A thesis submitted in partial fulfilment

of the requirements of the degree of

Doctor of Philosophy

Applications of the operator Schmidt decomposition in the quantification and exploitation of correlations

Matteo Caiaffa

University of Strathclyde

Department Physics

October 11, 2018

Acknowledgements

I am greatly thankful to Marco, for his guidance was constant. I will benefit from our countless discussions on the Method hereafter. His passion and care, together with our newly discovered closeness, brought us to bond far beyond any mere professional affair. I cannot forget Luca, his attitude towards life eased mine, and it always reminded me of what it does matter and what it does not. Thanks to friends and colleagues from the seventh floor, especially to whom I do not even need to name. Lastly, I salute Glasgow, which 'gave me more than it ever took away'.[1]

To Mom and Dad, thanks for paving the way of whatever road I have walked. To my Brother, for always destroying the boundaries of my ambition. To Roberto and Enrico, the foundations of a heretofore unrecognised consciousness. To my Girl, for her love and strength, for our sacrifices. To our life.

Edit: I have often learned new things during exams. This thesis defense was not to be outdone, thanks to the support and enlightening suggestions of the committee members. I owe a debt of gratitude to the Convenor of the Viva Professor Brian McNeil and to the examiners Doctor Paul Griffin and Professor Gerardo Adesso.

[1] Billy Connolly, *I wish I was in Glasgow*, 2012.

Abstract

In this thesis we analyse a particular decomposition of the density matrix into tensor product terms – known as the operator Schmidt decomposition (OSD) – showing how it can be used in order to measure and exploit correlations in bipartite quantum systems. Correlations rest at the heart of Quantum Information theory for both their foundational significance and their irreplaceable role in quantum computation and communication. However, because of their difficult characterisation, detecting and measuring correlations are usually believed arduous tasks. This is particularly true in the mixed-state domain, where the diversity of potential correlations represents a further complication. For these reasons, it would be advisable to define a common framework for examining and quantifying correlations of all kinds and degrees, both for pure and mixed states. Here we argue that the OSD is a powerful tool for this purpose, in that it can be used to devise measures of correlations, whether classical or quantum. In turn, these measures can be exploited in order to detect the presence of entanglement and steering, for example. The first part of this work is devoted to the definitions of such measures and the analysis of their properties. In the second part instead we consider the possibility of taking advantage of the OSD in the context of quantum process discrimination and tomography. These tasks are central to the implementation of quantum technologies, since the actual realisation of any application based upon quantum phenomena largely relies on the determination of quantum processes. We provide a set of tools – based on the OSD – that could serve as a means by which enabling or improving certain specific protocols of ancilla-assisted quantum process discrimination (AAPD) and tomography (AAPT). First, we present a quantifier for the performance of bipartite input states in AAPD. We show that the possibility to improve the discrimination power of this protocol – or to enable it altogether – is imprinted in the OSD of the input state. For what AAPT is concerned, we demonstrate how the OSD of the input state can be exploited in order to allow a characterization of an unknown local channel via a relatively small number of local transformations of the input state. More in general, we provide several results which show a sharp connection between the tasks of channel discrimination and tomography, the OSD of the input state, and the degree of correlations carried by the latter. We conclude the thesis with a collection of results of a seemingly different nature, but that were actually inspired by the examination of the previously mentioned ancilla-assisted tasks. In short, we define a family of state-dependent metrics on the space of quantum channels and show that they are deeply connected to the OSD of the state defining them. As a byproduct, the latest results entail a possible generalisation of the Choi–Jamiołkowski isomorphism, thus providing an interesting motivation to the extension of this research project.

Contents

List of papers

The research project pursued over the last three years has produced four papers, one of which has been published, one submitted and two are in preparation. The contents of the papers listed below are presented in Chapters 4, 5, 3 and 6, respectively.

[1] — M. Caiaffa and M. Piani, "Channel discrimination power of bipartite quantum states", *Physical Review A*, vol. 97, no. 3, p. 032334, 2018.

[2] — M. Caiaffa and M. Piani, "Correlation-assisted process tomography", submitted to Physical Review A on September 7 2018, *arXiv preprint arXiv:1808:10835*, 2018.

[3] — M. Caiaffa and M. Piani, "Measuring correlations via the operator Schmidt decomposition", Provisional title, paper in preparation.

[4] — M. Caiaffa and M. Piani, "Metrics and pseudometrics spaces of quantum channels", Provisional title, paper in preparation.

List of figures

Notation and acronyms

$\mathcal{H}_A$ — Hilbert space $\mathcal{H}$ associated to the system A, with elements (and operations between them) denoted in the usual Dirac notation

$\mathcal{H}_{AB} \equiv \mathcal{H}_A \otimes \mathcal{H}_B$ — tensor product of the Hilbert spaces associated to A and B, respectively, corresponding to the (pure) state space of the compound system AB

$\mathcal{L}(\mathcal{H}_A)$ — space of linear operators on $\mathcal{H}_A$

$\mathcal{L}(\mathcal{H}_{AB})$ — space of bipartite linear operators on $\mathcal{H}_{AB}$

$\mathcal{T}(\mathcal{H})$ — set of trace class operators on $\mathcal{H}$

$\mathcal{S}(\mathcal{H})$ — set of positive semidefinite, unite trace, linear operators on $\mathcal{H}$, with element denoted by lower case Greek letters

$\mathcal{S}_A \equiv \mathcal{S}(\mathcal{H}_A)$ — state space of the system A

$\mathcal{S}_{AB}$ — state space of the compound system AB

$\Lambda[\,\cdot\,]$ — capital Greek letters usually denotes quantum channels; i.e. completely positive, trace preserving linear mappings $\Lambda : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$

$\mathcal{C}(\mathcal{H}, \mathcal{H}')$ — space of quantum channels from $\mathcal{H}$ to $\mathcal{H}'$

id_A — identity map acting on $\mathcal{L}(\mathcal{H}_A)$

$\mathscr{L}$ — matrix representation of a channel $\Lambda[\,\cdot\,]$

$\|\cdot\|_p$ — Schatten p-norm of an operator

Notation and acronyms

AAPT — ancilla-assisted quantum process tomography

CAPT — correlation-assisted quantum process tomography

CCN — computable cross norm

CDP(ρ_{AB}) — channel discrimination power of the bipartite state ρ_{AB}

ESP — elementary symmetric polynomials

HS — Hilbert-Schmidt

LHV — local hidden variable

LO — local operation

LOCC — local operations and classical communication

MPDO — matrix product density operator

MPS — matrix product state

OGC — operator G-Concurrence

OSC — operator Schmidt coefficient

OSD — operator Schmidt decomposition

OSR — operator Schmidt rank

QMI — quantum mutual information

RC — realignment criterion

SPT — standard quantum process tomography

TC — total correlations

Introduction

The novelty and significance of the phenomena unveiled by Quantum Mechanics at the beginning of the last century have since prompted the interest of many scientists in dissecting the very founding elements of the theory, namely the states of a physical system, the correlations they possess, the measurement process and the rules governing the physical transformations in the quantum regime. After the awe for the conceptual revolution cooled down, and once the mathematical structure of the theory was given a solid and final ground, the possibility of exploiting the newly discovered features of nature with the purpose of conceiving technologies that would allegedly enhance our lives came up. The promise of impressive advances in the field of computation [5–8] and communication [9–11] was cast, encouraging the advent of new, profitable technologies [12]. The variety and potential usefulness of such applications, together with the early results supporting said promises, made Quantum Information science one of the most prolific research areas – and a cornerstone indeed – of contemporary Physics [13, 14]. Paramount to the effective realisation of these advancements is the study of correlations in quantum systems, as they underpin the whole library of algorithms and protocols of Quantum Information. Another key factor enabling the aforementioned applications is the ability to tell apart two or more physical processes, and to reconstruct the action of an unknown process by looking at its action on a set of quantum states [15, 16]. Indeed, quantum technologies rely extensively on the determination of quantum channels [15–17], as well as on probing their actual implementation versus their ideal description [18]. The study of correlations in quantum states and the characterization of quantum processes – which are the motivations at the heart of this theoretical thesis – will hopefully serve as a catalyst for the efficient implementation of certain specific tasks of Quantum Information. Our work boils down to two questions, which, in a way, divide the thesis in two parts. The first concerns the problem of measuring classical and quantum correlations in quantum systems, while the second, to a certain extent, is about using them. The main tool at our disposal – and actually the central object of our investigation – is the operator Schmidt decomposition (OSD), which is nothing but a generalisation of the ordinary Schmidt decomposition of state vectors to the case of bipartite density operators. Specifically, the OSD is a particular factorization of the density matrix into tensor product terms. As is well–known, the Schmidt decomposition characterises the set of entangled states in the pure case. Along the same line, and due to the similarity between the ordinary and the operator Schmidt decomposition at the algebraic level, one could wonder if the OSD can be used to make sense of correlations in mixed states as well. This might not be as direct as it looks, as mixed states exhibit several kinds of correlations, while pure states only comes in two flavours: either they are separable, or they are not. Mixed states can be classically or quantum correlated instead, or both at the same time. In their turn, quantum correlations can be classified through different categories, the latter drawing lines between steerable and non–steerable states, local and non–local states, and between bipartite states which are classical or not with respect to their subsystems, namely states with zero or non–zero discord. Trying to exploit the OSD in order to devise meaningful and sensible measures of correlations, whether they be classical or quantum, was the questions which originated this project in the first place. As it turned out, this is possible indeed: we are able to define several measures of total correlations based on the OSD, to study their property and to discuss their application to entanglement and steering detection. In the second part of this work we put into practice what we have learnt about the OSD and its relations to correlations to tackle the tasks of discrimination and tomography of quantum processes [1]. To be precise, the thesis is organized as follows.

In Chapter 1 we lay down the mathematical framework and the necessary notation needed to formulate the state of the art and the results of the research that we have pursued over the last three years. Special attention will be given to the definition of the different kinds of correlations in the classical and quantum settings, and to several useful representations of trace–preserving and completely–positive linear maps.

Chapter 2 is the literature review. We present several known results about the OSD and its application to entanglement detection, and we rederive those theorems that looked closer in spirit to the aims of our research project. We dedicate particular care to the well–known realignment criterion for separability from different perspective, including to its formulation in terms of the coefficients of the OSD, known as operator Schmidt coefficients (OSC, to be defined later). Also, the appearance of the OSD in the field of tensor networks is discussed.

Chapter 3, except for the first introductory sections, is the first original chapter. Here we delineate in an axiomatic fashion the requirements that a measure of total correlations must satisfy in order to be meaningful. Then, we define several measures. The latter are based, respectively, on the operator Schmidt rank (OSR, that is the number of terms in the OSD), on the OSC, on the distance between the square root of a quantum states and the product of the square root of its marginals, and on the OSD of the square root of a given state. We analyse their properties and their application to the certification of correlations. Lastly, we provide a relational expression between some of of the newly introduced quantities and the quantum mutual information. We also discuss the subject of vector majorization and its connection to the monotonicity of the elementary symmetric polynomial in the OSC. The content of this chapter is part of [3].

In Chapter 4 we present the result contained in our paper [1]. This is about what we have called channel discrimination power (CDP) of bipartite quantum states, and it concerns the definition and analysis of a worst–case scenario quantifier for the performance of a probe–ancilla state in channel discrimination. We provide general upper and lower bounds to the CDP of a state in terms of its OSD, and we compute the CDP of pure states exactly. Remarkably, we show that also correlated but separable states can have non-zero CDP, as long as they posses a certain amount of discord. More in general, we derive a non-trivial bound on the CDP of any state that passes the realignment criterion for separability.

In Chapter 5 we report the content of our second paper [2]. Here we prove that the correlations of a fixed bipartite state, of whatever degree and as measured by the logarithm of the OSR, can in principle be exploited to allow process tomography. In particular, we show that any single bipartite input state can be used to perform channel tomography through the protocol of ancilla–assisted quantum process tomography (AAPT). Indeed, even if an input state is not suitable for AAPT, it can be transformed through local operations, in order to obtain a set of states which altogether provide full tomography of any unknown channel. We argue that the number of local channels allowing channel tomography depends upon the correlations of the initial input state. For mixed states we provide examples showing how the presence of correlations dramatically reduces the number of local channels, when the latter are given by unitary operators. In particular, for a two-qubit state we find that discord is necessary and sufficient in order to attain channel tomography with less unitary than the ones required by SQPT. For pure states we provide the optimal number of such local transformations. We conclude the chapter showing how our protocol can be used to enhance the accuracy of channel tomography.

Chapter 6 concerns a possible genaralisation of the Choi-Jamiołkowski isomorphism, and it is based on the observation (already highlighted in the previous chapter) that in order to obtain a one–to–one mapping between bipartite operators and linear maps acting on single–system operators, the bipartite operator must have maximal OSR. We show how these state-dependent isomorphisms give rise to state-dependent metrics on the space of quantum channels. On the other hand, states with non maximal OSR cannot be bijectively associated to quantum channels. In turn, such states induces pseudometrics, which can be lifted to proper metrics through the introduction of particular equivalence classes of quantum channels. The main result of this chapter is to show that two equivalence classes of quantum channels induced by two different bipartite states are equal if and only if the local operators in the OSD of the latter generate the same subspaces. The content of this chapter is part of [4].

Chapter 1

Elements of quantum information

This chapter introduces the basic concepts and the notation adopted in the entire thesis. We first present the mathematical framework of quantum information, paying particular attention to the definition of correlations in quantum states. Then we introduce the mathematical representatives of quantum physical processes, together with several ways of expressing them. Finally, in the last section we recall the definition of Schatten $p-$norms and superoperator norms, which will be used often times in the thesis.

1.1 States and measurements

In quantum mechanics (QM) every physical system S is associated to a complex separable Hilbert space $\mathcal{H}$, known as the state space of the system [19]. We are interested in finite-dimensional systems, then each Hilbert space $\mathcal{H}$ is isomorphic to $\mathbb{C}^d$ for some finite dimension d, where $\mathbb{C}$ is the complex line. For every $|\psi\rangle, |\varphi\rangle \in \mathcal{H}$, we denote by $\langle\psi|\varphi\rangle$ the inner product and by $\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$ the induced norm on $\mathcal{H}$. In the original formulation of QM the states of a quantum system S, i.e. the *pure states* of S, are represented by unit length vectors of $\mathcal{H}$ [19–21]. In the modern interpretation of the theory [22–26] pure states correspond to *ensembles* of equally prepared quantum systems, rather than to individual ones. To be precise, pure states represent equivalence classes of preparation procedures, and in this framework the *statistical operator* (also referred to as density operator or density matrix) represents the more general and adequate description of the possible states of a quantum system. In order to define the set of statistical operators we first introduce the set $\mathcal{T}(\mathcal{H})$ of trace–class operators, which is a subset of the set $\mathcal{L}(\mathcal{H})$ of all linear operators on $\mathcal{H}$. A linear operator $\chi \in \mathcal{L}(\mathcal{H})$ belongs to $\mathcal{T}(\mathcal{H})$ if the trace of the absolute value $|\chi| = (\chi^\dagger \chi)^{1/2}$ is finite, where the trace $\mathrm{Tr}[\,\cdot\,]$ of an operator is defined as the linear mapping

$$\mathrm{Tr}[\chi] := \sum_i \langle u_i|\chi|u_i\rangle, \tag{1.1}$$

with $\{|u_i\rangle\}$ being any orthonormal basis of $\mathcal{H}$. Also notice that when $d < \infty$ the definition of trace of an operator coincides with that of a matrix, hence every finite-dimensional operator is trace class [27].

The set $\mathcal{S}(\mathcal{H})$ of statistical operators on $\mathcal{H}$ is that subset of $\mathcal{T}(\mathcal{H})$ made of positive semi-definite, unit trace operators:

$$\mathcal{S}(\mathcal{H}) = \{\rho \in \mathcal{T}(\mathcal{H}) \mid \rho \geq 0, \mathrm{Tr}[\rho] = 1\}. \tag{1.2}$$

The states space $\mathcal{S}(\mathcal{H})$ is convex, i.e. given a set of operators $\rho_k \in \mathcal{S}(\mathcal{H})$ and non-negative numbers λ_k with $\sum_k \lambda_k = 1$, also $\sum_k \lambda_k \rho_k \in \mathcal{S}(\mathcal{H})$. The extremal points of $\mathcal{S}(\mathcal{H})$ are one-dimensional projectors of the form $\rho = |\psi\rangle\langle\psi|$, with $|\psi\rangle \in \mathcal{H}$ a pure state. We refer to all the other states of $\mathcal{S}(\mathcal{H})$ as *mixed states*. Every mixed state $\rho \in \mathcal{S}(\mathcal{H})$ admits a *spectral decomposition*

$$\rho = \sum_k p_k \, |\psi_k\rangle\langle\psi_k| \tag{1.3}$$

where the collection $\{|\psi_k\rangle\}$ forms an orthogonal basis of $\mathcal{H}$ with $\langle\psi_k|\psi_l\rangle = \delta_{kl}$ and $\{p_k\}$ is a probability distribution. In other words, the mixed state in Eq. (1.3) represents an ensemble of a large number N of quantum systems, $\approx N p_k$ of which have been prepared in the pure state $|\psi_k\rangle$.

If, on the one hand, quantum states are associated to preparation procedures, observables corresponds to registration procedures, usually associated to some macroscopic experimental apparatus able to measure the value of a definite quantity. Their mathematical counterpart is given by positive operator-valued measures (POVMs) [13, 28]. Let Ω be a set of possible outcome of a measurement, and let $\sigma(\Omega)$ be a $\sigma-$algebra over Ω. A POVM is defined as the map sending any element of $\sigma(\Omega)$ into a positive (bounded) operator:

$$\begin{aligned} E(\,\cdot\,) : \sigma(\Omega) &\to \mathcal{L}(\mathcal{H}) \\ Z &\mapsto E(Z). \end{aligned} \tag{1.4}$$

Notice that a POVM satisfies the relation $0 \leq E(Z) \leq 1$, where it is understood that $E(\emptyset) = 0$ and $E(\Omega) = \mathbb{1}$. In addition, POVMs are completely additive on disjoint subset:

$$E(\cup_i Z_i) = \sum_i E(Z_i), \quad \text{if } Z_i \cap Z_j = \emptyset \text{ for } i \neq j. \tag{1.5}$$

Given a quantum state ρ, the probability that the quantity described by E takes on values in the subset Z is given by

$$p(E, Z) = \mathrm{Tr}[\rho E(Z)], \tag{1.6}$$

which is a classical probability measure, thanks to the properties of POVMs.

