$  that span  $\mathbf{B}(\mathbf{H})$ , 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 state  $\rho$  can be recovered by expanding  $\rho$  on the generators of the space,  $\rho = \sum_i c_i \rho_i$ , so that by linearity one obtains the action of the map on any state  $\rho$  as  $\mathcal{E}(\rho) = \sum_i c_i \mathcal{E}(\rho_i)$ .

As an example, consider the set of states given in [22] for quantum process tomography

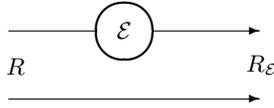
$$\left\{ |m\rangle, |\phi_{mn}\rangle = \frac{|m\rangle + |n\rangle}{\sqrt{2}}, |\psi_{mn}\rangle = \frac{|m\rangle + i|n\rangle}{\sqrt{2}} \right\} \quad (8.34)$$

it is a *faithful set of states*, as it is a set of generators for  $\mathbf{B}(\mathbf{H})$  because the elements of the basis  $|m\rangle\langle n|$  of  $\mathbf{B}(\mathbf{H})$  can be written as

$$|m\rangle\langle n| = |\phi_{mn}\rangle\langle\phi_{mn}| + i|\psi_{mn}\rangle\langle\psi_{mn}| - \frac{1+i}{2}|m\rangle\langle m| - \frac{1+i}{2}|n\rangle\langle n|. \quad (8.35)$$

Quantum process tomography has been realized with this method in liquid nuclear magnetic resonance systems [5–7], 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 [23], 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  $\dim(\mathbf{H})^2$  of



**Fig. 8.1.** 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  $\text{rank } A = \dim(\mathbf{H})$ .

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 many measurements).

A viable alternative to the above method of “spanning states” inspired by the operator representation of a channel was presented in [10] and [11], and experimentally implemented for polarization qubits in [14–16]. 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. 8.1, the output state  $R_{\mathcal{E}}$  reads

$$R_{\mathcal{E}} = (\mathcal{E} \otimes \mathcal{I}) [|A\rangle\rangle\langle\langle A|] = (I \otimes A^T) S_{\mathcal{E}} (I \otimes A^*) . \quad (8.36)$$

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

$$S_{\mathcal{E}} = [I \otimes (A^T)^{-1}] R_{\mathcal{E}} [I \otimes (A^*)^{-1}] , \quad (8.37)$$

and then the action of the map on a state  $\rho$  is found via (8.16), namely

$$\mathcal{E}(\rho) = \text{tr}_2 [ (I \otimes \rho^T) S_{\mathcal{E}} ] . \quad (8.38)$$

Summarizing, any bipartite state with maximal Schmidt number is faithful, namely by entering a quantum device it gets imprinted the full information about its channel. 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 [18], where a single measurement 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 (8.34) 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 complete 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 section, when a measure of “faithfulness” will be introduced, and also in the next section, where some practical applications of this framework for characterizing quantum devices will be exposed.

### 8.2.2 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

$$\begin{aligned} R_{\mathcal{E}} &= (\mathcal{E} \otimes \mathcal{I})[R] = \sum_l (I \otimes A_l^T) (\mathcal{E} \otimes \mathcal{I})[|I\rangle\rangle\langle\langle I|] (I \otimes A_l^*) = \\ &= \sum_l (I \otimes A_l^T) S_{\mathcal{E}} (I \otimes A_l^*) . \end{aligned} \tag{8.39}$$

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

$$\mathcal{R}(\rho) = \sum_l A_l^T \rho A_l^* , \tag{8.40}$$

it is immediate to notice that

$$R_{\mathcal{E}} = (\mathcal{I} \otimes \mathcal{R})[S_{\mathcal{E}}] , \tag{8.41}$$

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  $\mathcal{R}$  that is invertible. Recalling what we wrote in Sect. 8.1.2, and in particular (8.14), the invertibility of the CP map  $\mathcal{R}$  resorts to the invertibility of a customary operator. In fact, by considering the following equation involving vectors in  $\mathbf{H} \otimes \mathbf{H}$

$$|\mathcal{R}(\rho)\rangle\rangle = \left| \sum_l A_l^T \rho A_l^* \right\rangle\rangle = \left( \sum_l A_l^T \otimes A_l^\dagger \right) |\rho\rangle\rangle \doteq \check{S}_{\mathcal{R}} |\rho\rangle\rangle, \quad (8.42)$$

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  $\mathbf{H} \otimes \mathbf{H}$  is invertible. As we already noticed in Sect. 8.1.2, the action of the inverse map  $\mathcal{R}^{-1}$  can be defined through the relation

$$|\mathcal{R}^{-1}(\rho)\rangle\rangle \doteq \check{S}_{\mathcal{R}}^{-1} |\rho\rangle\rangle, \quad (8.43)$$

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

$$\check{S}_{\mathcal{R}} = (ER)^{T_2} E = (R^{T_2} E)^{T_1} \quad (8.44)$$

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  $l$ th 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

$$\mathcal{E}(\rho) = \text{tr}_2 \left[ (I \otimes \rho^T) (\mathcal{I} \otimes \mathcal{R}^{-1})[R_{\mathcal{E}}] \right]. \quad (8.45)$$

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  $\mathbf{H}$  and  $\mathbf{K}$ . We need now to consider vectors in either  $\mathbf{H} \otimes \mathbf{K}$ ,  $\mathbf{H}^{\otimes 2}$ , or  $\mathbf{K}^{\otimes 2}$ , and in all cases we will keep our notation  $|A\rangle\rangle$  for the vectors, with the corresponding operator  $A$  in  $\mathbf{B}(\mathbf{K}, \mathbf{H})$ ,  $\mathbf{B}(\mathbf{H})$ , or  $\mathbf{B}(\mathbf{K})$  respectively.