Finally, Eq. (1.6) and the spectral theorem yield the usual formula for computing the mean value of an observable associated to an Hermitian operator A when the system of interest is in the state ρ, thus recovering the Born rule

$$\langle A \rangle = \mathrm{Tr}[\rho A]. \tag{1.7}$$

1.2 Correlations in composite quantum systems

Consider two physical systems S_A and S_B associated with the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Then, the composite system S_{AB} will be associated to the tensor product $\mathcal{H}_{AB} := \mathcal{H}_A \otimes \mathcal{H}_B$[1], whose dimension is given by the product of the individual

[1]This axiom is usually substantiated [19, 29, 30] by the observation that if $\{|\psi\rangle_i^A\}$ and $\{|\varphi\rangle_j^B\}$ are orthonormal basis of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, then any state of $\mathcal{H}_{AB}$ can be expressed in the product

dimensions, i.e. $d_{AB} = d_A \cdot d_B$ (we indicate by d_X the dimension of the subsystem corresponding to $\mathcal{H}_X$). In general, a composite system $S_{AB\ldots N}$ is associated with the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \cdots \mathcal{H}_N$ of the individual Hilbert spaces. We focus on bipartite systems (i.e. systems comprising two subsystems) S_{AB} and, unless otherwise stated, we employ the notation $d = \min\{d_A, d_B\}$. From now on, if there is no possibility of confusion we write $\mathcal{S}_X$ in place of $\mathcal{S}(\mathcal{H}_X)$ and $\mathcal{S}_{XY}$ in place of $\mathcal{S}(\mathcal{H}_X \otimes \mathcal{H}_Y)$.

States in $\mathcal{S}_{AB}$ given by the tensor product

$$\rho_{AB} = \rho_A \otimes \rho_B \tag{1.8}$$

with $\rho_A \in \mathcal{S}_A$ and $\rho_B \in \mathcal{S}_B$ are called *product states.* They represent the simplest kind of state for composite quantum systems, portraying the physical situation of completely uncorrelated subsystems. The absence of statistical dependence between A and B means that product states can be obtained by independent preparation procedure on the two subsystems, and that probabilities of the measurement outcomes of independent experiments will always factorize [31]. In the language used throughout the thesis, we say that product states can be prepared (or transformed to other product states) by the means of *local operations* (LO) on the subsystems. If, besides LO, we also allow classical communication in the preparation of A and B, then we are able to introduce classical correlations between the subsystems. This preparation protocol is referred to as *local operations and classical communication* (LOCC), and states prepared (and transformed into one another) using said prescriptions are called *separable states.* Formally, they are convex combination of product states:

$$\rho_{AB} = \sum_i \lambda_i \rho_A^i \otimes \rho_B^i, \tag{1.9}$$

with $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$. In the light of the definition above we are in a position to introduce one of the main elements of QM: entanglement [32–36]. We say that any state that cannot be expressed as Eq. (1.9) is an *entangled state.* As known,

basis given by $\{|\psi\rangle_i^A \otimes |\varphi\rangle_j^B\}$.

entanglement has a prominent status in Physics, being one of those peculiar features of QM revealing an actual departure from classical mechanics [32, 37]. Along with its foundational relevance, entanglement is a central resource in quantum information and computation [5], from quantum teleportation [10] to superdense coding [9], as well as it enables a plethora of applications in quantum comminication [11]. As such, the set of entangled states is subject to a profound investigation. Nonetheless, despite the clear definition Eq. (1.9), it is generally difficult to determine if a given mixed states is indeed separable or not [38]. For pure state however, entanglement can be fully characterized. This result is achieved thanks to the renowned *Schmidt decomposition* theorem, which says that any state vector $|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ admits the decomposition [13]

$$|\psi\rangle_{AB} = \sum_{i=1}^{\mathrm{SR}(\psi)} \sqrt{p_i}\, |a_i\rangle_A \otimes |b_i\rangle_B \,, \tag{1.10}$$

where the $\{\sqrt{p_i}\}$ are positive numbers satisfying $\sum_{i=1}^{\mathrm{SR}(\psi)} (\sqrt{p_i})^2 = \sum_{i=1}^{\mathrm{SR}(\psi)} p_i = 1$, and $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$ are some special and $|\psi\rangle$-dependent orthonormal bases for $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. As customary goes, the elements of the probability distribution $\{p_i\}$ are usually taken to be ordered without loss of generality, so that we will assume $p_1 \geq p_2 \geq \ldots$. Here, $\mathrm{SR}(\psi)$ denotes the Schmidt rank of $|\psi\rangle_{AB}$, which is the number of non-zero p_i's, and satisfies $\mathrm{SR}(\psi) \leq \min\{d_A, d_B\}$. Then, as well known, $|\psi\rangle_{AB}$ is entangled if and only if $\mathrm{SR}(\psi) > 1$. This result reveals the importance of the Schmidt decomposition in QM, and has no analogue in the mixed states domain. Indeed, as we shall see in details in the remainder of the thesis, despite the fact that the Schmidt decomposition can be extended to the density operator formalism, the connection with mixed-state entanglement is not so striking as for pure states.

So far we have partitioned the set of quantum states in the two broad classes of separable and entangled ones. The former (strictly) contains the set of bipartite states whose correlations between the parties are classical. The latter strictly contains the subset of states whose correlations cannot be reproduced, loosely speaking, by any local realistic model.[2] As a matter of fact the situation is quite more involved, and one

[2]There are entangled states that are local in the sense that they do not give raise to any non–locality

recognises several degrees of *quantumness*. A finer classification of quantum correlations – initiated by the observations of Einstein, Podolsky and Rosen in the famous paper [37] – can be carried out by looking at the effects of local measurements on a given systems. We consider this alternative viewpoint in the next section.

1.2.1 Steering and nonlocality

Here we classify the various bipartite quantum correlations by considering the effects of local measurements performed on the components of a given bipartite system. To see how this can be accomplished, we begin with redefining the concept of entangled states in this framework. Let $\mathcal{S}_{AB}$ be the states space of a given system and denote with $\mathcal{D}_A$ the set of all observables on $\mathcal{H}_A$, i.e. the set of Hermitian operators acting on the state of the quantum system. For any observable $O_A \in \mathcal{D}_A$, the set of its eigenvalues will be denoted by $\{o_A\}$. Moreover, $P(o_A|O_A;\rho_{AB})$ indicates the probability that an observer in A will obtain the result o_A when measuring the observable O_A on a system in the state $\rangle$. Lastly, we denote with $M_A \subset \mathcal{D}_A$ the set of all measurements that Alice can perform. The same notation is used for the respective quantities on B, and we make use of the distant laboratories paradigm with Alice and Bob as observers.

We say that a state ρ_{AB} is separable if and only if for any $o_A \in \{o_A\}, o_B \in \{o_B\}$, and for any $O_A \in M_A, O_B \in M_B$, the following equation

$$P(o_A, o_B|O_A, O_B;\rho_{AB}) = \sum_\lambda p_\lambda \, P(o_A|O_A;\rho^A_\lambda)P(o_B|O_B;\rho^B_\lambda) \tag{1.11}$$

is satisfied. Here $\rho^A_\lambda, \rho^B_\lambda$ are quantum states on A, B respectively and p_λ is some probability distribution which involves the *local hidden variable* (LHV) λ, namely a probability distribution which is consistent with a LHV model [39].

On the same ground we can characterize states exhibiting Bell nonlocality [40]. We say that a state ρ_{AB} is Bell local if and only there exist $O_A \in M_A, O_B \in M_B$ and

(the latter intended as non–locality of joint conditional probability distributions). Werner states were introduced in [31] exactly to prove the difference between non–locality and entanglement.

$o_A \in \{o_A\}, o_B \in \{o_B\}$ such that the following equation

$$P(o_A, o_B|O_A, O_B; \rho_{AB}) = \sum_\lambda p_\lambda\, p(o_A|O_A; \lambda) p(o_B|O_B; \lambda) \tag{1.12}$$

is satisfied. Notice that here also $p(o_A|O_A; \lambda)$ and $p(o_B|O_B; \lambda)$ are probability distributions depending on the LHV λ . In other words, a state ρ_{AB} is Bell nonlocal if and only if the correlations between o_A and o_B cannot be justified by an underlying LHV model, i.e. if there exists a set of measurements $M_A \times M_B$ for which Eq. (1.12) is falsified.

To introduce the concept of steering, let us consider the following quantum information task. Alice can prepare a bipartite system and send a part to Bob, and repeats this procedure many times. At each round, they can measure the respective subsystems and communicate in a classical way. Alice's task is to convince Bob that she can prepare an entangled state, while Bob will not be convinced as long as there exists a particular classical model explaining the nature of the correlations between the outcomes of the respectively measured quantities. Such a model is called a local hidden state (LHS) model for Bob, and it consists of any local theory able to explain the correlations between the parties, whilst maintaining that Bob's system has a definite state (even if unknown to him). On the other hand, if the correlation cannot be explained by any such model, that would mean that Alice was able to create genuine entangled states, and thus to *steer* Bob's state with her measurements. This is the main idea behind the operational definition of steering; we refer to [36] for a detailed exposition of the task. For the purposes of this thesis, we settle for the formal characterization of steering, which can be formulated in the language of Eqs. (1.11) and (1.12). We say that the set of measurements $M_A \subset \mathcal{D}_A$ on a state ρ_{AB} exhibits steering if and only if there exist $O_A \in M_A, O_B \in M_B$ and $o_A \in \{O_A\}, o_B \in \{O_B\}$ such that the following equation

$$P(o_A, o_B|O_A, O_B; \rho_{AB}) = \sum_\lambda p_\lambda\, p(o_A|O_A; \lambda) P(o_B|O_B; \rho_\lambda^B) \tag{1.13}$$

is falsified. If such a subset M_A – called Alice's *strategy* – exists, then we say that ρ_{AB} is steerable by Alice.

1.2.2 Quantum discord

Quantum discord is historically the last refinement of the hierarchy of correlations of quantum states, and was introduced independently in [41, 42] and [43].

A bipartite state is said to be *classical on* A (or to have zero discord on A) if there exists a local basis on A in which an experimenter could perform measurements without modifying the state. From a structural standpoint, states which are classical on A can be expressed as

$$\rho_{AB} = \sum_i p_i \, |a_i\rangle\langle a_i|^A \otimes \rho_i^B \tag{1.14}$$

for some orthonormal basis $\{|a_i\rangle\}$ (analogous definition holds for states which are classical on B). States that are not classical on A are said to have non-zero *quantum discord* [41, 43, 44]. All entangled states necessarily possess discord, but also separable states can.

Discord plays a basic role in quantum information processing, being linked to the impossibility of local broadcasting of correlations and information [45], to quantum data hiding [46], to quantum data locking [47], to entanglement distribution [48, 49], to quantum metrology [50], to quantum cryptography [51]. In this thesis we are interested in the role role of discord in the channel discrimination and channel tomography, as discussed in Chapter 4 and 5, respectively.

1.3 Quantum channels

Quantum channels are the mathematical objects describing the possible transformations of the states of a quantum system. They are given by linear maps acting between spaces of linear operators. In particular we denote by $\mathcal{C}(\mathcal{H}, \mathcal{H}')$ the set of quantum channels from $\mathcal{L}(\mathcal{H})$ to $\mathcal{L}(\mathcal{H}')$. We are interested in one-step transformation, i.e. functions requiring an input state and returning an output state, regardless of any specific dynamical process. Such transformations are formally described by maps which must be linear and must preserve the characterizing properties of a quantum state, i.e. hermiticity, positivity and unit-trace condition. The linearity constraint makes the action of the map consistent when applied to convex mixtures. The remaining requirements preserve the probabilistic interpretation of the statistical operator. In addition, since we cannot neglect possible interactions between our system and another one, e.g. an external environment, we require our map to be positivity preserving[3] even when acting on a larger Hilbert space, namely when our system is only part of a larger one. In particular we require $\Lambda \otimes \mathrm{id}_d$ to be positive for any $d \in \mathbb{N}$, where d is the dimension of the supplementary Hilbert space. We have the following definition.

Definition 1.1. A linear map

$$\begin{aligned} \Lambda : \mathcal{L}(\mathcal{H}) &\to \mathcal{L}(\mathcal{H}) \\ \tau &\mapsto \Lambda[\tau], \end{aligned} \tag{1.15}$$

is called *completely positive* if and only if the map $\Lambda \otimes \mathrm{id}_d$ defined as

$$\begin{aligned} \Lambda \otimes \mathrm{id}_d : \mathcal{L}(\mathcal{H} \otimes \mathbb{C}^d) &\to \mathcal{L}(\mathcal{H} \otimes \mathbb{C}^d) \\ \tau \otimes \omega &\mapsto \Lambda[\tau] \otimes \omega, \end{aligned} \tag{1.16}$$

is positive for any $d \in \mathbb{N}$, where id_d is the identity operator on $\mathbb{C}^d$ and $\omega \in \mathcal{L}(\mathbb{C}^d)$.[4]

However, as proven by Choi in [52], completely positivity can be inferred just by looking at an ancillary space of the same dimension as the original domain of the map:

Theorem 1.1. *For $\mathcal{H} \equiv \mathbb{C}^d$, and with the notation of Definition Eq.* (1.1)*, the positivity of $\Lambda \otimes \mathrm{id}_d$ is enough to guarantee the complete positivity of Λ.*

In Sec. 1.3.1-1.3.3, different ways of representing completely positive (CP), trace preserving (TP) maps will be discussed.

[3]As customary, we usually refer to positivity preserving maps just as positive maps. The same convention will be employed for hermiticity preserving maps.

[4]Notice that τ and ω are generic operators of the respective linear spaces, not necessarily qauntum states. Moreover, since there always exists a tensor product basis for $\mathcal{L}(\mathcal{H} \otimes \mathbb{C}^d)$, then by linearity the action of Λ – as fixed by Eq. (1.16) – is defined for all quantum states in $\mathcal{L}(\mathcal{H} \otimes \mathbb{C}^d)$.

1.3.1 Linear maps as matrices

The space of linear operators $\mathcal{L}(\mathcal{H})$ is an Hilbert space when equipped with the Hilbert-Schmidt (HS) inner product. The latter is defined as

$$\langle A, B\rangle := \mathrm{Tr}(A^\dagger B), \quad A, B \in \mathcal{L}(\mathcal{H}). \tag{1.17}$$

Then, any set of linear operators $\{A_i\}_{i=1}^{d^2}$ orthonormal with respect to the product above forms a basis of $\mathcal{L}(\mathcal{H})$. This means that any $\sigma \in \mathcal{L}(\mathcal{H})$ can be expressed as

$$\sigma = \sum_i \mathrm{Tr}(A_i^\dagger \sigma) A_i. \tag{1.18}$$

Analogously, any linear map $\Lambda : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ can be written as

$$\Lambda[\sigma] = \sum_{ij} \mathscr{L}_{ij} \,\mathrm{Tr}(A_j^\dagger \sigma) A_i, \tag{1.19}$$

where

$$\mathscr{L}_{ij} = \mathrm{Tr}(A_i \Lambda[A_j]). \tag{1.20}$$

In conclusion, Eq. (1.18) enables the identification of $\mathcal{L}(\mathbb{C}^d)$ with the vector space $\mathcal{M}_{d\times d}(\mathbb{C})$ of complex matrices (which is itself a Hilbert space when equipped with the Hilbert-Schmidt inner product). In the same way Eq. (1.20) gives the isomorphism between the space of linear maps on $\mathcal{L}(\mathbb{C}^d)$ and $\mathcal{M}_{d^2\times d^2}(\mathbb{C})$.

The construction above can be easily generalized to channels with different input and output spaces, i.e. $\Lambda : \mathcal{L}(\mathbb{C}^d) \to \mathcal{L}(\mathbb{C}^{d'})$. In such cases the associated matrix algebra is given by $\mathcal{M}_{d'^2\times d^2}(\mathbb{C})$ and the transfer matrix of Λ is defined by the entries

$$\mathscr{L}_{ij} = \mathrm{Tr}(A'_i \Lambda[A_j]), \tag{1.21}$$

where $\{A_i\}$ and $\{A'_i\}$ are HS orthonormal basis of $\mathcal{L}(\mathbb{C}^d)$ and $\mathcal{L}(\mathbb{C}^{d'})$, respectively.

Definition 1.2. For any $\Lambda : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$, we refer to its matrix representation $\mathscr{L}$ with entries given by Eq. (1.21) as the *transfer matrix* of Λ.

We will repeatedly make use of transformations Eqs. (1.18) and (1.21). The latter in particular satisfies several properties which we mention here for future convenience. One first observation is that composition of maps $\Lambda_2 \circ \Lambda_1$ translates into the multiplication of the corresponding matrices $\mathscr{L}_1 \cdot \mathscr{L}_2$. Moreover, Eq. (1.21) implies that the dual $\Lambda^\dagger$ corresponds to $\mathscr{L}^\dagger$ if the map is Hermitian, and to $\mathscr{L}^T$ if the bases are Hermitian.[5] Also, if Λ is a quantum channel, then $\Lambda = \Lambda^\dagger$ if and only if $\mathscr{L} = \mathscr{L}^\dagger$.