Similarly to the previous reasoning lines, in relation to the bipartite input state  $R = \sum_l |A_l\rangle\rangle\langle\langle A_l|$  on  $\mathbf{H} \otimes \mathbf{K}$ , 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  $\mathbf{B}(\mathbf{H})$  to  $\mathbf{B}(\mathbf{K})$ . 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<sup>1</sup>. The operator  $\check{S}_{\mathcal{R}} = \sum_l A_l^T \otimes A_l^\dagger$  associated to  $\mathcal{R}$  now maps vectors in  $\mathbb{H}^{\otimes 2}$  to vectors in  $\mathbb{K}^{\otimes 2}$ , 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  $\mathbb{H}^{\otimes 2}$  to  $\mathbb{K}^{\otimes 2}$ , that in turn is equivalent to the condition  $\text{rank } \check{S}_{\mathcal{R}} = \dim(\mathbb{H})^2$ . Among all the possible left-inverses of  $\check{S}_{\mathcal{R}}$  one can use the Moore-Penrose pseudo-inverse  $\check{S}_{\mathcal{R}}^\ddagger$ , and thus define the left-inverse of the map  $\mathcal{R}$  as

$$|\mathcal{R}^{-1}(\rho)\rangle\rangle \doteq \check{S}_{\mathcal{R}}^\ddagger|\rho\rangle\rangle, \tag{8.46}$$

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

### 8.2.3 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 between two processes  $\mathcal{E}_1$  and  $\mathcal{E}_2$ . For input state  $R$ , their respective outputs will be

$$R_{\mathcal{E}_1} = (\mathcal{I} \otimes \mathcal{R})[S_{\mathcal{E}_1}] \quad \text{and} \quad R_{\mathcal{E}_2} = (\mathcal{I} \otimes \mathcal{R})[S_{\mathcal{E}_2}], \tag{8.49}$$

and thus we shall tune  $R$  in order to improve the distinguishability 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

---

<sup>1</sup> A generic operator  $T : \mathbb{H} \rightarrow \mathbb{K}$  is left-invertible iff  $\text{rank } T = \dim(\mathbb{H})$ . For having  $T$  left-invertible is therefore necessary that  $\dim(\mathbb{K}) \geq \dim(\mathbb{H})$ , 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  $T^\ddagger$  [24] is the most used one, due to its nice properties. Starting from the *singular values decomposition* (SVD) of  $T$

$$T = \sum_i \sigma_i |v_i\rangle\langle u_i|, \tag{8.47}$$

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

$$T^\ddagger = \sum_i \sigma_i^{-1} |u_i\rangle\langle v_i|. \tag{8.48}$$

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

map  $\mathcal{R}$  in (8.49) to keep its outputs as far as possible. By considering the singular value decomposition for the operator  $\check{S}_{\mathcal{R}}$

$$\check{S}_{\mathcal{R}} = \sum_i \sigma_i |V_i\rangle\langle\langle U_i|, \quad (8.50)$$

with  $\{|V_i\rangle\}$  and  $\{|\langle U_i|\rangle\}$  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  $\rho$  becomes

$$\mathcal{R}(\rho) = \sum_i \sigma_i \text{tr}[U_i^\dagger \rho] V_i, \quad (8.51)$$

whence it is clear that the smaller are the singular values  $\sigma_i$ , the nearer are the outputs of  $\mathcal{R}$ , since their components on the basis  $\{|V_i\rangle\}$  will be shrunk. 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

$$F(R) = \sum_i \sigma_i^2 = \text{tr}[\check{S}_{\mathcal{R}}^\dagger \check{S}_{\mathcal{R}}]. \quad (8.52)$$

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

$$\begin{aligned} \text{tr}[\check{S}_{\mathcal{R}}^\dagger \check{S}_{\mathcal{R}}] &= \sum_{ij} \text{tr}[A_i^T A_j^*] \text{tr}[A_i A_j^\dagger] = \sum_i \text{tr}[A_i^T A_i^*] \text{tr}[A_i A_i^\dagger] = \\ &= \sum_i (\langle\langle A_i|A_i\rangle\rangle)^2 = \text{tr}[R^\dagger R]. \end{aligned} \quad (8.53)$$

Therefore, from (8.52) one obtains

$$F(R) = \text{tr}[R^\dagger R], \quad (8.54)$$

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 that 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 reconstruct the output  $R_{\mathcal{E}}$ . The measured  $R_{\mathcal{E}}$  will be affected by experimental errors that 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  $\sigma_i^{-1}$ , then the smaller the singular values of  $\check{S}_{\mathcal{R}}$  are the higher the amplification of experimental errors on the measured  $S_{\mathcal{E}}$ .

For an unfaithful state  $R$ ,  $\check{S}_{\mathcal{R}}$  will have 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.

### 8.2.4 Faithful Ensembles of States

Now we will consider the case in which not a single bipartite state, but an ensemble of them  $\{R^{(n)}\}_{n=1}^N$  on  $H \otimes K$  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  $R_{\text{set}}$  on  $H \otimes K \otimes \mathbb{C}^N$  defined as

$$R_{\text{set}} = \sum_{n=1}^N p_n R^{(n)} \otimes |n\rangle\langle n|, \tag{8.55}$$

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

$$R_{\text{set}} \mathcal{E} = (\mathcal{E} \otimes \mathcal{I} \otimes \mathcal{I}) [R_{\text{set}}] = \sum_{n=1}^N p_n R_{\mathcal{E}}^{(n)} \otimes |n\rangle\langle n| \tag{8.56}$$

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  $R_{\text{set}}$  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  $R_{\text{set}}$ . The latter can be evaluated with the techniques exposed in the previous paragraph for bipartite states, by simply considering  $R_{\text{set}}$  as a bipartite state of  $H$  and  $K \otimes \mathbb{C}^N$ .

The nature of the state  $R_{\text{set}}$  can be interpreted from two subtly different points of view. On one hand, to use  $R_{\text{set}}$  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,  $R_{\text{set}}$  represents the situation in which the states  $\{R^{(n)}\}$  are employed in the characterization each with a frequency equal to  $p_n$ . In fact, measuring 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 reason, any quantity (e.g. the faithfulness  $F$ ) being defined for faithful states can be extended consistently to ensembles of states simply by

evaluating it on the corresponding  $R_{\text{set}}$ . For example, the faithfulness of a set of generating states  $\rho_n$  (employed with the same frequency) is equivalent to the faithfulness of the bipartite state  $R_{\text{set}} = \sum_n \frac{1}{n} \rho_n \otimes |n\rangle\langle n|$ , 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  $\mathcal{O}[1/\dim(\mathbf{H})]$ .

### 8.2.5 Patching Sets of Unfaithful States

An unfaithful state  $R$  can still be useful in encoding only some quantum channels, or at least in encoding 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  $\rho$  such that  $|\rho\rangle \in \text{Ker}(\check{\mathcal{S}}_{\mathcal{R}})$ ), one can still employ its pseudo-inverse  $\mathcal{R}^\ddagger$  defined as

$$|\mathcal{R}^\ddagger(\rho)\rangle\rangle \doteq \check{\mathcal{S}}_{\mathcal{R}}^\ddagger |\rho\rangle\rangle . \tag{8.57}$$

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

$$|\mathcal{Q}(\rho)\rangle\rangle = \check{\mathcal{S}}_{\mathcal{R}}^\ddagger \check{\mathcal{S}}_{\mathcal{R}} |\rho\rangle\rangle = \check{\mathcal{S}}_{\mathcal{Q}} |\rho\rangle\rangle , \tag{8.58}$$

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

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

$$\check{\mathcal{S}}_{\mathcal{E}} = (\mathcal{I} \otimes \mathcal{R}^\ddagger)[R_{\mathcal{E}}] = (\mathcal{I} \otimes \mathcal{Q})[S_{\mathcal{E}}] \tag{8.59}$$

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

For any bipartite  $R$  one can define a *number of faithfulness*  $\varphi$  as  $\varphi(R) = \text{rank}(\check{\mathcal{S}}_{\mathcal{R}})$ , 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  $\varphi(R) = \dim(\mathbf{H})^2$ . Notice that for  $\varphi(R) < \dim(\mathbf{H})^2$  one can have the situation in which  $\text{Ker}^\perp(\check{\mathcal{S}}_{\mathcal{R}}) = \text{Span}\{|\rho\rangle\rangle, \rho \in \mathcal{A}\}$ , with  $\mathcal{A}$  being 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  $\rho$  from its projections  $\mathcal{Q}^{(n)}(\rho)$ , and this is possible iff, given a basis  $\{B_i\}$  for  $\mathbf{B}(\mathbf{H})$ , one has  $\text{Span}\{\mathcal{Q}^{(n)}(B_i)\}_{i,n} = \mathbf{B}(\mathbf{H})$ . In such

circumstances, any element of the basis can be expressed as a linear combination of the  $\mathcal{Q}^{(n)}(B_i)$ , i.e.  $B_i = \sum_{jn} \lambda_{ij}^n \mathcal{Q}^{(n)}(B_j)$ , and therefore it is possible to recover  $M \equiv \sum_i \text{tr}[B_i^\dagger M] B_i$  by “patching” the projections  $\mathcal{Q}^{(n)}(M)$  as

$$M = \sum_{ijn} \lambda_{ij}^{n*} \text{tr}[B_j^\dagger \mathcal{Q}^{(n)}(M)] B_i . \tag{8.60}$$

Analogously, by patching the partial encodings  $\{\mathcal{S}_{\mathcal{E}}^{(n)}\}$  [see (8.59)] we get  $S_{\mathcal{E}}$  as

$$S_{\mathcal{E}} = \sum_{ijn} \lambda_{ij}^{n*} \text{tr}_2[(I \otimes B_j^\dagger) \tilde{\mathcal{S}}_{\mathcal{E}}^{(n)}] \otimes B_i . \tag{8.61}$$

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.

### 8.2.6 Generalization to QO’s and POVM’s

Suppose we have a quantum device performing the measurement described by the CP maps  $\mathcal{E}_i$ ,  $i = 1 \dots N$  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  $i$  will be

$$R_{\mathcal{E}_i} = \frac{(\mathcal{I} \otimes \mathcal{R})[S_{\mathcal{E}_i}]}{\text{tr}[(\mathcal{I} \otimes \mathcal{R})[S_{\mathcal{E}_i}]]} , \tag{8.62}$$

where the denominator is also the probability of occurrence for the outcome  $i$ . In the case of  $R$  faithful, from this output it is possible to recover  $S_{\mathcal{E}_i}$  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  $i$ , thus obtaining  $N$  different ensembles, each labeled by the corresponding  $i$ , and described by the states  $R_{\mathcal{E}_i}$ . The denominator of (8.62) can be evaluated as the fraction of systems of the original ensemble that have been transformed into the  $i$ -th state, therefore an exact reconstruction of all the  $S_{\mathcal{E}_i}$  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 for 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  $\rho_i$  on the unmeasured system is still available. It reads,

$$\rho_i = \text{tr}_1 R_{\mathcal{E}_i} = \frac{\mathcal{R}[\text{tr}_1 S_{\mathcal{E}_i}]}{\text{tr}[\mathcal{R}[\text{tr}_1 S_{\mathcal{E}_i}]]} = \frac{\mathcal{R}[P_i^T]}{\text{tr}[\mathcal{R}[P_i^T]]} , \tag{8.63}$$

where  $P_i$  is the POVM of the measurement relative to the outcome  $i$ . 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  $i$ .