1.3.2 Choi-Jamiołkowski isomorphism

The Choi-Jamiołkowski isomorphism [52, 53] is one of the central results of quantum information, establishing a correspondence between bipartite operators and linear maps acting on single–system operators. It allows to effectively encode or parametrize quantum transformations, with applications that go from the optimization of protocols in quantum information, to the analysis of rates in quantum communication, all the way to the consideration of the issue of causal order in physics (see, e.g., [54–56]). In particular, the isomorphism implies that for any channel $\Lambda \in \mathcal{C}(\mathcal{H}, \mathcal{H}')$, the state obtained by letting Λ acting on one part of the maximally entangles state encodes all the relevant properties of the channel itself. To be more precise, let us consider the maximally entangles state in $\mathcal{L}(\mathcal{H} \otimes \mathcal{H})$ given by

$$\left|\psi^+\right\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |ii\rangle \,, \tag{1.22}$$

where d is the dimension of the underlying Hilbert space. Then, we have the following:

Theorem 1.2 (Choi-Jamiołkowski isomorphism [52,53,57])**.** *The relations below establish a one-to-one correspondence between bipartite linear operators* $\chi := \chi \in \mathcal{L}(\mathcal{H}_A \otimes$

[5] We recall that in finite–dimensional linear spaces the definition of dual map coincides with that of adjoint operator. That is, for any $\Lambda \in \mathcal{L}(\mathcal{H})$, its adjoint $\Lambda^\dagger$ is defined via

$$\langle \Lambda^\dagger[\tau], \omega\rangle = \langle \tau, \Lambda[\omega]\rangle, \quad \forall\ \tau, \omega \in \mathcal{L}(\mathcal{H}).$$

$\mathcal{H}_B$) *and linear maps* $\Lambda \in \mathcal{C}(\mathcal{H}_B, \mathcal{H}_A)$*:*

$$\chi = (\Lambda \otimes \mathrm{id}) \left|\psi^+\right\rangle\!\left\langle\psi^+\right| \tag{1.23}$$

$$\mathrm{Tr}(A\Lambda[B]) = d\,\mathrm{Tr}(\chi A \otimes B^T), \tag{1.24}$$

for all $A \in \mathcal{M}_{d'}$ *and* $B \in \mathcal{M}_d$ *(where* $d = \dim \mathcal{H}$ *and* $d' = \dim \mathcal{H}'$*) and* $|\psi^+\rangle$ *as in* Eq. (1.22)*. Moreover,* Eq. (1.23) *and* Eq. (1.24) *are mutual inverse and satisfy what follows:*

i) $\chi = \chi^\dagger$ *if and only if* $\Lambda[B^\dagger] = \Lambda[B]$ *for all* $B \in \mathcal{M}_d$*, i.e.* χ *is Hermitian if and only if* Λ *is hermiticity preserving,*

ii) Λ *is completely positive if and only if* $\chi \geq 0$*,*

iii) $\Lambda[\mathbb{1}] = \mathbb{1}$ *if and only if* $\mathrm{Tr}_B(\chi) = \mathbb{1}_{d'}/d$*,*

iv) $\Lambda^\dagger[\mathbb{1}] = \mathbb{1}$ *if and only if* $\mathrm{Tr}_A(\chi) = \mathbb{1}_d/d$*,*

v) $\mathrm{Tr}(\chi) = \mathrm{Tr}(\Lambda^\dagger[\mathbb{1}])/d$*.*

The matrix χ is called the *Choi matrix* of Λ. It is interesting to notice that, differently from the transfer matrix $\mathscr{L}$, the Choi matrix gives a direct way for checking the complete positivity of the associated channel. Lastly, let us notice that the Choi matrix associated to Λ is the transfer matrix $\mathscr{L}$ represented in the basis made of matrix units $|i\rangle\langle j|$, where $|i\rangle$ and $|j\rangle$ are elements of the computational basis. In particular, for any Λ the correspondence between χ and $\mathscr{L}$ is given by

$$\mathscr{L} = d\chi^R, \tag{1.25}$$

where the map $\chi \mapsto \chi^R$ is defined as[6]

$$\langle m, n|\chi^R|k, l\rangle = \langle m, k|\chi|n, l\rangle. \tag{1.26}$$

[6]The involution R of Eq. (1.26) is sometime called *reshuffling operation* (cf. Sec 2.2).

1.3.3 Operator-sum representation

We have seen that the duality between quantum states and channels provides a direct way to infer the properties of a linear map by looking at the associated bipartite matrix, and vice versa. Beyond that, the relevance of the Choi-Jamiołkowski isomorphismy relies also upon a very useful representation known as *Kraus decomposition* (or operator sum representation) [52, 58] of completely positive maps.

Theorem 1.3 (Kraus decomposition [52, 57, 58])**.** *A linear map $\Lambda \in \mathcal{C}(\mathcal{H}, \mathcal{H}')$ is completely positive if and only if it admits a decomposition of the form*

$$\Lambda[A] = \sum_{i=1}^{n} K_i A K_i^\dagger. \tag{1.27}$$

Given a completely positive linear map Λ, the associated set of Kraus operators satisfies the following properties:

i) Λ is trace preserving if and only if $\sum_i K_i^\dagger K_i = \mathbb{1}$, and it is unital if and only if $\sum_i K_i K_i^\dagger = \mathbb{1}$,

ii) the minimal number of operators in the decomposition, called the Kraus rank r, satisfies $r \leq dd'$,

iii) there always exists a representation with r HS orthogonal Kraus operators, i.e. such that $\mathrm{Tr}(K_i^\dagger K_j) = \delta_{ij}$,

iv) given two sets of Kraus operators $\{K_i\}$ and $\{F_j\}$, they represent the same map Λ if and only if there exists a unitary U such that $K_i = \sum_j U_{ij} F_j$ (the smaller set is padded with zeroes).

The Kraus representation can be derived directly from the famous *Stinespring's dilation theorem* [59]. It asserts that the action of a quantum channel Λ on a state ρ can be simulated by introducing an auxiliary system E (sometimes referred to as the *environment*), applying a unitary interaction between the system and E, and then discarding the ancillary system. To be precise, we have the following:

Theorem 1.4 (Stinespring's dilation [59]). *Let $\Lambda \in \mathcal{C}(\mathcal{H}, \mathcal{H}')$ be a quantum channel. Then there exist a Hilbert space $\mathcal{H}_E$ and a unitary operation U on $\mathcal{L}(\mathcal{H} \otimes \mathcal{H}_E)$ such that*

$$\Lambda[\rho] = \mathrm{Tr}_E[U(\rho \otimes |0\rangle\langle 0|_E)U^\dagger], \tag{1.28}$$

for all $\rho \in \mathcal{S}(\mathcal{H})$, where $\mathrm{Tr}_E[\,\cdot\,]$ denotes the partial trace over $\mathcal{H}_E$.

1.4 Norms

In the remainder of the thesis we shall make an extensive use of norms. In particular, we will consider the family of Schatten $p-$norms, defined as

$$\|L\|_p := \left[\mathrm{Tr}\left((L^\dagger L)^{\frac{p}{2}}\right)\right]^{\frac{1}{p}}, \tag{1.29}$$

with $1 \leq p \leq \infty$ and with L a generic linear operator [60].[7] In particular, we are interested in the special cases $p = 1, 2, \infty$.

The Schatten $p-$norms satisfies a series of interesting properties, which we list here for future reference. First of all, they are monotonic in p, meaning that for $1 \leq p \leq p' \leq \infty$ one has

$$\|L\|_1 \geq \|L\|_p \geq \|L\|_{p'} \geq \|L\|_\infty. \tag{1.30}$$

Moreover, they are isometrically invariant: for any two isometries U and V, it holds that $\|ULV\|_p = \|L\|_p$ for any $1 \leq p \leq \infty$. Finally, it is convenient to recall that the $p-$norms are submultiplicative, i.e.

$$\|LR\|_p \leq \|L\|_p \|R\|_p \tag{1.31}$$

[7]Notice that the norms in Eq. (1.29) can be equivalently expressed in terms of the singular values of L.

and satisfy the Hölder inequality:

$$\|LR\|_1 \leq \|L\|_p \|R\|_q \tag{1.32}$$

where $1/p + 1/q = 1$ and $L, T \in \mathcal{C}_(\mathcal{H}, \mathcal{H}')$.

The norms above appear in different context in quantum information, and some of them possess a particular physical meaning. For example, let us consider the trace distance between two density matrices ρ and σ , which is defined in terms of the $1-$norm as $D(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_1$ [13]. Its operational meaning is that of bias in the optimal discrimination of the two states: the probability of correctly identifying the state of a system that is a priori in the state ρ or σ each with 50% chance, in the single-shot scenario when one is given one copy of the state to measure, is $(1 + D(\rho, \sigma))/2$. The trace distance varies between 0 (for identical states) to 1 (for perfectly distinguishable states, which are mathematically orthogonal, $\mathrm{Tr}(\rho\sigma) = 0$).

A similar notion of distance can be attributed to pairs of channels too [61]. For this, one needs the notion of superoperator norms, defined for any $\Lambda \in \mathcal{C}(\mathcal{H}, \mathcal{H}')$ and any $1 \leq p, q \leq \infty$ as

$$\|\Lambda\|_{q\to p} = \sup_{X\neq 0} \frac{\|\Lambda[X]\|_p}{\|X\|_q}. \tag{1.33}$$

In the particular case when the supremum is taken over Hermitian operators, the norm above will be denoted with $\|\cdot\|^H_{q\to p}$. Given two channels $\Lambda, \Gamma \in \mathcal{C}(\mathcal{H}, \mathcal{H}')$, the quantity $\|\Lambda - \Gamma\|^H_{1\to 1}$ represents a way to measure the distance between the channels. In particular, since

$$\|\Lambda - \Gamma\|^H_{1\to 1} = \max_{\||\psi\rangle\|_2 = 1} \left\{ \|\Lambda[|\psi\rangle\langle\psi|] - \Gamma[|\psi\rangle\langle\psi|]\|_1 \right\}, \tag{1.34}$$

it characterizes the maximum probability of distinguishing the two channels over all pure states. However, Kitaev showed in [62] that this norm is not stable with respect to the tensorization with the identity superoperator. In other words, there exist quantum channels Λ', Γ' such that $\|\Lambda' - \Gamma'\|^H_{1\to 1} < \|\Lambda' \otimes \mathrm{id} - \Gamma' \otimes \mathrm{id}\|^H_{1\to 1}$. To resolve this problem, he introduced the so-called *diamond norm*, defined as

$$\|\Lambda\|_{\diamond} = \|\Lambda \otimes \mathrm{id}\|_{1\to 1}, \tag{1.35}$$

where the identity superoperator is defined on a space of the same dimension as the domain of Λ. We will make use of the diamond norm in Sec. 4.1.

Chapter 2

Operator Schmidt decomposition

In this chapter we introduce the central object of our investigation, i.e. the operator Schmidt decomposition (OSD). This is nothing but a generalization of the ordinary Schmidt decomposition for state vectors, sketched in Sec. 1.2, to the case of density matrices. As already said, the Schmidt decomposition plays a central role in the characterization of pure state entanglement, in that it explicitly discriminates between separable and entangled state. However, in the general case of density operators the situation is more involved, since entanglement is not the only correlation which the state can exhibit, as discussed in Sec. 1.2. We argue that the OSD accounts for this variety of correlations and it offers means by which meaningful measures of total correlations can be devised.

We begin with the derivation of the OSD itself, which takes advantage of the vector space structure of $\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then, the principal results concerning the exploitation of the OSD in the detection and quantification of entanglement will be discussed in detail. In this chapter we review the main results contained in [63–69].

2.1 Operator Schmidt decomposition

The ordinary Schmidt decomposition is a consequence of a critical result from linear algebra known as the singular value decomposition. This is nothing but a particular factorization of complex matrices and, as such, it does apply to the elements of any vector space. This also means that density matrices – vectors of $\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ – can be decomposed using the singular value decomposition (see [65,67] and references therein).

Here we provide an explicit derivation of the OSD. Before that however, an important definition that will be used repeatedly throughout the thesis is needed.

Definition 2.1 (Correlation matrix)**.** Let $\rho_{AB} \in \mathcal{S}_{AB}$ be a bipartite state, and let $\{C_i\}$ and $\{D_i\}$ be HS orthonormal bases of $\mathcal{L}(H_A)$ and $\mathcal{L}(H_B)$, respectively. We define the *correlation matrix* $\mathcal{C}(\rho_{AB})$ of ρ_{AB} as the matrix with entries

$$\mathcal{C}(\rho_{AB})_{kl} = \mathrm{Tr}(C_k^\dagger \otimes D_l^\dagger \rho_{AB}), \tag{2.1}$$

for $k = 1, \ldots, d_A^2$ and $l = 1, \ldots, d_B^2$.

It is worth noticing that a change in the choice of the orthonormal bases $\{C_i\}$ and $\{D_i\}$ induces only a unitary change in the correlation matrix. This implies that the value of unitarily invariant functionals of the correlation matrix will not be affected by the choice of the local bases.

Theorem 2.1 (Operator Schmidt decomposition)**.** *Any bipartite state $\rho_{AB} \in \mathcal{S}_{AB}$ admits a decomposition of the form*

$$\rho_{AB} = \sum_{i=1}^{m} r_i A_i \otimes B_i, \tag{2.2}$$

where $\{A_i\}_{i=1}^{d_A^2}$ and $\{B_i\}_{i=1}^{d_B^2}$ form (ρ_{AB}-dependent) bases of $\mathcal{L}(\mathcal{H}_A)$ and $\mathcal{L}(\mathcal{H}_B)$, respectively, orthonormal with respect to the HS inner product. Moreover, the r_i are nonnegative real numbers and $m \leq \min\{d_A^2, d_B^2\}$.

Proof. Fix HS orthonormal bases $\{C_i\}$ and $\{D_i\}$ on $\mathcal{L}(H_A)$ and $\mathcal{L}(H_B)$, respectively. In view of Definition 2.1, ρ_{AB} can be expressed as

$$\rho_{AB} = \sum_{kl} \mathcal{C}(\rho_{AB})_{kl}\, C_k \otimes D_l. \tag{2.3}$$

Consider the singular value decomposition of the correlation matrix $\mathcal{C}(\rho_{AB}) = U\Sigma V$, where U, V are unitary and $\Sigma_{ii} = r_i$. One has

$$\begin{aligned}\rho_{AB} &= \sum_{kl} \sum_{i} U_{ki} r_i V_{il} C_k \otimes D_l \\ &= \sum_i r_i \left(\sum_k U_{ki} C_k \right) \otimes \left(\sum_l V_{il} D_l \right). \end{aligned} \tag{2.4}$$

Thus, upon defining $A_i := \sum_k U_{ki} C_k$ and $B_i := \sum_l V_{il} D_l$ one gets the wanted decomposition (2.2). Indeed, the new bases A_i and B_i are orthonormal since the original bases were orthonormal and the coefficients of the linear combination form unitary matrices. Finally, the orthonormality requirement on the operators A_i and B_i just defined implies that the number of terms m in Eq. (2.4) is bounded by $\min\{d_A^2, d_B^2\}$. □

Definition 2.2. From now on, we refer to the r_i of Eq. (2.2) as the *operator Schmidt coefficients* (OSC) of ρ_{AB}, and we assume that they are ordered: $r_1 \geq r_2 \geq \ldots \geq r_{\mathrm{OSR}(\rho_{AB})}$. Moreover, we indicate by $\mathrm{OSR}(\rho_{AB})$ the *operator Schmidt rank* (OSR) of ρ_{AB}. Notice that the OSR is the minimum number of terms necessary to decompose ρ_{AB} in the form $\sum_i C_i \otimes D_i$ (here the local operators need not be necessarily orthonormal).

Finally, let us report an immediate but important fact about the OSC, here stated in the form of a proposition for future reference.

Proposition 2.1. *The sum of the squares of the OSC of any $\rho_{AB} \in \mathcal{S}_{AB}$ equals its purity:*

$$\sum_i r_i^2 = \|\rho_{AB}\|_2^2. \tag{2.5}$$

Remark. *Since ρ_{AB} is Hermitian, one can argue that the two orthonormal operator bases in* (2.2) *can be (but need not be) chosen to be made of Hermitian operators. Also, notice that the SD of a pure state $|\psi\rangle_{AB}$ given in Eq.* (1.10) *and the OSD of the corresponding density matrix $|\psi\rangle\langle\psi|_{AB}$ are related: $r_i = \sqrt{p_k}\sqrt{p_l}$, $A_i = |a_k\rangle\langle a_l|$, and $B_i = |b_k\rangle\langle b_l|$, for $i = (k, l)$ a multi-index. It follows in particular that* $\mathrm{OSR}(|\psi\rangle\langle\psi|_{AB}) = \mathrm{SR}(\psi_{AB})^2$.

2.2 The realignment criterion for separability

A well-known application of the OSD is the so-called *computable cross norm* (CCN) criterion for separability, also known as the *realignment criterion* [63, 64], which establishes a necessary condition satisfied by separable states. This feature is achieved by the identification of an upper bound to a particular functional, when the latter is evaluated on separable states. As the names suggest, this result was originally derived in two different settings: as a bound on a particular norm, the CCN [64, 70] indeed, and in the context of matrix reordering [63]. For completeness, here we state and prove this result in the two original settings. After that, we argue how the two forms of the criteria – henceforth simply called the realignment criterion (RC) – can be restated in terms of the OSC only.

We begin by defining the CCN of a linear operator.

Definition 2.3 (Computable cross norm [64])**.** The CCN of an operator[1] $\tau \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is given by

$$\|\tau\|_{\mathrm{CC}} := \inf \left\{ \sum_{i=1}^{k} \|A_i\|_2 \|B_i\|_2 \;\middle|\; \tau = \sum_{i=1}^{k} A_i \otimes B_i \right\}, \tag{2.6}$$

where $A_i \in \mathcal{L}(H_A)$, $B_i \in \mathcal{L}(H_B)$. The infimum is taken over all the decompositions of τ into a finite sum of tensor products of local operators. We refer to the single elements $A_i \otimes B_i$ of the decomposition, for for any i, as *simple tensors.*

In view of the definition above we can prove what follows.