### 8.2.7 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. Let us see some examples of separable faithful states.

The Werner's states for dimension  $d$

$$R_f = \frac{1}{d(d^2 - 1)}[(d - f)I + (df - 1)E], \quad -1 \leq f \leq 1, \quad (8.64)$$

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

$$(ER_f)^{T_2} = \frac{1}{d(d^2 - 1)}[(d - f)|I\rangle\rangle\langle\langle I| + (df - 1)], \quad (8.65)$$

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 [15] an experiment employing these states for quantum process tomography was presented.

Similarly, the ‘‘isotropic’’ states

$$R_f = \frac{f}{d}|I\rangle\rangle\langle\langle I| + \frac{1-f}{d^2-1}(I - \frac{1}{d}|I\rangle\rangle\langle\langle I|), \quad (8.66)$$

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  $\frac{f}{d}$ .

### 8.2.8 Faithfulness in Infinite Dimensions

For infinite dimensions (the so-called ‘‘continuous variables’’ in quantum optics), one needs to restrict  $\mathbf{B}(\mathbf{H})$  to the Hilbert space of Hilbert-Schmidt operators on  $\mathbf{H}$ , 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  $\mathbf{B} = \{B_j\}$  in  $\mathbf{B}(\mathbf{H})$  used for representing the channel map. As an example, let's consider a twin beam

from parametric down-conversion of vacuum

$$|\Psi\rangle\rangle = \Psi \otimes I|I\rangle\rangle, \quad \Psi = (1 - |\xi|^2)^{\frac{1}{2}} \xi a^\dagger a, \quad |\xi| < 1 \quad (8.67)$$

for a fixed  $\xi$ ,  $a^\dagger$  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  $\Psi^{-1}$  is unbounded, whence the inverse map  $\mathcal{R}^{-1}$  is also unbounded. In a photon number representation  $\mathbf{B} = \{|n\rangle\langle m|\}$ , 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* [25]

$$\mathcal{N}_\nu(\rho) = \int_{\mathbb{C}} \frac{d\alpha}{\pi\nu} \exp[-|\alpha|^2/\nu] D(\alpha)\rho D^\dagger(\alpha), \quad (8.68)$$

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, analogous of the depolarizing channel for infinite dimension. The maps  $\mathcal{N}_\nu$  for varying  $\nu$  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

$$R = \mathcal{I} \otimes \mathcal{N}_\nu(|\Psi\rangle\rangle\langle\langle\Psi|). \quad (8.69)$$

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

$$R = \frac{1}{\nu} (\Psi \otimes I) \exp[-(a - b^\dagger)(a^\dagger - b)/\nu] (\Psi^\dagger \otimes I), \quad (8.70)$$

thus its partial transposed on the second space reads

$$R^{T_2} = (\nu + 1)^{-1} (\Psi \otimes I) \left( \frac{\nu - 1}{\nu + 1} \right)^{\frac{1}{2}(a-b)^\dagger(a-b)} (\Psi^\dagger \otimes I), \quad (8.71)$$

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 [26], therefore, for  $\nu > 1$ ,  $R$  is separable (see also [27]), yet it is *formally* faithful, since the operator  $\Psi$  and the map  $\mathcal{N}_\nu$  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  $\mathbf{B} = \{B_j\}$ , 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  $\mathbf{B} = \{|\alpha\rangle\langle\beta|\}$ , with  $|\alpha\rangle$  and  $|\beta\rangle$  coherent states, since from the identity

$$\mathcal{N}_\nu(|\alpha\rangle\langle\alpha|) = \frac{1}{\nu+1} D(\alpha) \left( \frac{\nu}{\nu+1} \right)^{a^\dagger a} D^\dagger(\alpha), \quad (8.72)$$

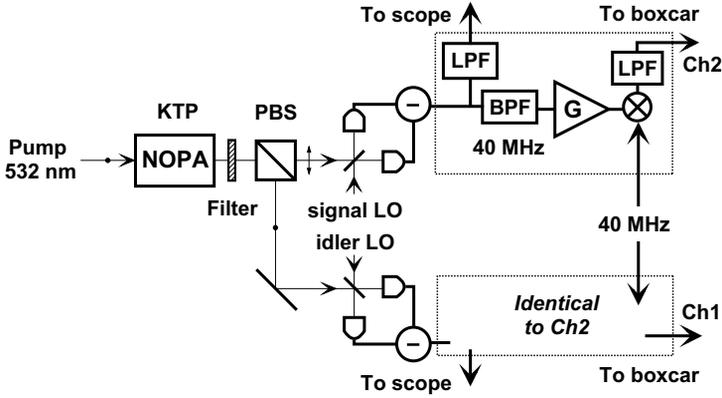
one obtains

$$\mathcal{N}_\nu^{-1}(|\alpha\rangle\langle\alpha|) = \frac{1}{1-\nu} D(\alpha) (1-\nu^{-1})^{-a^\dagger a} D^\dagger(\alpha), \quad (8.73)$$

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 [28]. Therefore, we say that the state is *formally* faithful, however, we are constrained to representations that are analytical for the inverse map  $\mathcal{R}^{-1}$ .

### 8.3 Homodyne Tomography of Channels and POVM's

Once the information about a device is encoded into quantum states, all the techniques of quantum tomography, which are also reviewed in this set of Lecture Notes, 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 [5–7] or in quantum optics [8, 9, 14, 15], the latter also deserving an entire chapter in these Lecture Notes. 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 [29], in which quantum homodyne tomography of the (joint number probability distribution of) a twin-beam was achieved using the setup depicted in Fig. 8.2. After a brief introduction on homodyne tomography, we report as an example of quantum process tomography the result presented in [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. For the tomography of the unitary transformation the tomographic reconstruction will be performed by the method of pattern function averaging. For the tomographic of the photo-detector, on the other hand, we will also consider maximum likelihood methods, to show how they can give a huge boost to the precision of the characterization, at the sole expense of greater computational complexity.



**Fig. 8.2.** 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  $\phi'$  and  $\phi''$  with respect to the local oscillators. (From [29])

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.

### 8.3.1 Homodyne Tomography

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

$$X_{\phi} = \frac{a^{\dagger} e^{i\phi} + a e^{-i\phi}}{2}, \tag{8.74}$$

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  $\phi$ . In the Fock basis  $|n\rangle$  the (unnormalizable) eigenstate  $|x\rangle_{\phi}$  of the quadrature  $X_{\phi}$  is given by

$$|x\rangle_{\phi} = \sum_{n=0}^{\infty} \left(\frac{2}{\pi}\right)^{\frac{1}{4}} \frac{1}{\sqrt{2^n n!}} \exp(-x^2) H_n(\sqrt{2}x) e^{in\phi} |n\rangle, \tag{8.75}$$

$H_n(x)$  denoting Hermite polynomials. Once the phase  $\phi$  is fixed, the ideal measurement realizes the POVM  $\text{Hom}(x; \phi) = |x\rangle_{\phi} \langle x|$  for the “continuous

variable"  $x$ , with a probability density distribution of the outcomes given by product

$$p(x; \phi) = \text{tr} [\rho \text{Hom}(x; \phi)] , \quad (8.76)$$

$\rho$  being the state of the system. In the non-ideal situation on 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  $\eta$  denoting the *quantum efficiency* of photo-detectors used in the homodyne.

Homodyne tomography is a method for estimating the state  $\rho$  from a finite sample of homodyne data, i.e distributed according to  $p(x; \phi)$  in (8.76). 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

$$\langle O \rangle = \text{tr}[\rho O] = \int_0^\pi \frac{d\phi}{\pi} \int_{-\infty}^{+\infty} dx p_\eta(x; \phi) \mathcal{P}_\eta[O](x, \phi) , \quad (8.77)$$

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

$$\frac{1}{N} \sum_{n=0}^N \mathcal{P}_\eta[O](x_n, \phi_n) \xrightarrow{N \rightarrow \infty} \langle O \rangle \quad (\text{with probability } 1) . \quad (8.78)$$

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 the first chapter of this set of Lecture Notes.

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  $X_\phi$  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

$$\mathcal{P}[O_1 \otimes O_2](x_n, \phi_n, x'_n, \phi'_n) = \mathcal{P}[O_1](x_n, \phi_n) \mathcal{P}[O_2](x'_n, \phi'_n) , \quad (8.79)$$

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

$$\frac{1}{N} \sum_{n=0}^N \mathcal{P}_\eta[|j\rangle\langle i|](x_n, \phi_n) \mathcal{P}_\eta[|m\rangle\langle l|](x'_n, \phi'_n) \rightarrow \langle i|\langle l|R|j\rangle|m\rangle . \quad (8.80)$$

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

$$\mathcal{L}[\rho] = \sum_n \ln \text{tr}[\rho \text{Hom}_\eta(x_n; \phi_n)] . \quad (8.81)$$

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 [30], 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 a bipartite system. A survey on the use of maximum likelihood methods in quantum mechanics is presented in this set of Lecture Notes.

### 8.3.2 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

$$R_z = [D(z) \otimes I] |\Psi\rangle\rangle \langle\langle \Psi | [D^\dagger(z) \otimes I] = (1 - |\xi|^2) |D(z) \xi^{a^\dagger a}\rangle\rangle \langle\langle D(z) \xi^{a^\dagger a} |, \quad (8.82)$$

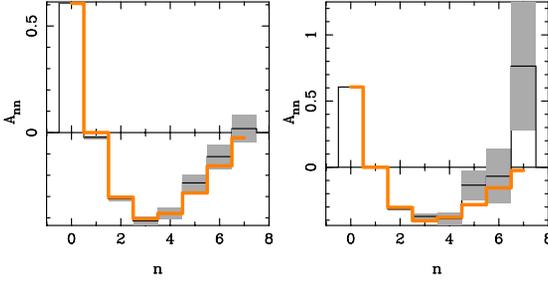
which is then measured with two balanced homodyne setups, one for each mode.

In Fig. 8.3 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

$$A_{nn} = \langle n | D(z) | n \rangle = (1 - |\xi|^2)^{-1/2} \xi^{-n} \frac{\langle n | \langle n | R_z | 0 \rangle | 0 \rangle}{\sqrt{\langle 0 | \langle 0 | R_z | 0 \rangle | 0 \rangle}}, \quad (8.83)$$

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 [29]. Improving quantum efficiency and increasing the amplifier gain (toward a maximally entangled state) 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 [29], 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 measurement of the diagonal matrix elements  $\langle n | \langle m | R | n \rangle | m \rangle$  only 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.



**Fig. 8.3.** From [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 a thin solid line on an extended abscissa range, with their respective error bars in a 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  $\eta$ . Left:  $z = 1$ ,  $\bar{n} = 5$ ,  $\eta = 0.9$ , and  $1.5 \times 10^6$  data have been used. Right:  $z = 1$ ,  $\bar{n} = 3$ ,  $\eta = 0.7$ , and  $6 \times 10^7$  data have been used. The last plot corresponds to the same parameters of the experiment in [29].