Theorem 2.2 (CCN criterion for separability [64])**.** *Let $\rho_{AB} \in \mathcal{S}_{AB}$ be separable. Then, $\|\rho_{AB}\|_{CC} \leq 1$.*

Proof. The subadditvity of the Hilbert-Schmidt norm together with Definition 2.3 yield $\|\rho_{AB}\|_{\mathrm{CC}} \geq \|\rho_{AB}\|_2$, and the inequality is saturated for simple tensors, i.e. $\|A_i \otimes B_i\|_{\mathrm{CC}} = \|A_i\|_2 \|B_i\|_2$. On the other hand, for any simple tensor $A_i \otimes B_i$ the monotonicity

[1] Notice that it is not required for the operator to be positive semidefinite.

of the Schatten p-norms implies

$$\begin{aligned}\|A_i \otimes B_i\|_{\mathrm{CC}} &= \|A_i\|_2 \|B_i\|_2 \\ &\leq \|A_i\|_1 \|B_i\|_1.\end{aligned} \tag{2.7}$$

The property above is known as *subcross property.*[2] In conclusion, for a separable state $\sigma_{AB} = \sum_i^r \lambda_i \sigma_i^A \otimes \sigma_i^B$, subadditivity and subcross property imply

$$\|\sigma_{AB}\|_{\mathrm{CC}} \leq \sum_i^r \lambda_i \|\sigma_i^A \otimes \sigma_i^B\|_{\mathrm{CC}} \leq \sum_i^r \lambda_i \|\sigma_i^A\|_1 \|\sigma_i^B\|_1 = 1. \tag{2.8}$$

□

We now prove the criterion above from another perspective, i.e. in terms of a simple inspection of the density matrix entries. To do so, the author of [63] made use of an operation known as *vectorization* of a matrix [71].

Definition 2.4. For any $m \times n$ matrix $A = [a_{ij}]$, where a_{ij} is the entry corresponding to the $i-$th row and $j-$th column, we define the vectorization $|A\rangle\rangle$ of A as the following vector:

$$|A\rangle\rangle = [a_{11}, \ldots, a_{m1}, a_{12}, \ldots, a_{m2}, \ldots, a_{1n}, \ldots, a_{mn}]^T, \tag{2.9}$$

were T denotes transposition. Moreover, if G is a $m \times m$ block matrix with blocks $G_{i,j}$, of size $n \times n$, we define the realigned matrix G^R as the following $m^2 \times n^2$ matrix:

$$G^R = \begin{pmatrix} |G_{1,1}\rangle\rangle^T \\ \vdots \\ |G_{m,1}\rangle\rangle^T \\ \vdots \\ |G_{1,m}\rangle\rangle^T \\ \vdots \\ |G_{m,m}\rangle\rangle^T \end{pmatrix}. \tag{2.10}$$

[2] To be more specific, we say that the CCN satisfies the subcross property with respect to the trace norm.

In order to give a concrete example of this representation, consider a 4×4 bipartite density matrix ρ_{AB}. Then, Eqs. (2.9) and (2.10) yield the transformation below:

$$\rho_{AB} = \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & \rho_{22} & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & \rho_{33} & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & \rho_{44} \end{pmatrix} \rightarrow (\rho_{AB})^R = \begin{pmatrix} \rho_{11} & \rho_{21} & \rho_{12} & \rho_{22} \\ \rho_{31} & \rho_{41} & \rho_{32} & \rho_{42} \\ \rho_{13} & \rho_{23} & \rho_{14} & \rho_{24} \\ \rho_{33} & \rho_{43} & \rho_{34} & \rho_{44} \end{pmatrix}. \tag{2.11}$$

We have the following.

Theorem 2.3 (Realignment criterion for separability [63])**.** *Let $\rho_{AB} \in \mathcal{S}_{AB}$ be separable. Then, $\|(\rho_{AB})^R\|_1 \leq 1$.*

Proof. Let $\rho_{AB} \in \mathcal{S}_{AB}$, with $d_A = m, d_B = n$, and assume that it admits the decomposition $\rho_{AB} = \sum_i \lambda_i \rho_i^A \otimes \rho_i^B$. In addition, let U_i and V_i be the unitary matrices that diagonalise ρ_i^A and ρ_i^B , respectively. Then, by denoting with $E_{11}^{(k,l)}$ the $k \times l$ matrix with a one in the $(1,1)$ position and zeros elsewhere, ρ_{AB} can be written as $\rho_{AB} = \sum_i \lambda_i (U_i E_{11}^{(m,m)} U_i^\dagger) \otimes (V_i E_{11}^{(n,n)} V_i^\dagger)$. By using the following properties of the vectorization operation: (see [71–73])

$$|ABC\rangle\rangle = (C^T \otimes A)\, |B\rangle\rangle \tag{2.12}$$

$$(A \otimes B)^R = |A\rangle\rangle\, |B\rangle\rangle^T, \tag{2.13}$$

one obtains

$$\begin{aligned} \|(\rho_i^A \otimes \rho_i^B)^R\|_1 &= \left\| \left(U_i^* \otimes U_i\right) \left|E_{11}^{(m,m)}\right\rangle\right\rangle \left|E_{11}^{(n,n)}\right\rangle\right\rangle^T \left(V_i^* \otimes V_i\right)^\dagger \right\|_1 \\ &= \left\| (U_i^* \otimes U_i) E_{11}^{(m^2,n^2)} (V_i^* \otimes V_i)^\dagger \right\|_1 \\ &= \left\| E_{11}^{(m^2,n^2)} \right\|_1 \\ &= 1. \end{aligned} \tag{2.14}$$

The last two lines above depend upon the facts that the trace norm is unitarily invariant and that by construction $E_{11}^{(m^2,n^2)}$ possesses a unique singular value equal to one. One concludes using subadditivity:

$$\begin{aligned}\left\|(\rho_{AB})^R\right\|_1 &\leq \sum_i \lambda_i \left\|(\rho_i^A \otimes \rho_i^B)^R\right\|_1 \\ &= \sum_i \lambda_i \\ &= 1. \end{aligned} \tag{2.15}$$

□

As anticipated, Theorems 2.2 and 2.3 can be restated in terms of the OSC. We have what follows.

Theorem 2.4. *Let $\rho_{AB} \in \mathcal{S}_{AB}$ be separable. Then $\|\mathcal{C}(\rho_{AB})\|_1 \leq 1$.*

Proof. We give a sketch of the proof, which is straightforward. Notice that the correlation matrix of a pure product state is just a rank one matrix with unit one–norm. Then, the $1-$norm of the correlation matrix of a pure product state is equal to one. One concludes by using the linearity of the correlation matrix in the state and the convexity of the $1-$norms.

We observe that the result above holds for any Schatten norm: separable states are such that $\|\mathcal{C}(\rho_{AB})\|_p \leq 1$ for any p. However, the monotonicity of the $p-$norms implies that the criterion corresponding to $p = 1$ is the more stringent. □

The connection between this new form of the criterion involving only the OSC and the CCN criterion is somewhat lengthy, and for convenience we refer to [70] for a detailed discussion. Of course, what it is proved there is that the CCN of a quantum state equals the sum of its OSC, ensuring the equivalence between the two approaches.

On the other hand, the relationship between the sum of the OSC and the $1-$norm of the realigned matrix ρ_{AB}^R is more interesting, and it comes from the comparison of Eq. (2.1) with the transformation defined in Eq. (2.10). To be more precise, consider a bipartite state $\rho_{AB} \in \mathcal{S}_{AB}$ and let $\{E_i\}$ and $\{F_j\}$ be the basis of $\mathcal{L}(\mathcal{H}_A)$ and $\mathcal{L}(\mathcal{H}_B)$

given by *matrix units* of the form $E_i = |k\rangle\langle l|$. It holds that [73, 74]

$$[(\rho_{AB})^R]_{ij} = \mathrm{Tr}[\rho_{AB}(E_i \otimes F_j)]. \tag{2.16}$$

In other words, the realigned matrix $(\rho_{AB})^R$ is nothing but the correlation matrix $\mathcal{C}(\rho_{AB})$ expressed with respect to a particular basis. Hence Eq. (2.16) implies that the singular values of $(\rho_{AB})^R$ equal the OSC of ρ_{AB}, so that $\|(\rho_{AB})^R\|_1 = \sum_i r_i$ and the equivalence of Theorem 2.4 and 2.3 follows.

2.3 Realignment method for steering detection

Just as for entanglement, it is possible to devise a realignment method to detect steering. This result was proved in [75], as a consequence of an operational form of steering certification. To see how such certification takes place, let us define the uncertainty of an observable O_X in terms of its variance $\delta^2(O_X) = \langle O_X^2\rangle - \langle O_X\rangle^2$, with $\langle O_X\rangle = \mathrm{Tr}(X\rho)$ is the expectation of O_X. Now assume that Alice and Bob share a bipartite state ρ_{AB} and perform measurements $\{O_{A,i}\}$ and $\{O_{B,i}\}$ on the respective subsystems. Then we have the following.

Proposition 2.2 (Local uncertainty relation for steering [75])**.** *A bipartite state ρ_{AB} is steerable (by A) if the inequality*

$$\sum_i \delta^2[\alpha_i(O_{A,i} \otimes \mathbb{1}) + (\mathbb{1} \otimes O_{B,i})] \geq \min_{\rho_B} \sum_i \delta^2(O_{B,i}) \tag{2.17}$$

is violated, with $\{\alpha_i\} \in \mathbb{R}$.

Proof. It is a general fact from probability theory that for a random variable $x \in X$ with probability distribution $p(X) = \sum_i p_i\, p(x|i)$, its variance satisfies

$$\delta^2(X) \geq p_i \sum_i \delta^2(X)_i, \tag{2.18}$$

where $\delta^2(X)_i$ is the variance of X with probability distribution $p(x|i)$. On the other hand, the defining condition for steering Eq. (1.13) implies that assuming a local hidden state for Bob means

$$\begin{aligned}\delta^2[\alpha_i(O_{A,i}\otimes\mathbb{1})+(\mathbb{1}\otimes O_{B,i})] &\geq \sum_\lambda p_\lambda\,\delta^2[\alpha_i(O_{A,i}\otimes\mathbb{1})+(\mathbb{1}\otimes O_{B,i})]_\lambda \\ &= \sum_\lambda p_\lambda\,[\alpha_i^2\delta^2(O_{A,i})_\lambda+\delta^2(O_{B,i})_{\rho_\lambda^B}],\end{aligned} \tag{2.19}$$

where $\delta^2(O_{A,i})_k$ is the variance of $O_{A,i}$ with probability distribution $p(o_{A,i}|O_{A,i};\lambda)$, while $\delta^2(O_{B,i})_{\rho_B}$ is the variance of $O_{B,i}$ with probability distribution $p(o_{B,i}|O_{B,i};\rho_\lambda^B)$, cf. Eq. (1.13). By applying Eq. (2.18) to the two terms in the last line of Eq. (2.19) we obtain, on the one hand, $\sum_i\alpha_i^2\delta^2(O_{A,i})\geq 0$, since Alice can prepare any probability distribution. On the other hand, Bob's local states satisfy $\sum_i\delta^2(O_{B,i})\geq\min_{\rho_B}\sum_i\delta^2(O_{B,i})$. Hence, Eq. (2.17) is satisfied for any unsteerable state and the claim follows. □

Before proving the realignment criterion for steering, we need one last technical proposition which sets a lower bound to the sum of the variances of a basis of observables.

Proposition 2.3. *Let $\{H_i\}$ be an HS orthonormal basis for the space $\mathcal{L}(\mathcal{H})$ of linear operator on $\mathcal{H}=\mathbb{C}^d$ and let $H_i=H_i^\dagger$ for any i. Then,*

$$\sum_i\delta^2(H_i)\geq d-1 \tag{2.20}$$

for any $\rho\in\mathcal{S}(\mathcal{H})$.

Proof. From the definition of variance, the claim above can be rewritten as $\sum_i\langle H_i^2\rangle-\langle H_i\rangle^2\geq d-1$. Observe that, since any $\rho\in\mathcal{S}(\mathcal{H})$ can be expressed as $\rho=\sum_i\langle H_i\rangle H_i$, one has $\sum_i\langle H_i\rangle^2=\mathrm{Tr}(\rho^2)\leq 1$. Then, it remains to prove $\sum_i\langle H_i^2\rangle=d$. To do so, let us identify $\{H_i\}$ with the Hermitian basis of $\mathcal{L}(\mathcal{H})$ given by the generalization of the Pauli matrices, i.e. the basis given by the $d(d-1)$ matrices $(|k\rangle\langle l|+|l\rangle\langle k|)/\sqrt{2}$ and $(|k\rangle\langle l|-|l\rangle\langle k|)/i\sqrt{2}$ with $1\leq k<l\leq d$, plus the d matrices $|k\rangle\langle k|$ with $1\leq k\leq d$. With this choice one directly verifies that $\sum_i H_i^2=d\,\mathbb{1}$, hence $\sum_i\langle H_i^2\rangle=d$. This result is basis-independent because any other Hermitian basis $\{F_i\}$ will be related to $\{H_i\}$ by $F_i=\sum_k O_{ik}H_k$ for an orthogonal O, thus $\sum_i F_i^2=\sum_{ikl}O^T_{ki}O_{il}H_kH_l=\sum_k H_k^2$. □

We can now prove the sought criterion.

Theorem 2.5 (Realignment criterion for steerability [75])**.** *Let $\rho_{AB} \in \mathcal{S}_{AB}$ be unsteerable and $\mathcal{C}(\rho_{AB})$ its correlation matrix. Then, $\|\mathcal{C}(\rho_{AB})\|_1 \leq \sqrt{d_B}$.*

Proof. Let $\rho_{AB} \in \mathcal{S}_{AB}$ and consider its OSD with respect to local orthonormal Hermitian bases $\{H_{A,i}\}$ and $\{H_{B,i}\}$, i.e. $\rho_{AB} = \sum_i r_i H_{A,i} \otimes H_{B,i}$. Since $H_{A,i}$ and $H_{B,i}$ can be thought to be Alice and Bob's local observables, Propositions 2.2 and 2.3 imply that the violation of the following inequality

$$\sum_i \delta^2[\alpha_i(H_{A,i} \otimes \mathbb{1}) + (\mathbb{1} \otimes H_{B,i})] \geq d_B - 1, \tag{2.21}$$

with $\{\alpha_i\} \in \mathbb{R}$, is a sufficient condition for ρ_{AB} to be steerable by Alice. By choosing $\alpha := \alpha_i = -\sum_i r_i/d_B$ one sees that Eq. (2.21) will be violated anytime that

$$d_B\alpha^2 + 2\alpha \sum_i r_i - \sum_i (\alpha \langle H_{A,i}\rangle + \langle H_{B,i}\rangle)^2 < -1, \tag{2.22}$$

where we have used $\langle H_{A,i} \otimes H_{B,i}\rangle = r_i$. Then, by omitting the (positive) quadratic term in Eq. (2.22), the violation witnessing the presence of steering takes place when $\|\mathcal{C}(\rho_{AB})\| = \sum_i r_i > \sqrt{d_B}$. □

2.4 Symmetric polynomials in the Schmidt coefficients

As seen in the Sec. 2.2, the OSC can be successfully employed for entanglement detection. This feature is achieved, e.g. through the evaluation of the 1−norm of the correlation matrix, which reveals the entanglement of a subset of all entangled states. One can easily check indeed that this criterion will leave certain entangled states undetected [64], and the natural question addressed in [65] is whether it is possible to formulate separability criteria which are stronger than the RC while still relying, as the latter, only on the OSC. An attempt in this direction was made with the introduction of the *elementary symmetric polynomials*[3] (ESP) in the OSC. In the following we review

[3] In commutative algebra, a *symmetric polynomial* is a polynomial $p(x_1, x_2, \dots, x_n)$ which is invariant under the permutation of its n variables, i.e. $p(x_1, x_2, \dots, x_n) = p(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)})$, where $\sigma(k)$

the main results obtained in [65–67].

Definition 2.5. Let $\rho_{AB} \in \mathcal{S}_{AB}$ be a bipartite state and $\mathcal{C}(\rho_{AB})$ its correlation matrix. The $l^{th}-$order elementary symmetric polynomials in the OSC of ρ_{AB} are defined as the following quantities:

$$\begin{aligned} M^{[1]}(\rho_{AB}) &= \sum_{1 \leq k \leq d^2} r_k \\ M^{[2]}(\rho_{AB}) &= \sum_{1 \leq k_1 < k_2 \leq d^2} r_{k_1} r_{k_2} \\ &\vdots \\ M^{[l]}(\rho_{AB}) &= \sum_{1 \leq k_1 < \cdots < k_l \leq d^2} r_{k_1} \cdots r_{k_l} \\ &\vdots \\ M^{[d^2]}(\rho_{AB}) &= \prod_{1 \leq k \leq d^2} r_k. \end{aligned} \tag{2.23}$$

By direct inspection it is clear that the RC can be equivalently reformulated in term of the $1^{st}-$order ESP of the OSC, by saying that $M^{[1]}(\rho_{AB}) \leq 1$ for a separable ρ_{AB}. In general, the RC induces a family of criteria for separability which read, for each $1 \leq l \leq d^2$,

$$\rho_{AB} \text{ separable } \Rightarrow M^{[l]}(\rho_{AB}) \leq \binom{d^2}{l} \left(\frac{1}{d^2}\right)^l. \tag{2.24}$$

This can be further refined by bringing into play the OSR of (ρ_{AB}), which here we denote with K for convenience. We get

$$\rho_{AB} \text{ separable } \Rightarrow M^{[l]}(\rho_{AB}) \leq \binom{K}{l} \left(\frac{1}{K}\right)^l \tag{2.25}$$

if $l \leq K$, and zero otherwise. Moreover, as for the case of Eq. (2.24), the RC is fully recovered when $l = 1$.

denotes permutation [76]. Every symmetric polynomial can be expressed as a polynomial in the n elementary symmetric polynomials $e_k(x_1, x_2, \ldots, x_n)$, defined as

$$e_k(x_1, x_2, \ldots, x_n) = \sum_{1 \leq j_1 < j_2 \cdots < j_k \leq n} x_{j_1} x_{j_2} \cdots x_{j_k}, \quad \text{for } k = 1, 2, \cdots, n.$$

The separability criteria derived above are easily verifiable and, as the reader has certainly noticed, they are weaker than (or equivalent to) the RC, having been derived from it. Nonetheless, the authors of [65] argue that the approach through which these results were obtained (the use of the set of the OSC *as a whole*) could hopefully provide RC-independent separability criteria. The idea (then pursued in [77]) is based on the question whether it is possible to derive an upper bound $\tilde{B}(d, D)$ such that

$$M^{[1]} \leq 1 \ \Rightarrow M^{[l]}(\rho_{AB}) \leq \tilde{B}(d, D) < \binom{K}{l} \left(\frac{1}{K}\right)^{l} \tag{2.26}$$

where, consistently with the definition of d, we have set $D := \max\{d_A, d_B\}$. In other words, Eq. (2.26) establishes a strict upper bound for the symmetric polynomials in the OSC of the states passing the RC. If Eq. (2.26) holds, then a fortiori also the following criteria will be satisfied:

$$\rho_{AB} \text{ separable } \Rightarrow M^{[l]}(\rho_{AB}) \leq B(d, D) < \binom{K}{l} \left(\frac{1}{K}\right)^{l}, \tag{2.27}$$

where $B(d, D) \leq \tilde{B}(d, D)$. It follows that if $\tilde{B}(d, D) > B(d, D)$, then there exists a state ρ_{AB} that satisfies the RC, but such that $M^{[l]}(\rho_{AB}) > B(d, D)$. Thus, there would be at least an entangled state detected as such by $B(d, D)$ but not by the RC. Numerical tests performed in [65] for the cases $d = 2, D = 2$ and $d = 2, D = 3$ suggest that $\tilde{B}(2, 2) = B(2, 2)$ and $\tilde{B}(2, 3) > B(2, 3)$. These estimations were partially validated in [77], where the authors proved that $\tilde{B}(n, n) = B(n, n)$ for any n, meaning that when $d = D$ the bound on the symmetric polynomial cannot be used to derive separability criteria which are independent of the RC. Despite this negative result – which by the way agrees with the numerical test of [65] – the possibility of deriving new RC-independent criteria for other combination of d, D remains open. In particular, the numerical results for $d = 2, D = 3$ still corroborates the plausibility of the approach sketched so far, therefore making room for future studies.

2.5 Beyond the realignment criterion

In this section we review a special result contained in [66], a special instance of which was derived also in [78]. We start by reporting a technical theorem from [66], summarised in the following. This is quite specific, and neglecting the proof will not compromise the understanding of the subsequent discussion. However, it will give us the possibility to introduce an example that shows how the OSC of a particular transformed operator can be used to detect entanglement beyond the RC.

Theorem 2.6. *Let $\rho_{AB} \in \mathcal{S}_{AB}$ and consider, for $n \geq 1$, the set of $2\,n$ jointly linear (or jointly antilinear) superoperators*

$$\Gamma_i^A : \mathcal{H}_A \to \mathcal{H}_A, \quad \Gamma_i^B : \mathcal{H}_B \to \mathcal{H}_B, \quad \textit{for } i = 1, \dots, n \tag{2.28}$$

such that for some $\epsilon_A, \epsilon_B \geq 0$ and for any $\sigma_A^i \in \mathcal{S}_A$ and $\sigma_B^i \in \mathcal{S}_B$, with $i = 1, \dots, n$, one has

$$\sum_{i=1}^{n} \left\| \Gamma_i^A[\sigma_A^i] \right\|_2^2 \leq n\,\epsilon_A \quad \textit{and} \quad \sum_{i=1}^{n} \left\| \Gamma_i^B[\sigma_B^i] \right\|_2^2 \leq n\,\epsilon_B. \tag{2.29}$$

Moreover, define the following linear operator on $\mathcal{S}_{AB}$:

$$\rho_{AB}\left(\Gamma_{1,\dots,n}^{(A,B)}\right) := \frac{1}{n}\left(\sum_{i=1}^{n} \Gamma_i^A \otimes \Gamma_i^B[\rho_{AB}] + \sum_{i \neq j}^{n} \Gamma_i^A \otimes \Gamma_j^B[\rho_A \otimes \rho_B],\right) \tag{2.30}$$

where ρ_A, ρ_B are the reduced density matrices of ρ_{AB}. If ρ_{AB} is separable, then

$$\begin{aligned} \left\| \left[\rho_{AB}\left(\Gamma_{1,\dots,n}^{(A,B)}\right)\right]^R \right\|_1 \leq & \left(\epsilon_A + \frac{1}{n}\sum_{i<j}\left(\langle \Gamma_i^A[\rho_A], \Gamma_j^A[\rho_A]\rangle + c.c.\right)\right)^{1/2} \\ & \times \left(\epsilon_B + \frac{1}{n}\sum_{i<j}\left(\langle \Gamma_i^B[\rho_B], \Gamma_j^B[\rho_B]\rangle + c.c.\right)\right)^{1/2}, \end{aligned} \tag{2.31}$$

where c.c. stands for complex conjugate and $\langle \cdot\,, \cdot \rangle$ is the HS inner product in $\mathcal{L}(\mathcal{H}_{A(B)})$.

Proof. See Appendix A. $\square$

Corollary 2.6.1 (Generalised RC [66]). *Let $\rho_{AB} \in \mathcal{S}_{AB}$ be separable, and define*

$$\tau := \left(\frac{e^{i\omega} + e^{i(\omega+\theta+\phi)}}{2}\right)\rho_{AB} + \left(\frac{e^{i(\omega+\theta)} + e^{i(\omega+\phi)}}{2}\right)\rho_A \otimes \rho_B. \tag{2.32}$$

Then,

$$\left\|\tau^R\right\|_1 \leq \sqrt{\left[1 + \cos\theta \, \mathrm{Tr}(\rho_A^2)\right]\left[1 + \cos\phi \, \mathrm{Tr}(\rho_B^2)\right]}, \tag{2.33}$$

for any $\omega, \theta, \phi \in \mathbb{R}$.[4] *In particular, for $\omega = 0$ and $\phi = -\theta$ one has*

$$\left\|\left[\rho_{AB} + \cos\theta(\rho_A \otimes \rho_B)\right]^R\right\|_1 \leq \sqrt{\left[1 + \cos\theta \, \mathrm{Tr}(\rho_A^2)\right]\left[1 + \cos\theta \, \mathrm{Tr}(\rho_B^2)\right]} \tag{2.34}$$

for any $\theta \in [0, \pi]$.

The specific choice of ω and ϕ that yield the family of inequalities (2.34) is interesting for two reasons. On the one hand, the RC is clearly recovered for $\theta = \pi/2$. On the other hand, depending on the value of θ, new criteria which in some cases (i.e. for certain classes of quantum states) are stronger than the RC can be obtained. This result was first proved in [78] for the specific choice $\theta = \pi$, which gives the new criteria according to which separable states $\rho_{AB} \in \mathcal{S}_{AB}$ satisfy

$$\left\|\left[\rho_{AB} - \rho_A \otimes \rho_B\right]^R\right\|_1 \leq \sqrt{\left[1 - \mathrm{Tr}(\rho_A^2)\right]\left[1 - \mathrm{Tr}(\rho_B^2)\right]}. \tag{2.35}$$

The example below gives a class of states violating inequality Eq. (2.35).

Example. Let us consider the one-parameter family of 3×3 bound entangled states

[4] Although the presence of the parameter ω in the definition of τ is trivial, it will be convenient for deriving some specific inequalities. See [66, 78] for a comprehensive discussion about the possible cases encompassed by the inequalities (2.33).

introduced in [79], i.e.

$$\rho(a) = \frac{1}{8a+1}\begin{pmatrix} a & 0 & 0 & 0 & a & 0 & 0 & 0 & a \\ 0 & a & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a & 0 & 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 & a & 0 & 0 & 0 & a \\ 0 & 0 & 0 & 0 & 0 & a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1+a}{2} & 0 & \frac{\sqrt{1+a^2}}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a & 0 \\ a & 0 & 0 & 0 & a & 0 & \frac{\sqrt{1+a^2}}{2} & 0 & \frac{1+a}{2} \end{pmatrix} \tag{2.36}$$

with $0 \leq a \leq 1$. Then let us consider the two-parameter family of states given by the noisy version of the operators above, i.e.

$$\hat{\rho}(p,a) = p\,\rho(a) + (1-p)\frac{\mathbb{1}}{9}, \tag{2.37}$$

with $0 \leq p \leq 1$. A straightforward application of the RC and inequality (2.35) shows two things. First of all, all entangled states of the form Eq. (2.37) which are detected by the RC are also detected by criterion (2.35). Furthermore one finds that there exist values of p and a for which $\hat{\rho}(p,a)$ is entangled but only the new criterion (2.35) is able to reveal it, while the RC fails.

Finally, let us remark that $\theta = \pi$ is not the only choice allowing to conceive new criteria which outperforms the RC. For example, it is possible to show that also $\theta = 3\pi/4$ works. Nonetheless, numerical tests carried out in [66] showed that as long as the class of states of Eq. (2.37) is considered, $\theta = \pi$ returns the strongest separability criteria.

2.6 Operator Schmidt decomposition of transformed density operators

One of the aims of this thesis is to show that it is possible to exploit the OSD in order to detect and measure correlations. So far we have discussed detection, while throughout the next chapter we shall examine how to measures total correlations (TC) by decomposing the state describing the system of interest. Nevertheless, there is an instance in literature studying a measure which is not defined upon the OSD of the state itself, but on the decomposition of a certain transformed density operator [68]. In particular, instead of $\rho_{AB} \in \mathcal{S}_{AB}$ one can consider

$$\tilde{\rho}_{AB} := \left(\mathbb{1}_A \otimes \rho_B^{-1/2}\right) \rho_{AB} \left(\rho_A^{-1/2} \otimes \mathbb{1}_B\right), \tag{2.38}$$

where the inverse of the marginals ρ_A, ρ_B are defined on their supports. We refer to $\tilde{\rho}_{AB}$ as the *transformed operator* associated to ρ_{AB}. Before stating the results of [68] however, we first need to mention the most critical and indeed the defining features of the generic measure of TC. Specifically, we say that a function $f : \mathcal{S}_{AB} \to \mathbb{R}^+ \cup \{0\}$ is a measure of total correlations if it does not increase under the action of local channels and it if is minimal if and only it is evaluated on product states[5]. As a matter of fact, it is possible to prove that the OSC $\{\tilde{r}_i\}_{i=1}^d$ of $\tilde{\rho}_{AB}$ are indeed monotonic under LO, whilst the requirement about minimality is trivially satisfied whenever $i \geq 2$ (since $\tilde{r}_i = 0$ for any $i \geq 2$ when ρ_{AB} is a product).

Theorem 2.7 (total correlations from the OSD of the transformed operator [68])**.** *For any $\rho_{AB} \in \mathcal{S}_{AB}$, let $\{\tilde{r}_i\}$ and $\{\tilde{s}_i\}$ be the OSC of $\tilde{\rho}_{AB}$ and $\tilde{\sigma}_{AB}$, i.e. of the transformed operators associated to ρ_{AB} and $\sigma_{AB} := (\Lambda_A \otimes \Lambda_B)[\rho_{AB}]$, respectively. Then $\tilde{r}_i \geq \tilde{s}_i$ for any i, and in particular $\tilde{r}_1 = \tilde{s}_1 = 1$.*

In conclusion, except for the first OSC of the transformed operator, each other coefficient can serve as a proper measures of TC.[6]

[5] A more rigorous characterisation of measures of total correlations can be found in Sec. 3.2.

[6] The second OSC of the transformed operator satisfies an additional, stronger result: it is not

2.7 Operator Schmidt decomposition and tensor networks

In an attempt to review a relatively broad spectrum of occurrences of the OSD in the literature, here we take a departure from the topics considered so far, i.e. detection and quantification of correlations, and discuss the appearance of the OSD in the field of tensor network (TN). In particular, we take a look at the description of quantum many-body systems based on the formalism of matrix product states (MPS) and matrix product density operators (MPDO) [80–87].

TN methods are a set of algebraic and computational tools which provide an efficient way of approximating certain classes of quantum states. As well known in QM, such approximations are particularly useful when the size of the system increases, as this corresponds to an exponential growth of the Hilbert space that rapidly become intractable. With the term MPS one refers to a particular instance of TN methods whose significance mainly relies upon the existence of a *canonical form* [81], ensuring the possibility of representing exactly any given quantum state. In particular, the MPS representation is known to approximate a class of one–dimensional gapped systems with remarkable accuracy [88]. For what mixed states are concerned, one can still define the class of MPDO [89, 90] by analogy with the pure state case, but no canonical form has been found yet. One of the problems when dealing with MPDO is that in general local truncations do not preserve the positivity of the total tensor. However, if one consider the MPS representation of the purification of a mixed state instead of the associated MPDO, then local positivity is preserved. Yet, one could still wonder whether an efficient MPDO representing a mixed state implies the existence of efficient MPS representation of the purifying state. This question was considered in [69] and here we review the main results. To do so, let us rephrase the problem above in a proper mathematical setting.

Let ρ be a one-dimensional mixed state of N $d-$level systems with open boundary

additive under tensor product. This fact yield a certain data process inequality, that in turn bounds the set of states that can be generated under LO, when an arbitrary number of copies of a resource state is given. For it goes beyond the focus of this chapter, the reader is advised to refer to [68] for details.

conditions:

$$\rho = \sum_{\substack{i_1,\ldots,i_N=1 \\ j_1,\ldots,j_N=1}}^{d} \varrho_{i_1,\ldots,i_N}^{j_1,\ldots,j_N} \, |i_1,\ldots,i_N\rangle\langle j_1,\ldots,j_N| \,. \tag{2.39}$$

The MPDO representation of ρ is defined as

$$\rho = \sum_{\alpha_1=1}^{K_1} \sum_{\alpha_2=1}^{K_2} \cdots \sum_{\alpha_{N-1}=1}^{K_{N-1}} M_1^{\alpha_1} \otimes M_2^{\alpha_1,\alpha_2} \otimes \cdots \otimes M_N^{\alpha_{N-1}}, \tag{2.40}$$

where $M_i^{\alpha_{i-1},\alpha_i}$, for $1 < i < N$, are $d \times d$ matrices, $M_1^{\alpha_1}$ is a row vector and $M_N^{\alpha,\alpha_{N-1}}$ is a column vector of size d. Moreover, K_i for $1 \leq i \leq N$ is the minimal dimension such that Eq. (2.40) holds. In this multipartite setting, the OSR of ρ corresponds to

$$OSR(\rho) := \max_i K_i = K. \tag{2.41}$$

Now we consider the representation of ρ as MPS of the *local purification.* The latter is obtained by purifying the given state $\rho \in \mathcal{S}_S$ into the pure state $|\psi\rangle \in (\mathcal{H}_S \otimes \mathcal{H}_E)$, i.e. $\rho = \mathrm{Tr}_E \, |\psi\rangle\langle\psi|$ and consider the MPS of the purifying state:

$$|\psi\rangle = \sum_{\beta_1=1}^{K'_1} \sum_{\beta_2=1}^{K'_2} \cdots \sum_{\beta_{N-1}=1}^{K'_{N-1}} A_1^{\beta_1} \otimes A_2^{\beta_1,\beta_2} \otimes \cdots \otimes A_N^{\beta_{N-1}}, \tag{2.42}$$

where $A_i^{\beta_{i-1},\alpha_i}$, for $1 < i < N$, are $d \times d_{E_i}$ matrices where d_{E_i} is the size of the $i-$th ancilla, $A_1^{\beta_1}$ is a row vector of size $d \cdot d_{E_1}$ and $A_N^{\beta,\beta_{N-1}}$ is a column vector of size $d \cdot d_{E_{N-1}}$. The *purification rank* of ρ is defined as

$$\mathrm{rank_p}(\rho) := \max_i K'_i = K'. \tag{2.43}$$

We are now able to rephrase the original question in a more concrete way. Are the MPDO representation of ρ and the MPS description of the purifying state $|\psi\rangle$ equivalent by any means? Or can the latter be arbitrarily more costly (computationally speaking) than the former? In other words, we wonder if it is possible to find an upper bound to K' as a function of K only. The answer to this question is negative, and was proved through a counterexample in [69]. Without going into the detail, here it is enough to say that it is possible to construct a family of classical bipartite states such that their OSR is constant, while the purification rank grows unboundedly with the (arbitrary) dimension of the Hilbert space of the system.

Despite the negative result sketched above, [69] provides also two constructive purification methods which indeed return the sought upper bound, with the clause of considering also the eigenvalues of ρ. Specifically, these two constructions (that we do not review here for brevity) imply, respectively, that

$$K' \leq Dn^2, \tag{2.44}$$

$$K' \leq \mathcal{O}(D^{m-1}), \tag{2.45}$$

where n is the number of eigenvalues and m the number of different eigenvalues of ρ.

To conclude, the results just mentioned prove that it is impossible to find a description of mixed states in the MPDO formalism which is both efficient and locally positive semidefinite. Nonetheless, there exist approximations which are good, provided that ρ has a suitable spectrum in the sense of Eqs. (2.44) and (2.45).