### Comments on the Maximum-Likelihood Strategy

The reconstruction can be made much more efficient by ML methods [31–37] also reviewed in these Lecture Notes, 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. (8.36)],  $S_{\mathcal{E}}$  being the operator corresponding to the quantum process under analysis, which is positive and satisfies  $\text{tr}_1 S_{\mathcal{E}} = I$ . The probability distribution of the result  $(x, \phi, x', \phi')$  of a double homodyne detection on the two modes becomes

$$\begin{aligned} \Pr(x, \phi, x', \phi'; S_{\mathcal{E}}) &= \text{tr} [\text{Hom}_{\eta}(x; \phi) \otimes \text{Hom}_{\eta}(x'; \phi') R_{\mathcal{E}}] , \\ &= \text{tr} [\text{Hom}_{\eta}(x; \phi) \otimes (\Psi^* \text{Hom}_{\eta}(x'; \phi') \Psi^T) S_{\mathcal{E}}] . \end{aligned} \quad (8.84)$$

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

$$\mathcal{L}[S_{\mathcal{E}}] = \sum_{n=0}^N \ln [\Pr(x_n, \phi_n, x'_n, \phi'_n; S_{\mathcal{E}})] , \quad (8.85)$$

within the simplex defined by the constraints  $S_{\mathcal{E}} \geq 0$  and  $\text{tr}_1 S_{\mathcal{E}} = I$ . 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.

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

$$F(\boldsymbol{\theta})_{mn} = \left\langle \frac{\partial \ln \Pr(x; \boldsymbol{\theta})}{\partial \theta_m} \frac{\partial \ln \Pr(x; \boldsymbol{\theta})}{\partial \theta_n} \right\rangle_x \tag{8.86}$$

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

$$\Sigma_{mn} = \left\langle (\hat{\theta}_m - \theta_m)(\hat{\theta}_n - \theta_n) \right\rangle_{x_1 \dots x_N} . \tag{8.87}$$

The two matrices satisfy the Cramer-Rao bound

$$\Sigma F(\boldsymbol{\theta}) \geq \frac{1}{N}, \tag{8.88}$$

which puts a limit on the efficiency of the estimation that is independent of the estimator. It is possible to prove that if there exists 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 [38], a much more convenient way to compute the constrained CR bound was presented, based on the distinction between regular points of  $\Theta$  (i.e. the points in the closure of the set of interior points of  $\Theta$ ) and non-regular points. As an example, for  $\Theta$  defined by the constraints  $0 \leq \theta_i \leq 1$ , all the points are regular, whereas for  $\Theta$  equal to a lower dimensional manifold embedded in  $\mathbb{R}^d$  (e. g. a surface defined by some equality constraints) all points are nonregular. The result is that if  $\boldsymbol{\theta}$  is a regular point, then the CR matrix is unaltered, 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 [39], along with a proof

that, 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

$$\Sigma \geq \frac{1}{N} [F^{-1} - F^{-1}G(G^T F^{-1}G)^{-1}G^T F^{-1}] , \quad (8.89)$$

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

For the general problem of quantum process tomography, the likelihood functional  $\mathcal{L}[\mathcal{S}_{\mathcal{E}}]$  of (8.85) is defined for a parameter  $\mathcal{S}_{\mathcal{E}}$  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  $\mathbf{Q}(N)$  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 an infinite size sample. This method—called *sieved maximum likelihood*—has been analyzed in [40] 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, consistently with the fact that the faithfulness of the input state rapidly vanishes for larger photon numbers.

The only downside of the ML approach is the difficulty involved in the maximization of the nonlinear functional in (8.85), which can be tackled either with standard techniques of numerical constrained maximization or with suitable modifications [41] of the iterative algorithms of the kind *expectation-maximization* (EM) for maximum likelihood [30,42]. In practice, several technical problems may arise, as we will discuss concretely in the following example.

### 8.3.3 Homodyne Tomography of an On/Off Detector

In what follows we exploit the ideas of Sect. (8.2.6) 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 Sect. 8.2.6, looking at (8.63), the reduced states of the remaining beam after the measurement will be

$$\rho^{(i)} = \frac{\Psi^T \Pi^{(i)T} \Psi^*}{\text{tr}[\Psi^T \Pi^{(i)T} \Psi^*]} , \quad (8.90)$$

$i$  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.

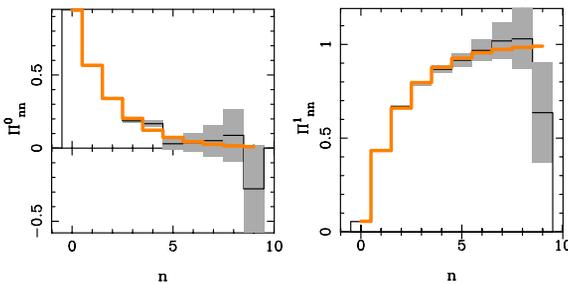
As a model of ON/OFF detector with non unit quantum efficiency and dark current, we will use an ideal ON/OFF photodetector preceded by a beam-splitter of transmissivity  $\tau$  with one port entered by the mode of interest and with the other port fed by a thermal radiation state with mean photon number  $\mu$  [43]. The POVM element for the OFF outcome reads

$$\Pi^{(0)} = \frac{1}{\nu + 1} \sum_{n=0}^{\infty} \left(1 - \frac{\tau}{\nu + 1}\right)^n |n\rangle\langle n|, \quad (8.91)$$

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

### *Reconstruction Using Pattern-Function Averaging Strategy*

The graphs in Fig. 8.4 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 photodetector, for its intrinsic detection mechanism), one can focus the attention



**Fig. 8.4.** Homodyne tomography of an On/Off photo-detector having transmissivity  $\tau = 0.4$  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. 8.3

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 it 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

$$\rho_{nn}^{(i)} = \frac{\sum_m R_{mn} \Pi_{mm}^{(i)}}{\text{tr}[\sum_{mn} R_{mn} \Pi_{mn}^{(i)}]}, \quad (8.92)$$

so that, once measured  $\rho_{nn}^{(i)}$ , it is possible to recover  $\Pi_{mm}^{(i)}$  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.