Chapter 3

Measuring total correlations

This chapter is devoted to the definition and the analysis of possible measures of total correlation based on the OSD. As we have seen in the introductory chapter, correlations in pure states are due to entanglement only. On the other hand, mixed states can posses several kinds of quantum correlations, as well as several degrees of entanglement. For example, one can easily picture a situation such that a mere classical statistical dependency between the components of a bipartite quantum state is established. This is the case, e.g., of Alice and Bob receiving two local states drawn according to a probability distribution (which is known to both). The overall bipartite state is given by convex combination of the products of the possible states. This kind of correlations are fully classical in nature, and can appear either alone or in combination with quantum correlations, as in Eq. (1.9). Together, classical and quantum correlations stand at the basis of all the tasks and algorithms of quantum information theory [5,9–11,91–93]. For this reason, it would be useful to have a common framework for quantifying correlations altogether, and the OSD will turn out to be a suitable tool for this purpose.

This chapter is organized as follows. We first review some notions from classical and quantum information theory, paving the way for the introduction to the most widely used measure of total correlations in quantum information theory, i.e. the quantum mutual information. After that, we identify the criteria for testing the candidate measures of total correlations proposed in the remainder of this chapter. In Sec. 3.3 we recall some facts about the simplest measure based on the OSD that one can think of: (the logarithm of) the OSR, whose monotonicity under local channels and relation with discord was already observed by [94–96]. In Sec. 3.4 we introduce our first novel measure of total correlations based on the OSD, which has been defined along the lines of an entangled measure known as the G-Concurrence. In Sec. 3.5 we analyse a new possible measure based on the notion of *affinity* between quantum states. In Sec. 3.6, our last measure obtained by exploiting the OSD of the square root of the density matrix is presented. Lastly, in Sec. 3.7 we will compare the newly introduced measures of total correlations of Secs. 3.5 and 3.6 with the quantum mutual information.

3.1 Entropy in classical and quantum information theory

Entropy is the central concept of any theory that attempts to formalize the idea of information. Intuitively, a random variable X possesses a certain information content which can be thought of as the amount of new knowledge obtained – or the amount of previous uncertainty cleared away – once the value of X is unveiled. In classical information theory the *Shannon entropy* is used to quantify such information content carried by X. It is defined as $H(X) := -\sum_x p(x)\log p(x)$, where $p(x)$ belongs to a probability distribution and denotes the probability that X assumes the value x. Moreover, the logarithm is taken in base 2, and it is agreed that $0\log 0 = 0$ (this convention also holds for quantum versions of the Shannon entropy introduced in the remainder of this section). A more concrete way to look at the Shannon entropy is to consider some kind of source producing a string of independent, identically distributed random variables. Then H estimates the minimum number of bits per symbol needed to encode the information produced by the source. The fundamental concept of the Shannon can be exploited in order to obtain some means by which comparing the information content of several variables. For example, the *classical relative entropy* measures the similarity between two probability distributions $p(x)$ and $q(x)$ over a common index set, and it is defined as $H(p||q) := \sum_x p(x)\log(p(x)/q(x))$. An interesting instance of the relative entropy answers a more sophisticated questions about the mutual statistical dependency between two random variables. In particular, given two random variables

X and Y, their *classical mutual information* quantifies how much we can learn about X, once we know Y. It is given by $H(X:Y) := H(X) + H(Y) - H(X,Y)$, where $H(X,Y) := -\sum_{x,y} p(x,y) \log p(x,y)$ is the joint entropy of X and Y, which expresses the total uncertainty about the pair (X,Y). Finally, notice that the mutual information of X and Y equals the *closeness* (as measured by the classical relative entropy) between their joint probability distribution $p_{X,Y}$ and the product of the two individual ones p_X and q_Y, i.e. $H(X:Y) = H(p_{XY}||p_X \cdot p_Y)$.

The quantities defined above generalise readily to the quantum setting. Here, the density operator takes on the role of a probability distribution, and the Shannon entropy is replaced by the *von Neumann entropy*, defined as

$$S(\rho) := -\operatorname{Tr}(\rho \log \rho) = -\sum_i \lambda_i \log \lambda_i, \tag{3.1}$$

where λ_i are the eigenvalues of ρ. The von Neumann entropy it is often said to measure the *mixedness* of a quantum state [97]. To make sense of this statement it is enough to observe the values that $S(\rho)$ assumes, and to recognise that they reflect the degree of ignorance of an observer about the state of the system he/she is interested in. Indeed, pure states (corresponding to the situation of an observer having complete knowledge about the state of a quantum system) have zero entropy, while the completely mixed state (corresponding to the situation of an observer having complete ignorance about the state of a quantum system) returns the maximum value for S, that is $\log d$. The extremal values of S mirror one of the distinctive features of QM, when coming to the maximally entangled state. Without a classical analog in fact, the maximally entangled state has zero entropy, but it is such that the information retained by the composite state is completely lost when one of the two subsystems is traced out, revealing completely disordered marginals. Furthermore, the von Neumann entropy satisfies a number of physically desirable properties, for a careful analysis of which we recommend the usual, distinguished references [13] and [28].

As for classical probability distributions, the quantum version of the relative entropy allows to compare the similarity between two states. It is defined, for any ρ and σ, as

$S(\rho||\sigma) := \mathrm{Tr}(\rho \log \rho) - \mathrm{Tr}(\rho \log \sigma)$.[1] Now let $\rho_{AB} \in \mathcal{S}_{AB}$ be bipartite and consider the relative entropy between ρ_{AB} and the tensor product of its marginals, i.e. $S(\rho_{AB}||\rho_A \otimes \rho_B)$. The departure of ρ_{AB} from being a product state gives the *quantum mutual information* (QMI), which in general is defined as

$$\begin{aligned} \mathbf{I}(A:B)_{\rho_{AB}} &= S(\rho_{AB}||\rho_A \otimes \rho_B) \\ &= S(A) + S(B) - S(AB), \end{aligned} \tag{3.2}$$

where $S(X)$ denotes the von Neumann entropy of ρ_X.

The QMI satisfies several desirable properties. To name a few, the QMI is non negative and it is symmetric under interchange of variables. Moreover, besides being a natural measure of distance between a bipartite quantum state and the product of its marginals (as already observed), it is possible to prove that the QMI is indeed the minimum of the relative entropy between ρ_{AB} and *any* tensor product state [13]. Also, the definition Eq. (3.2) directly implies that the QMI vanishes if and only if ρ_{AB} is a product state [13,28]. Another important result concerns the monotonicity of the QMI:

Proposition 3.1 (Monotonicity of QMI [98])**.** *For any $\rho_{AB} \in \mathcal{S}_{AB}$ and quantum channels Λ_A, Λ_B, one has*

$$\mathbf{I}(A:B)_{\rho'_{AB}} \leq \mathbf{I}(A:B)_{\rho_{AB}}, \tag{3.3}$$

where $\rho'_{AB} = (\Lambda_A \otimes \Lambda_B)[\rho_{AB}]$.

Proof. We prove that the monotonicity of the QMI under local channels is equivalent to its monotonicity under local trace, and for simplicity we do it for one local channel only, say Λ_A. Then, one can conclude the proof by noticing that the monotonicity under local trace, in turn, corresponds to the strong subadditivity of the von Neuman entropy (for a complete proof of the latter we refer to [13]). Thanks to Stinespring's dilation theorem 1.4, the action of Λ_A on the subsystem A can be simulated by introducing a

[1]In both its classical and quantum version, the relative entropy between two probability distributions p and q is finite if and only if the support of p is included in the support of q.

third system C, initially in the state $|0\rangle\langle 0|_C$, and a unitary interaction U_{AC} between A and C. Then, the action of Λ_A on A is equivalent to the action of U_{AC} followed by tracing out the ancilla. It is convenient to define the states of the system of interest at different stages, the latter corresponding to the operations used to characterise the channels. Thus we denote the state after the tensorization as $\rho_{ABC} = \rho_{AB} \otimes |0\rangle\langle 0|_C$, after the unitary operation as $\rho'_{ABC} = U_{AC}(\rho_{AB} \otimes |0\rangle\langle 0|_C)U^\dagger_{AC}$, and the final state after the partial trace as $\rho'_{AB} = \mathrm{Tr}_C(\rho'_{ABC})$, that is $\rho'_{AB} = (\Lambda_A \otimes \mathrm{id})[\rho_{AB}]$. Now since C starts in a product state with AB, it follows that

$$\mathbf{I}(A:B)_{\rho_{AB}} = \mathbf{I}(A:BC)_{\rho_{ABC}} = \mathbf{I}(A:BC)_{\rho'_{ABC}}. \tag{3.4}$$

On the other hand, since the partial trace cannot increase the QMI [13], one has

$$\mathbf{I}(A:B)_{\mathrm{Tr}_C(\rho'_{ABC})} \leq \mathbf{I}(A:B)_{\rho'_{ABC}}. \tag{3.5}$$

The last two observations together yield $\mathbf{I}(A:B)_{\rho'_{AB}} \leq \mathbf{I}(A:B)_{\rho_{AB}}$. □

3.2 Requirements for a measures of total correlations

The QMI is used in QM as a measure of total correlation. As seen earlier, it satisfies several properties, including being minimal if and only if it is evaluated on product states and being non increasing under local channels. These two requirements assume a central role in this chapter, since they can be considered as the minimal requirements that any measure of total correlations must obey in order to be meaningful. To put it more rigorously, let us first consider the following definition.

Definition 3.1. Given a function f from the set of quantum states to the (nonnegative) real numbers, we say that f is a *measure of correlations* if the following conditions hold true:

1. ρ_{AB} is a product state $\Rightarrow f(\rho_{AB}) = 0$,

2. f is monotonically non increasing under local operation, i.e. for any $\rho_{AB} \in \mathcal{S}_{AB}$

and any local maps $\Lambda_A : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_{A'})$ and $\Lambda_B : \mathcal{L}(\mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_{B'})$,

$$f(\rho_{AB}) \geq f(\Lambda_A \otimes \Lambda_B[\rho_{AB}]). \tag{3.6}$$

Moreover, if $f(\rho_{AB}) = 0$ if and only if ρ_{AB} is a product state, we say that f is a *measure of total correlations* or that is a *faithful* measure of correlations.

The first requirement follows from the statistical independence existing between the parties of a product states. In other words, since product states carry no correlations whatsoever, any measure of correlations – whether they be faithful or not – must account for this fact. The second requirement comes from the observation that local maps cannot increase any global property of the state, e.g. cannot increase entanglement. On the other hand, local operations alone cannot increase classical correlations either, since some sort of communication between A and B would be needed for that, for example a classical communication channel which allow the establishment of shared randomness between the two parties. Eq. (3.6) reflects, mathematically, the observation above. Of course, there are other plausible requirements that one might want to add to the list. It would be reasonable, for example, to devise measures which are also completely additive, indicating that the correlations of the product of several bipartite states would equals the sum of the correlations of the individual states. Again, it looks fair to ask that the sought measure of total correlation be maximal for the maximally entangled state. However, these extra conditions do not qualify as necessary. For this reason, in the reminder of this chapter we will test the proposed measures of correlations only against the two necessary requirements given in Definition 3.1, while for the measures of total correlations the additional requirement of being faithful will be demanded.

3.3 The operator Schmidt rank

The simplest measure satisfying the desired requirements and which can be obtained from the OSD is as a function of the OSR. In particular we claim that

$$\mathbf{O}(\rho_{AB}) = \log(\mathrm{OSR}(\rho_{AB})) \tag{3.7}$$

is a meaningful measure of total correlations, i.e. the following theorem holds.

Theorem 3.1. $\mathbf{O}(\rho_{AB})$ *attains its minimum if and only if* ρ_{AB} *is a product state, and it is non–increasing under local operations, i.e. for any* $\rho_{AB} \in \mathcal{S}_{AB}$ *and any local channels* Λ_A, Λ_B *one has*

$$\mathbf{O}(\Lambda_A \otimes \Lambda_B[\rho_{AB}]) \leq \mathbf{O}(\rho_{AB}). \tag{3.8}$$

Proof. The first claim follows trivially from the properties of the logarithm function, being zero if and only if its argument (here the OSR) equals one, which happens only for product states. For the monotonicity, let $\mathscr{L}_A$ and $\mathscr{L}_B$ be the transfer matrices associated to Λ_A and Λ_B, respectively, cf. Definition 1.2. For simplicity, first consider the case of one local map only, e.g. when $\rho'_{AB} = \Lambda_A \otimes \mathrm{id}_B[\rho_{AB}]$. Let $\{C_k\}$ and $\{D_l\}$ be local bases for the subsystem A and B, respectively, and consider the expansion of ρ_{AB} on these bases, i.e. $\rho_{AB} = \sum_{kl} \mathcal{C}(\rho_{AB})_{kl} C_k \otimes D_l$, with $\mathcal{C}(\rho_{AB})_{kl}$ the entries of the correlations matrix for such a choice of local bases. Then, $\rho'_{AB} = \sum_{kl} \mathcal{C}(\rho_{AB})_{kl}\, \Lambda[C_k] \otimes D_l$. It follows that the generic element of the correlation matrix of ρ'_{AB} reads

$$\begin{aligned}
\mathcal{C}(\rho'_{AB})_{ij} &= \mathrm{Tr}\left[C_i^\dagger \otimes D_j^\dagger \left(\sum_{kl} \mathcal{C}(\rho_{AB})_{kl}\, \Lambda_A[C_k] \otimes D_l\right)\right] \\
&= \sum_{kl} \mathcal{C}(\rho_{AB})_{kl}\, \mathrm{Tr}(C_i^\dagger \Lambda_A[C_k])\, \mathrm{Tr}(D_j^\dagger D_l) \\
&= \sum_{k} \mathcal{C}(\rho_{AB})_{kj} (\mathscr{L}_A)_{ik},
\end{aligned} \tag{3.9}$$

where we have used $(\mathscr{L}_A)_{ik} = \mathrm{Tr}(C_i^\dagger \Lambda_A[C_k])$. Then,

$$\mathcal{C}(\rho'_{AB}) = \mathscr{L}_A \mathcal{C}(\rho_{AB}). \tag{3.10}$$

In a similar fashion, if we consider also the action of the local map on B we obtain

$$\mathcal{C}(\rho'_{AB}) = \mathcal{L}_A \mathcal{C}(\rho_{AB}) \mathcal{L}_B^T. \tag{3.11}$$

We conclude by observing that since matrix multiplication cannot increase the rank, then

$$\begin{aligned} \mathrm{OSR}(\rho'_{AB}) &= \mathrm{rank}[\mathcal{L}_A \mathcal{C}(\rho_{AB}) \mathcal{L}_B^T] \\ &\leq \mathrm{rank}[\mathcal{C}(\rho_{AB})] \\ &= \mathrm{OSR}(\rho_{AB}), \end{aligned} \tag{3.12}$$

and the claim follows from the monotonicity of the logarithm function. □

Although the logarithm of the OSR gives a straightforward and direct way of making sense of the amount of correlations of a given state, it suffers of a lack of sensitivity due to its integer valued nature. Here we present an example of this fact which employs the class of isotropic states, which we recall are given by convex mixtures of the completely mixed and the maximally entangled state, parametrized by the relative weight of the combination.

Example. Let us consider the generic isotropic state [99] in $\mathcal{S}_{AB}$

$$\rho_{AB}(p) = (1-p)\frac{\mathbb{1}}{d^2} + p\,|\psi^+\rangle\langle\psi^+|, \tag{3.13}$$

where $0 \leq p \leq 1$, $\mathbb{1} \equiv \mathbb{1}_{AB}$ and $|\psi^+\rangle = \frac{1}{\sqrt{d}}\sum_i^d |ii\rangle$ is the maximally entangled state in $\mathcal{H}_{AB}$. Further assume that $d = d_A = d_B$. In order to compute the OSC, consider the orthonormal basis of $\mathcal{L}(\mathcal{H}_A)$ and $(\mathcal{H}_B)$ given by the matrix units $\{|i\rangle\langle j|\}_{i,j=1}^d$ and

$\{|k\rangle\langle l|\}_{k,l=1}^d$. The correlation matrix elements of the state in Eq. (3.13) are easily found:

$$\begin{aligned}
\mathcal{C}[\rho_{AB}(p)]_{ij,kl} &= \mathrm{Tr}\left[|j\rangle\langle i| \otimes |l\rangle\langle k|\, \rho_{AB}(p)\right] \\
&= \frac{1-p}{d^2}\,\mathrm{Tr}\left(|j\rangle\langle i| \otimes |l\rangle\langle k|\right) + \frac{p}{d}\,\mathrm{Tr}\left(|j\rangle\langle i| \otimes |l\rangle\langle k| \sum_{a,b=1}^{d} |a\rangle\langle b| \otimes |a\rangle\langle b|\right) \\
&= \frac{1-p}{d^2}\;\delta_{ij}\delta_{kl} + \frac{p}{d}\sum_{a,b=1}^{d} \delta_{ia}\delta_{bj}\delta_{ka}\delta_{bl} \\
&= \frac{1-p}{d^2}\;\delta_{ij}\delta_{kl} + \frac{p}{d}\delta_{ik}\delta_{jl},
\end{aligned} \tag{3.14}$$

where in the second equality we have used the fact that the maximally entangled state can be expressed as $\sum_{a,b=1}^{d} |a\rangle\langle b| \otimes |a\rangle\langle b|$. Then Eq. (3.14) returns

$$\begin{aligned}
\mathcal{C}[\rho_{AB}(p)] &= \sum_{ij,kl} [\mathcal{C}\rho_{AB}(p)]_{ij,kl}\, |i\rangle\langle j| \otimes |k\rangle\langle l| \\
&= \frac{1-p}{d}\; |\psi^+\rangle\langle\psi^+| + \frac{p}{d}\,\mathbb{1},
\end{aligned} \tag{3.15}$$

which gives $\mathrm{OSC}[\rho_{AB}(p)] = (1/d, p/d, \ldots, p/d)$. It is now clear why $\mathbf{O}(\rho_{AB})$ is a very coarse measure. In fact, it is zero for $p = 1$, i.e. for the maximally mixed state, while it is always maximal for any $p > 0$,[2]. On the other hand, notice that isotropic states are entangled only when $p > 1/(d+1)$.

As a final remark, notice that since the OSR of a pure state $|\psi\rangle\langle\psi|$ equals the square of the Schmidt rank of ψ, one concludes that the OSR of a pure state is as coarse-grained as the corresponding SR.

3.4 The operator G-Concurrence

In this section we review the properties of a function of the OSC, which we call *operator G-concurrence* (OGC), that is remarkably more sensitive[3] than $\mathbf{O}(\rho_{AB})$. We begin

[2]This observation can be rephrased by saying that when restricted to isotropic states, the OSR coincides with an indicator function of correlations: it is constant and non-zero whenever there are correlations (in the class).

[3]As we are going to see in the following, the OGC vanishes for any state with non–maximal OSR. Nevertheless, when restricted to state with maximal OSR, the OGC varies smoothly with the OSC,

by reporting the origin of the OGC, which was inspired by a family of entanglement measures known as *concurrence monotone* [100], member of which is the renowned *G-Concurrence.*

Definition 3.2 (Concurrence monotones [100]). Consider a $d\times d-$dimensional bipartite pure state $|\psi\rangle$ with Schmidt coefficients λ_i, with $i = 0, \dots, d-1$. The concurrence monotones $C_k(\psi)$, with $k = 1, \dots, d$, of $|\psi\rangle$ are given by

$$C_k(\psi) := \left(\frac{M^{[k]}(\lambda)}{M^{[k]}(1/d)}\right)^{1/k}, \tag{3.16}$$

where $M^{[k]}(\lambda)$ are the $k^{th}-$order elementary symmetric polynomial of $\lambda_0, \lambda_1, \dots, \lambda_{d-1}$ defined as in Eqs. (2.23). The normalizing factor $M^{[k]}(1/d)$ is the $k^{th}-$order elementary symmetric polynomial when $\lambda_i = 1/d$ for any $i = 0, \dots, d-1$. For bipartite mixed states ρ_{AB}, the concurrence monotones $C(\rho_{AB})$ of ρ_{AB} are defined as the average C_k of the pure states in the decomposition, minimized over all the possible decomposition of ρ_{AB}, in formulae

$$C_k(\rho_{AB}) = \min_i p_i\, C_k(|\psi\rangle), \quad \rho_{AB} = \sum_i p_i\, |\psi_i\rangle\langle\psi_i|\,. \tag{3.17}$$

Finally, the G-concurrence is given by $C_k(|\psi\rangle)$ with $k = d$, and it is equal to the geometric mean of the Schmidt coefficients

$$C_d(\psi) = d(\lambda_0\lambda_1 \dots \lambda_{d-1})^{1/d}. \tag{3.18}$$

By analogy with the definition above one could try to define a family of measures of correlations which exploits the ESP in the OSC of a generic quantum states, instead of using functions of the Schmidt coefficients and their convex roof extensions. However, this is possible only for ESP of a certain order. For example, we will show at the end of this section that the $1^{st}-$order ESP in the OSC is not monotone under LO, thus it does not suit our purposes. On the other hand, the generalization of Eq. (3.18) to a function involving the OSC looks promising and led us to consider the following

and can be considered a highly sensitive measure of correlations, especially if compared with $\mathbf{O}(\rho_{AB})$.

candidate measure of correlations.

Definition 3.3. The operator G-concurrence $\mathbf{G}(\rho_{AB})$ of a bipartite state $\rho_{AB} \in \mathcal{S}_{AB}$ is the geometric mean of its OSC, i.e.

$$\mathbf{G}(\rho_{AB}) := \prod_{i=1}^{d^2} r_i^{1/d^2}, \tag{3.19}$$

where the product runs over all the coefficients, including the null ones, and we are denoting $d = \min\{d_A, d_B\}$. Notice moreover that the OGC can be written as a function of the correlation matrix:

$$\mathbf{G}(\rho_{AB}) := \left[\det\left(\mathcal{C}(\rho_{AB})^\dagger \mathcal{C}(\rho_{AB})\right)\right]^{1/(2d^2)}, \tag{3.20}$$

and as a function of the $(d^2)^{th}-$order ESP defined in Eq. (2.23):

$$\mathbf{G}(\rho_{AB}) := \left[M^{[d^2]}(\rho_{AB})\right]^{1/d^2}. \tag{3.21}$$

That the OGC attains its minimum for product states follows trivially from having included in its definition also the possible zero coefficients. This implies, in particular, that any state with OSR strictly less than maximal has vanishing OGC. However, this argument deserves a separate discussion which we reserve to deal with at the end of this section, after assessing several properties of the OGC. To begin with, we demonstrate here the monotonicity of the OGC under LO. We consider two cases. First, we regard the action of the local maps as being *deterministic*, namely we consider the evolved state as the result of one single transformation – described by $\Lambda_A \otimes \Lambda_B$ – applied to a generic input state. Then, we investigate the monotonicity *on average*, namely when the decomposition of a quantum channel into subchannels is taken into account. For the sake of clarity, let us recall the following definition.

Definition 3.4. Given a channel $\Lambda \in \mathcal{C}(\mathcal{H}, \mathcal{H}')$, a collection of linear maps $\{\Lambda_i\}$ such that $\Lambda = \sum_i \Lambda_i$ is a channel is called an *instrument*, while the individual Λ_i are called *subchannels*.[4]

Then, given an instrument $\{\Lambda_i\}$, the evolution ρ under $\Lambda = \sum_i \Lambda_i$ can be seen as made of *branches* $\rho \mapsto \Lambda_i[\rho]$, each one occurring with probability $q_i := \mathrm{Tr}(\Lambda_i[\rho])$. On the other hand, the global evolution $\Lambda[\rho]$ can be regarded as the situation when the information about which branch of the evolution the state went through is lost. With this wording, a candidate measure of total correlations f is said to be monotonically non increasing, *on average*, under LO when

$$f(\rho) \geq \sum_i q_i f(\Lambda_i[\rho]). \tag{3.22}$$

In what follows we show that the OGC does not increase under deterministic LO. On the other hand, our reasoning allows to establish the monotonicity on average only when the subchannels are given by single Kraus operators. Finally, we consider the general case of arbitrary subchannels and show that the average monotonicity holds if the LO are unital.

In order to tackle the monotonicity of the OGC under deterministic LO, certain results about the spectral properties of quantum channels are needed. When we speak about the *spectrum* of a quantum channel $\Lambda : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$, we are tacitly implying that its domain and range coincide, i.e. that $\dim \mathcal{H} = \dim \mathcal{H}'$. In this situation, the transfer matrix of Λ is square, and it becomes meaningful to consider its eigenvalues. From another perspective, one can see the eigenvalues of Λ are those complex numbers λ such that

$$\Lambda[A] = \lambda A, \tag{3.23}$$

where the vector space structure of $\mathcal{L}(\mathcal{H})$ makes it possible to think of the operator $A \in \mathcal{L}(\mathcal{H})$ as a vector – an eigenvector of Λ in fact. Another observation which will turn useful is that since channels are hermiticity preserving, linearity implies $\Lambda[A^\dagger] = \bar{\lambda} A^\dagger$. For the sake of clarity we provide the following formal definition.

Definition 3.5 (Spectral radius of positive linear maps [57])**.** Given a linear map

[4]Observe that since channels are trace preserving, subchannels are trace non–increasing.

$\Lambda : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$, its *spectrum* $\mathrm{spec}(\Lambda)$ is defined as the set of eigenvalues of the transfer matrix $\mathscr{L}$ associated to Λ. Moreover, its *spectral radius* is defined as $\varrho(\Lambda) := \sup\{|\lambda| \mid \lambda \in \mathrm{spec}(\Lambda)\}$.

In view of the last definition, here we present a lemma which is necessary to prove a useful result about the determinant of positive and trace non–increasing preserving (PTN) linear maps.

Lemma 3.1 (Spectrum of positive maps [57])**.** *For a positive linear map* $\Lambda : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$*, one has*

$$\varrho(\Lambda) \leq \|\Lambda[\mathbb{1}]\|_\infty. \tag{3.24}$$

In particular, if Λ *is trace preserving or unital, then* $\varrho(\Lambda) = 1$*.*

Proof. The proof comes from a result of functional analysis, known as Russo-Dye theorem [101], which implies that a positive map Λ is such that $\|\Lambda[A]\|_\infty \leq \|\Lambda[\mathbb{1}]\|_\infty \|A\|_\infty$. By comparing this fact with Eq. (3.23) one has $|\lambda| \|A\|_\infty \leq \|\Lambda[\mathbb{1}]\|_\infty \|A\|_\infty$, which in turn implies Eq. (3.24). For the second part of the lemma observe that when Λ is unital, i.e. $\Lambda[\mathbb{1}] = \mathbb{1}$, there must be an eigenvalue equal to one. The same result holds when Λ is trace preserving, i.e. $\Lambda^\dagger[\mathbb{1}] = \mathbb{1}$, since the transfer matrix of $\Lambda^\dagger$ is given by $\mathscr{L}^T$ (where $\mathscr{L}$ is the transfer matrix of Λ), and transposition does not change the spectrum. □

We finally have the following.

Theorem 3.2 (Determinant of PTN linear maps [57])**.** *Let* $\Lambda : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ *be a positive and trace preserving linear map. Then* $\det(\mathscr{L}) \leq 1$*, where* $\mathscr{L}$ *is the transfer matrix of* Λ*.*

Proof. Since Λ is hermiticity preserving, Eq. (3.23) together with its adjoint version tell us that the eigenvalues of $\mathscr{L}$ are either real or come in complex conjugate pairs. It follows that also their product, i.e. the determinant of Λ, must be real. In addition, since Λ is trace preserving, Lemma 3.1 gives $\varrho(\Lambda) = 1$, hence the claim. □

Theorem 3.3. $\mathbf{G}(\rho_{AB})$ *is non–increasing under LO, i.e. for any* $\rho_{AB} \in \mathcal{S}_{AB}$ *and any local channels* $\Lambda_A : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_A)$ *and* $\Lambda_B : \mathcal{L}(\mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_B)$*, one has*

$$\mathbf{G}(\rho_{AB}) \geq \mathbf{G}(\Lambda_A \otimes \Lambda_B[\rho_{AB}]). \tag{3.25}$$

Proof. For notational convenience, we first prove the monotonicity under the action of a local channel acting on A only, i.e. we show that $\mathbf{G}(\Lambda_A \otimes \mathrm{id}_B[\rho_{AB}]) \leq \mathbf{G}(\rho_{AB})$. Define $\rho'_{AB} := (\Lambda_A \otimes \mathrm{id}_B[\rho_{AB}])$. Then, Eq. (3.10) gives

$$\begin{aligned}
\mathbf{G}(\rho'_{AB}) &= \left[\det(\mathcal{C}(\rho'_{AB})^\dagger \mathcal{C}(\rho'_{AB}))\right]^{1/2(d^2)} \\
&= \left[\det(\mathcal{C}(\rho_{AB})^\dagger \mathscr{L}_A^\dagger \mathscr{L}_A \mathcal{C}(\rho_{AB}))\right]^{1/(2d^2)} \\
&= \mathbf{G}(\rho_{AB}) |\det(\mathscr{L}_A)|^{1/d^2} \\
&\leq \mathbf{G}(\rho_{AB}),
\end{aligned} \tag{3.26}$$

where the last inequality comes from Theorem 3.2. In order to obtain the monotonicity for local channels acting on both subsystems instead, it is enough to define $\rho'_{AB} := (\Lambda_A \otimes \Lambda_B[\rho_{AB}])$, invoke Eq. (3.11) rather than Eq. (3.10), and follow similar steps as above. $\square$

In order to assess the average monotonicity of the OGC, we first consider the case of a local instrument whose subchannels are given by single Kraus operators, as anticipated. We conveniently define

$$\Lambda_X[\,\cdot\,] = \sum_i \Lambda_{X,i}[\,\cdot\,] \tag{3.27}$$

$$\Lambda_{X,i}[\,\cdot\,] = K_{X,i} \cdot K_{X,i}^\dagger, \quad \text{for any } i, \tag{3.28}$$

where $X = A, B$ denote one part of a bipartite quantum system. We have the following theorem.

Theorem 3.4. $\mathbf{G}(\rho_{AB})$ *is non–increasing, on average, under any LO whose subchan-*

nels are given by single Kraus operators, i.e.

$$\mathbf{G}(\rho_{AB}) \geq \sum_{ij} q_{ij} \mathbf{G}(\rho'_{AB,ij}), \tag{3.29}$$

with $q_{ij} = \mathrm{Tr}(\Lambda_{A,i} \otimes \Lambda_{B,i}[\rho_{AB}])$*,* $\rho'_{AB,ij} = (\Lambda_{A,i} \otimes \Lambda_{B,j}[\rho_{AB}])/q_{ij}$*, and* $\Lambda_{A,i}, \Lambda_{B,j}$ *are as in Eqs.* (3.27)–(3.28).

Proof. For simplicity, we consider again the action of a local map acting on one subsystem only – say A – hence we redefine $q_i = \mathrm{Tr}(\Lambda_{A,i} \otimes \mathrm{id}_B[\rho_{AB}])$ and $\rho'_{AB,i} = (\Lambda_{A,i} \otimes \mathrm{id}_B[\rho_{AB}])/q_i$. Let $\Lambda_A = \sum_i \Lambda_{A,i}$. Moreover, let $\mathscr{L}_A$ be the transfer matrix of Λ_A and, accordingly, let $\mathscr{L}_{A,i}$ represents $\Lambda_{A,i}$ for any i. Similarly to the proof of Theorem 3.3 and invoking again Eq. (3.10), observe that

$$\begin{aligned}\mathcal{C}(\rho'_{AB,i}) &= \mathcal{C}\left(\frac{\Lambda_{A,i} \otimes \mathrm{id}_B[\rho_{AB}]}{q_i}\right) \\ &= \frac{1}{q_i}\mathscr{L}_{A,i}\mathcal{C}(\rho_{AB}).\end{aligned} \tag{3.30}$$

Then, by using

$$\begin{aligned}\det\Big(\mathcal{C}(\rho'_{AB,i})^\dagger \mathcal{C}(\rho'_{AB,i})\Big) &= q_i^{-2d^2} \det\Big(\mathcal{C}(\rho_{AB})^\dagger \mathscr{L}_{A,i}^\dagger \mathscr{L}_{A,i} \mathcal{C}(\rho_{AB})\Big) \\ &= q_i^{-2d^2}\, |\det \mathscr{L}_{A,i}|^2 \det\Big(\mathcal{C}(\rho_{AB})^\dagger \mathcal{C}(\rho_{AB}),\Big)\end{aligned} \tag{3.31}$$

one finds that

$$\begin{aligned}\sum_i q_i \mathbf{G}(\rho'_{AB,i}) &= \sum_i q_i \left[\det\Big(\mathcal{C}(\rho'_{AB,i})^\dagger \mathcal{C}(\rho'_{AB,i})\Big)\right]^{1/(2d^2)} \\ &= \sum_i q_i \left[q_i^{-2d^2}\, |\det \mathscr{L}_{A,i}|^2 \det\Big(\mathcal{C}(\rho_{AB})^\dagger \mathcal{C}(\rho_{AB}),\Big)\right]^{1/(2d^2)} \\ &= \sum_i |\det \mathscr{L}_{A,i}|^{1/d^2}\, \mathbf{G}(\rho_{AB}).\end{aligned} \tag{3.32}$$

To conclude, it is enough to prove that $\sum_i |\det(\mathscr{L}_{A,i})|^{1/d^2} \leq 1$. To do so one can use the arithmetic-geometric mean inequality [102–104], which says that for any positive definite matrix M of dimension $d^2 \times d^2$ one has

$$\det(M)^{1/d^2} \leq \frac{Tr(M)}{d^2}. \tag{3.33}$$

By applying the inequality above to our context, and recalling that the matrix representation of each $\Lambda_{A,i}$ is given by $\mathscr{L}_{A,i} = \sum_i K_{A,i} \otimes \overline{K}_{A,i}$, we find:

$$\begin{aligned}
\sum_i \left[\det\left(\mathscr{L}_{A,i}^{\dagger}\mathscr{L}_{A,i}\right)\right]^{1/2d^2} &\leq \frac{1}{d}\sum_i \sqrt{\mathrm{Tr}\left(\mathscr{L}_{A,i}^{\dagger}\mathscr{L}_{A,i}\right)} \\
&= \frac{1}{d}\sum_i \sqrt{\mathrm{Tr}\left[\left(K_{A,i}\otimes\overline{K}_{A,i}\right)^{\dagger}\left(K_{A,i}\otimes\overline{K}_{A,i}\right)\right]} \\
&= \frac{1}{d}\sum_i \mathrm{Tr}(K_{A,i}^{\dagger}K_{A,i}) \\
&= 1,
\end{aligned} \tag{3.34}$$

where in the last line the trace preserving property of Λ_A, corresponding to $\sum_i K_i^{\dagger} K_i = I$, has have been used. □

It remains to study the case of local instruments whose subchannels are given by an arbitrary combination of Kraus operators. In particular, here we consider maps of the form

$$\Lambda_X[\,\cdot\,] = \sum_m \Lambda_{X,m}[\,\cdot\,] \tag{3.35}$$

$$\Lambda_{X,m}[\,\cdot\,] = \sum_{i\in I_m} K_{X,i}\cdot K_{X,i}^{\dagger}, \tag{3.36}$$

where the index sets I_m are arbitrary but satisfy $I_m \cap I_{m'} = \emptyset$, for $m \neq m'$. We have the following.