### *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 purpose 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

$$\text{Hom}_\eta(x, \phi) = \sqrt{\eta} \sum_{j=0}^{\infty} V_j^\dagger e^{i\phi a^\dagger a} |\sqrt{\eta}x\rangle \langle \sqrt{\eta}x| e^{-i\phi a^\dagger a} V_j, \quad (8.93)$$

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  $\eta$  denotes the quantum efficiency of the detectors (this scheme is equivalent to having an ideal homodyne detector preceded by a beam-splitter with transmissivity  $\eta$  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 (8.93), which yields a diagonal POVM  $\text{Hom}_\eta(x)$  (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 photo-counter signaling outcome  $i$  and the homodyne measuring  $x$ , is given by

$$\begin{aligned} \Pr(i, x; \Pi) &= \text{tr}[R \Pi^{(i)} \otimes \text{Hom}_\eta(x)] = \sum_{mn} R_{mn} \Pi_{mm}^{(i)} \text{Hom}_{\eta mn}(x) = \\ &= \sum_m \Pi_{mm}^{(i)} A_m(x), \end{aligned} \tag{8.94}$$

where  $R$  denotes twin-beam state, and the positive coefficients  $A_m(x)$  are defined as

$$A_m(x) = \sqrt{\eta} \sum_{n \geq h} R_{mn} \binom{n}{h} \eta^h (1 - \eta)^{n-h} \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \frac{e^{-2\eta x^2}}{2^h h!} H_h^2(\sqrt{2\eta}x) > 0. \tag{8.95}$$

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

$$\mathcal{L}[\Pi] = \sum_l \ln \left[ \sum_m \Pi_{mm}^{(i_l)} \cdot A_m(x_l) \right], \tag{8.96}$$

with  $\Pi$  restricted to the simplex of diagonal POVM’s.

First, one must choose the dimension of the subspace on which one is 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 nonlinear convex programming problem, and can be faced with several different kind of algorithms as the *simplex method* (see contribution of D’Ariano *et al.* to these Lecture Notes), or the methods of *sequential quadratic programming* (SQP), or the methods of *expectation-maximization* (EM) type (see chapter by Hradil *et al.* in these Lecture Notes) that can be easily impemented in this particular example since, for fixed  $m$  and varying  $i$ , the numbers  $\Pi_m^{(i)}$  define a probability distribution. 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  $\theta$  defining  $\Pi$  can be very small, so that very different values of the parameters will give almost the same likelihood, thus making it 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.

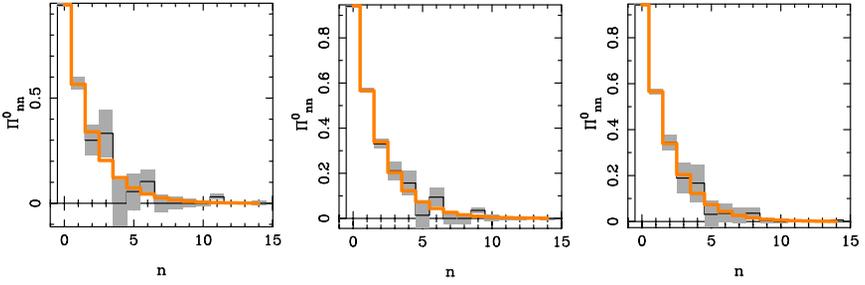
Notice that the Fisher information matrix (8.86) for the probability distribution  $\Pr(i, x; \Pi)$  can be expressed in terms of the expectation value of the derivatives of the likelihood with respect to the independent parameters  $\theta_m$  defining  $\Pi$

$$F(\Pi)_{mn} = \frac{1}{N} \left\langle \frac{\partial \mathcal{L}[\Pi]}{\partial \theta_m} \frac{\partial \mathcal{L}[\Pi]}{\partial \theta_n} \right\rangle_{(i_1, x_1) \dots (i_N, x_N)}, \quad (8.97)$$

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 (8.88). This, however, needs the calculation of the Fisher information matrix in correspondence of the unknown true value of  $\Pi$ , 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 a 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 [44], 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 of 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.

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 (8.94), for the POVM model presented in (8.91), with the same parameters as Fig. 8.4, and various values of the quantum efficiency  $\eta$ . The detector POVM has been estimated with the maximum likelihood method, with the only hypothesis of diagonal POVM, and putting the dimensional 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



**Fig. 8.5.** 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  $\eta$ . Left:  $N = 10^5$ ,  $\eta = 0.7$ ; Middle:  $N = 10^4$ ,  $\eta = 0.9$ ; Right:  $N = 10^6$ ,  $\eta = 0.7$ .

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. 8.5 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. 8.4 evidences the much higher efficiency of maximum likelihood reconstruction. The graph on the left shows how the 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, however, 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.

The maximization has been performed numerically by the routine `donlp2` [45] that implements an SQP algorithm, and then the self-consistency of the solution has been checked by means of a few EM type iterations applied alternatively to the probability distributions corresponding to the elements  $\Pi_m^{(i)}$ , for fixed  $m$ . Of course, it would be much easier to employ the only EM algorithm, being a recursive application of an easily implementable iteration. 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 well the theoretical POVM. This paradoxical behavior is an artifact due to the smoothness of the theoretical curve of the considered model, as one can easily check by changing the model.

## Acknowledgements

This work has been sponsored by INFN through the project PRA-2002-CLON, and by EEC and MIUR through the cosponsored ATESIT project IST-2000-29681 and Cofinanziamento 2002.

## References

1. W. F. Stinespring: Proc. Am. Math. Soc. **6**, 211 (1955).
2. K. Kraus: Ann. Phys. **64**, 311 (1971).
3. I. L. Chuang, M. A. Nielsen: J. Mod. Opt. **44**, 2455 (1997).
4. J. F. Poyatos, J. I. Cirac, P. Zoller: Phys. Rev. Lett. **78**, 390 (1997).
5. M. A. Nielsen, E. Knill, R. Laflamme: Nature **396**, 52 (1998).
6. A. M. Childs, I. L. Chuang, D. W. Leung: Phys. Rev. A, 012314 (2001).
7. N. Boulant, T. F. Havel, M. A. Pravia, D. G. Cory: Phys. Rev. A **67**, 042322 (2003).
8. P. Kwiat, J. Altepeter, D. Branning, E. Jeffrey, N. Peters, T. C. Weh: Taming entanglement. In: *Proceedings of the 6th International Conference on Quantum Communications, Measurement and Computing*, ed by J. Shapiro, J. and O. Hirota (Rinton Press, Princeton, NJ, 2003).
9. M. W. Mitchell, C. W. Ellenor, S. Schneider, A. M. Steinberg: Phys. Rev. Lett. **91**, 120402 (2003).
10. G. M. D'Ariano, P. Lo Presti: Phys. Rev. Lett. **86**, 4195 (2001).
11. D. Leung: Towards robust quantum computation, PhD thesis, Stanford University (2001).
12. A. Jamiolkowski: Rep. Math. Phys. **3**, 275 (1972).
13. M.-D. Choi: Linear Algebra Appl. **10**, 285 (1975).
14. F. De Martini, A. Mazzei, M. Ricci, G. M. D'Ariano: Phys. Rev. A **67**, 062307 (2003).
15. J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, P. G. Kwiat, R. T. Thew, J. L. O'Brien, M. A. Nielsen, A. G. White: Phys. Rev. Lett. **90**, 193601 (2003).
16. A. G. White, A. Gilchrist, G. J. Pryde, J. L. O'Brien, M. J. Bremner, N. K. Langford, (2003): arXiv:quant-ph/0308115 (2003).
17. G. M. D'Ariano, P. Lo Presti: Phys. Rev. Lett. **91**, 047902 (2003).
18. G. M. D'Ariano, P. Lo Presti, M. G. A. Paris: Phys. Rev. Lett. **87**, 270404 (2001).
19. K. Kraus: *States, Effects, and Operations* (Springer, Berlin, 1983).
20. A. S. Holevo: *Probabilistic and statistical aspects of quantum theory*, vol 1 of *Series in Statistics and Probability* (North-Holland, Amsterdam, New York, Oxford, 1982).
21. G. M. D'Ariano, P. Lo Presti: unpublished (2003).

22. I. L. Chuang, M. A. Nielsen: *Quantum Information and Quantum Computation* (Cambridge University Press, Cambridge UK, 2000).
23. A. Luis: Phys. Rev. A **62**, 054302 (2000).
24. R. B. Bapat: *Linear algebra and linear models* (Springer, Berlin, 2000).
25. M. J. W. Hall: (1994) *Phys. Rev. A* **50**, 3295 (1994).
26. R. Simon: Phys. Rev. Lett. **84**, 2726 (2000).
27. M. G. A. Paris: J. Opt. B: Quantum Semiclass. Opt. **4**, 442 (2002).
28. G. M. D'Ariano, N. Sterpi: J. Mod. Opt. **44**, 2227 (1997).
29. M. Vasilyev, S.-K. Choi, P. Kumar, G. M. D'Ariano: Phys. Rev. Lett. **84**, 2354 (2000).
30. Y. Vardi, D. Lee: J. R. Statist. Soc. B **55**, 569 (1993).
31. Z. Hradil: Phys. Rev. A **55**, R1561 (1997).
32. Z. Hradil, J. Summhammer, G. Badurek, H. Rauch: Phys. Rev. A **62**, 014101 (2000).
33. J. Fiurášek, Z. Hradil: Phys. Rev. A **63**, 020101 (2001).
34. K. Banaszek, G. M. D'Ariano, M. G. A. Paris, M. F. Sacchi: Phys. Rev. A **61**, R010304 (2000).
35. M. G. A. Paris, G. M. D'Ariano, M. F. Sacchi: Maximum-likelihood method in quantum estimation. In: *Bayesian inference and maximum entropy methods in science and engineering*, vol 568 of *AIP Conf. Proc.* p 456; preprint arXiv:quant-ph/0101071 (2001).
36. M. F. Sacchi: Phys. Rev. A **64**, 022106 (2001).
37. M. F. Sacchi: Phys. Rev. A **63**, 054104 (2001).
38. J. D. Gorman, A. O. Hero: IEEE Trans. Inf. Th. **26**, 1285 (1990).
39. T. L. Marzetta: IEEE Trans. on Signal Processing **41**, 2247 (1993).
40. R. Gill, M. I. Guță: arXiv:quant-ph/0303020 (2003).
41. J. Rehacek, Z. Hradil, M. Jezek: Phys. Rev. A **63**, R040303 (2001).
42. A. P. Dempster, N. M. Laird, D. B. Rubin, J. R. Statist. Soc. B **39**, 1 (1977).
43. M. G. A. Paris: Phys. Lett. A **489**, 167 (2001).
44. B. Efron, R. Tibshirani: *An Introduction to the Bootstrap* (Chapman and Hall, New York, 1994).
45. P. Spellucci: Math. Prog. **82**, 413 (1998).