Theorem 3.5. $\mathbf{G}(\rho_{AB})$ *is non–increasing, on average, under any unital LO, i.e.*

$$\mathbf{G}(\rho_{AB}) \geq \sum_{mn} q_{mn}\mathbf{G}(\rho'_{AB,mn}), \tag{3.37}$$

where $q_{mn} = \mathrm{Tr}(\Lambda_{A,m}\otimes\Lambda_{B,n}[\rho_{AB}])$*,* $\rho'_{AB,mn} = (\Lambda_{A,m}\otimes\Lambda_{B,n}[\rho_{AB}])/q_{mn}$*, and* $\Lambda_A[\mathbb{1}] =$

$\Lambda_B[\mathbb{1}] = \mathbb{1}$.

Proof. As before, we prove the claim for a local channel acting on A, i.e. we consider $q_m = \sum_m \mathrm{Tr}(\Lambda_{A,m} \otimes \mathrm{id}_B[\rho_{AB}])$ and $\rho'_{AB,m} = (\Lambda_{A,m} \otimes \mathrm{id}_B[\rho_{AB}])/q_m$. The extension to general two-sided LO is trivial and easily attainable using Eq. (3.11). Following similar steps as in Theorem 3.4 [cf. Eq. (3.32)] we obtain

$$\sum_m q_m \mathbf{G}(\rho'_{AB,m}) = \sum_m |\det(\mathscr{L}_{A,m})|^{1/d^2} \mathbf{G}(\rho_{AB}), \tag{3.38}$$

thus the theorem is proven if we show that $\sum_m |\det(\mathscr{L}_{A,m})|^{1/d^2} \leq 1$.

Let Λ_A and $\Lambda_{A,m}$ be as in Eqs. (3.35) and Eqs. (3.36), respectively, and define

$$\begin{aligned} \hat{\Lambda}_A = \Lambda_A^\dagger \Lambda_A &= \sum_m \Lambda_{A,m}^\dagger \sum_n \Lambda_{A,n} \\ &= \sum_{m=n} \Lambda_{A,m}^\dagger \Lambda_{A,n} + \sum_{m\neq n} \Lambda_{A,m}^\dagger \Lambda_{A,n}, \end{aligned} \tag{3.39}$$

where the composition of maps $\Lambda_A^\dagger \Lambda_A \equiv \Lambda_A^\dagger \circ \Lambda_A$ is understood. Since both Λ_A and $\Lambda_A^\dagger$ are CPTP maps, also $\hat{\Lambda}_A$ is CPTP. Hence, the two sums on the rhs of Eq.(3.39) can be regarded as subchannels. In particular, this implies that $\sum_m \Lambda_{A,m}^\dagger \Lambda_{A,m}[\,\cdot\,]$ is a CP trace non-increasing maps. Every $\Lambda_{A,m}$ is Hermiticity preserving, and the matrix associated to $\Lambda_{A,m}^\dagger$ is given by $\mathscr{L}_{A,m}^\dagger$. Furthermore, since composition of maps translates into the product of the associated matrices, the matrix representation of $\sum_m \Lambda_{A,m}^\dagger \Lambda_{A,m}[\,\cdot\,]$ is given by $\sum_m \mathscr{L}_{A,m}^\dagger \mathscr{L}_{A,m}$. Then Theorem 3.2 applies, giving

$$\det\left(\sum_m \mathscr{L}_{A,m}^\dagger \mathscr{L}_{A,m}\right) \leq 1. \tag{3.40}$$

We conclude by enforcing the Brunn–Minkowski inequality [105,106] for determinants, saying that for any pair of positive semidefinite $n \times n$ matrices A and B, one has $(\det(A+B))^{1/n} \geq (\det A)^{1/n} + (\det B)^{1/n}$. This means that

$$\sum_m \det(\mathscr{L}_{A,m}^\dagger \mathscr{L}_{A,m})^{1/2d^2} \leq \det\left(\sum_m \mathscr{L}_{A,m}^\dagger \mathscr{L}_{A,m}\right)^{1/2d^2}, \tag{3.41}$$

which together with Eq. (3.40) concludes the proof. □

3.4.1 Properties of the OGC

We have seen that a necessary condition that a measure of correlations must satisfy is to be zero for product states, the latter carrying no correlation whatsoever. However, it is also of interest to know the maximum value that a measure can attain. For what the logarithm of the OSR is concerned the result is trivial, as $\mathbf{O}(\rho_{AB})$ is maximal for any state with maximal OSR, regardless of the presence of entanglement. The OGC shows instead a more compelling behaviour regarding the states for which it reaches its maximum value. We have the following.

Proposition 3.2. *$\mathbf{G}(\rho_{AB})$ is maximal if and only if ρ_{AB} is the maximally entangled state.*

Proof. The proof is divided in two parts. We first show that $\mathbf{G}(\rho_{AB})$ is maximal if and only if the OSC of ρ_{AB} are all equal to $1/d$. In the second part we prove that a state gives rise to such OSC if and only it is a maximally entangled state.

For the first part, recall that $\mathbf{G}(\rho_{AB})$ is non-zero only when OSR is maximal, then $r_i > 0$ for any i. Furthermore, we argue that the maximum of $\mathbf{G}(\rho_{AB})$ is achieved for pure states. To see this, let $\{a_i\}_{i=1}^n$ and $\{b_i\}_{i=1}^n$ be two sets of positive numbers with $\sum_i a_i \leq \sum_i b_i$. The arithmetic–geometric mean inequality [cf. Eq. (3.33)] implies that

$$\prod_{i=1}^n a_i \leq \left(\sum_{i=1}^n \frac{a_i}{n}\right)^n \leq \left(\sum_{i=1}^n \frac{b_i}{n}\right)^n \leq \prod_{i=1}^n b_i. \tag{3.42}$$

In other words, the set of coefficients with the larger sum will return the larger product. This means that we can restrict our maximization problem to pure states, for which $\sum_i r_i^2$ is maximal (see Proposition 2.1). In conclusion, finding the maximum value of $\mathbf{G}(\rho_{AB})$ over pure states translates into maximising a sequence of d^2 positive numbers r_i subject to the constraint $\sum_i r_i^2 = 1$. Using the method of Lagrange multipliers, we set

$$L(r_i, \lambda) = \prod_{i=1}^{d^2} r_i - \lambda\left(\sum_{i=1}^{d^2} r_i^2 - 1\right) \tag{3.43}$$

and solve the system of equations

$$\begin{cases} \dfrac{\partial}{\partial r_j} L(r_i, \lambda) = r_j^{-1} \displaystyle\prod_{i=1}^{d^2} r_i - 2 r_j \lambda = 0 \quad \forall\, j \in \{1, \dots, d^2\} \\ \dfrac{\partial}{\partial \lambda} L(r_i, \lambda) = \displaystyle\sum_{i=1}^{d^2} r_i^2 - 1 = 0, \end{cases} \tag{3.44}$$

which gives $r_i = 1/d$ for all $i \in \{1, \dots, d^2\}$.

For the second part, first observe that from the definition of maximally entangled state it is direct to see that its OSC are all equal to $1/d$. For the reverse implication, we show that if $OSC(\rho_{AB}) = (1/d, \dots, 1/d)$, then ρ_{AB} is pure and has completely mixed marginals, namely ρ_{AB} is maximally entangled. Purity is a consequence of Prop. 2.1, which gives $\mathrm{Tr}(\rho_{AB}^2) = \sum_{i=1}^{d^2} d^{-2} = 1$. To compute the marginals instead, consider orthonormal basis $\{A_i\}, \{B_i\}$ of $\mathcal{L}(\mathcal{H}_A), \mathcal{L}(\mathcal{H}_B)$, respectively, so that $\rho_{AB} = \frac{1}{d} \sum_{i=1}^{d^2} A_i \otimes B_i$. Since ρ_{AB} is pure, $\rho_A = \mathrm{Tr}_B(\rho_{AB}) = \mathrm{Tr}_B(\rho_{AB}^2)$, thus we can obtain the value of $\mathrm{Tr}(\rho_A^2)$ by computing

$$\begin{aligned} \mathrm{Tr}(\rho_A^2) &= \mathrm{Tr}\left[\left(\sum_{i,j=1}^{d^2} \frac{1}{d^2} \mathrm{Tr}_B[(A_i \otimes B_i)^\dagger (A_j \otimes B_j)] \right)^2 \right] \\ &= \mathrm{Tr}\left[\left(\frac{1}{d^2} \sum_{i,j=1}^{d^2} A_i^\dagger A_j \otimes \mathrm{Tr}_B(B_i^\dagger B_j) \right)^2 \right] \\ &= \mathrm{Tr}\left[\left(\frac{1}{d^2} \sum_{i=1}^{d^2} A_i^\dagger A_i \right)^2 \right] \\ &= \frac{1}{d}. \end{aligned} \tag{3.45}$$

It follows that ρ_A is completely mixed (and the same holds for ρ_B). The equality in the last line of Eq. (3.45) is due to $\sum_{i=1}^{n^2} A_i^\dagger A_i = d\mathbb{1}$, which can be proved as follows. Let $A_i = |k\rangle\langle l|$, where we are identifying the index i with the couple (k, l), with $i = 1, \dots, n^2$

and $k, l = 1, \ldots, n$. Then one has

$$\sum_{i=1}^{d^2} A_i^\dagger A_i = \sum_{k,l=1}^{d} |l\rangle\langle k|k\rangle\langle l| = d\,\mathbb{1}. \tag{3.46}$$

To show that Eq. (3.46) is independent of the chosen basis, let $F_j = \sum_i u_{ji} A_i$ be a new orthonormal basis, where $u_{ji} = \mathrm{Tr}(A_i^\dagger F_j)$. The result $\sum_{i=1}^{d^2} A_i^\dagger A_i = \sum_{j=1}^{d^2} F_j^\dagger F_j$ follows from the unitarity of U. □

Here we focus on the possibility of detecting entanglement and steering in bipartite quantum states by looking at their OGC. The following result is consequence of the realignment criterion for separability given in Theorem 2.4.

Theorem 3.6. *A bipartite state $\rho_{AB} \in \mathcal{S}_{AB}$ is entangled if $\mathbf{G}(\rho_{AB}) > d^{-2}$, and it is steerable if $\mathbf{G}(\rho_{AB}) > d^{-3/2}$, where $d = \min\{d_A, d_B\}$.*

Proof. To prove the first claim we show that $\mathbf{G}(\rho_{AB})$ is upper bounded by d^{-2} on the set of separable states. To do so, it is enough to reformulate the maximization problem in Eqs. (3.43) and (3.44) of Proposition 3.2 with the new constraint $\sum_{i=1}^{d^2} r_i \leq 1$, that is nothing but the RC (see Theorem 2.4). As discussed in Proposition 3.2, we can restrict the maximization problems to states such that $\sum_{i=1}^{d^2} r_i$ is exactly one, cf. Eq. (3.42)

The new Lagrangian reads

$$L(r_i, \lambda) = \prod_{i=1}^{d^2} r_i - \lambda\left(\sum_{i=1}^{d^2} r_i - 1\right). \tag{3.47}$$

Solving the system of equations $\frac{\partial}{\partial r_j} L(r_i, \lambda) = \frac{\partial}{\partial \lambda} L(r_i, \lambda) = 0$ gives $r_i = d^{-2}$ for any $i = 1, \ldots, d^2$, implying that

$$\sup\{\mathbf{G}(\rho) \mid \rho \text{ is separable}\} \leq d^{-2}. \tag{3.48}$$

The claim about steerable states is proven in the same way, with the exception of the constraint, which is dictated by Theorem 2.5 and takes on the form $\sum_{i=1}^{d^2} r_i \leq \sqrt{d}$.

The optimization here yields

$$\sup\{\mathbf{G}(\rho) \mid \rho \text{ is unsteerable}\} \leq d^{-3/2}. \tag{3.49}$$

□

Remark. *The results in Theorem 3.6 depend on the general fact that, given a sequence of positive numbers with finite sum, their product is maximized when they are all equal. However, there is no ρ which saturates the bounds of Eqs.* (3.48) *and* (3.49). *Indeed, the maximally entangled state $|\psi^+\rangle\langle\psi^+|$ is the only state having maximal* OSR *and such that the* OSC*s are all equal to each other. The proof of this fact is by contradiction. Assume there exists a density matrix σ_{AB} with* $\mathrm{OSC}(\sigma_{AB}) = (1/c, \ldots, 1/c)$ *and such that $c \neq d$, i.e. there exist local orthonormal bases $\{A_i\}$ and $\{B_i\}$ such that*

$$\sigma_{AB} = \frac{1}{c}\sum_i A_i \otimes B_i, \quad \sigma_{AB} \neq |\psi^+\rangle\langle\psi^+|. \tag{3.50}$$

On the one hand, if we assume $c < d$, Proposition 2.1 would imply $\mathrm{Tr}(\sigma_{AB}^2) = d^2c^{-2} > 1$. *On the other hand, $c > d$ returns* $\mathrm{Tr}(\sigma_A) = \mathrm{Tr}(\sigma_B) = dc^{-2} < 1/d$. *To see this last point, first notice that σ_{AB} must be pure in order to be a possible solution to the maximization problems of Theorem 3.6 (as argued before). This means $\sigma_{AB} = \sigma_{AB}^2$, which implies*

$$\begin{aligned}
\mathrm{Tr}(\sigma_A) &= \mathrm{Tr}\left[\mathrm{Tr}_B(\sigma_{AB}^2)\right] \\
&= \frac{1}{c^2}\,\mathrm{Tr}\left[\sum_{ij} A_i^\dagger A_j \otimes \mathrm{Tr}(B_i^\dagger B_j)\right] \\
&= \frac{d}{c^2},
\end{aligned} \tag{3.51}$$

where we have used $\sum_i \mathrm{Tr}(A_i^\dagger A_i) = d$*, see the discussion after Eq.* (3.45). *Soth assumptions $c < d$ and $c > d$ give absurd conditions, in that they violate the bounds on the purity of a quantum state and the unit trace condition, respectively. Hence, we conclude that the only open possibility is $c = d$, which is actually realized by any maximally* *entangled state.*

As anticipated, an issue with the OGC is that it vanishes for any state with OSR less than maximal, even if highly correlated. This means that the OGC is a measure of correlations, i.e. it is not faithful. Loosely speaking, we could argue that the OGC qualifies as a measure of total correlations only for state with maximal OSR. Although the set of OSR-deficient states has zero measure, from a theoretical perspective it is advisable to look for a generalization of the OGC which would measure the amount of total correlations for every possible state in the Hilbert space. One is tempted to consider the other ESP of Eqs. (2.23), as they don't vanish like the OGC as soon as a state is OSR–deficient. In particular, if a quantum state ρ_{AB} has $\mathrm{OSR}(\rho_{AB}) = K$, then all the $l^{th}-$order ESP with $l \leq K$ are nonzero. Moreover, the ESP are Schur-convex [107], meaning that the monotonicity of the single coefficients would imply the monotonicity of all the ESP. However, the monotonicity of the individual OSC under LO does not hold in general. To see this point it is enough to consider two product state ρ_{AB} and σ_{AB} with different purities, i.e. such that $r_1 \neq s_i$, where r_1 and s_1 are the single OSC of ρ_{AB} and σ_{AB}, respectively (see Proposition 2.1). Then, since we can always map ρ_{AB} into σ_{AB} and vice versa by means of an opportune LO, the first OSC – hence the $1^{st}-$order ESP – is not monotonic and cannot be used to define any measure of total correlations. We were able to find a counterexample only to the monotonicity of the $1^{st}-$order ESP. The possibility of considering the other ones remains open.

In Sec. 3.5 we introduce two new measures which do not suffer of the drawbacks of $\mathbf{O}(\rho)$ and $\mathbf{G}(\rho)$, i.e. they depend smoothly on the OSC and do not vanish for OSR-rank deficient states.

3.4.2 Majorization and local operations

One of the central problems of QM is to find conditions under which an entangled state can be transformed into another one by local operations and classical communication (see Sec. 1.2). In particular, identifying the set of quantum states that can be prepared through LOCC starting form a given state ultimately translates to the question of what tasks can be accomplished using a given physical resource [108]. In this section we introduce a result showing how the theory of entanglement transformation is closely related to the algebraic concept of *majorization*, which is a preorder relations on vectors of real numbers [109–111] and it is defined as follows.

Definition 3.6 (Majorization of vectors [112]). Let us consider two vectors of real numbers $\mathbf{p} = (p_q, p_2, \ldots, p_n)$ and $\mathbf{q} = (q_1, q_2, \ldots, q_n)$. We say that $\mathbf{a}$ is *weakly majorized* by $\mathbf{b}$ (equivalently, $\mathbf{b}$ *weakly majorizes* $\mathbf{a}$), written as $\mathbf{a} \prec_w \mathbf{b}$, if and only if

$$\sum_{i=1}^{k} a_i^{\downarrow} \leq \sum_{i=1}^{k} b_i^{\downarrow} \quad \text{for any } k = 1, 2, \ldots, n, \tag{3.52}$$

where $\downarrow$ indicates that the components of the vectors are taken in decreasing order. Moreover, if the additional condition $\sum_{i=1}^{d} a_i = \sum_{i=1}^{d} b_i$ holds, we say that $\mathbf{a}$ is *majorized* by $\mathbf{b}$ (equivalently, $\mathbf{b}$ *majorizes* $\mathbf{a}$), written as $\mathbf{a} \prec \mathbf{b}$.[5]

The connection between majorization and entanglement was given in [108],[6] where the following theorem for pure states was derived:

Theorem 3.7 (Majorization and entanglement [108]). *Let us consider two bipartite vector $|\psi\rangle, |\varphi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. Let λ_ϕ and λ_φ be the vectors of eigenvalues of $\mathrm{Tr}_B |\psi\rangle\langle\psi|$ and $\mathrm{Tr}_B |\varphi\rangle\langle\varphi|$, respectively. Then, $|\psi\rangle$ can be transformed to $|\varphi\rangle$ using local operations and classical communication if and only if λ_ψ is majorized by λ_φ, in formulae*

$$|\psi\rangle \xrightarrow[\text{LOCC}]{} |\varphi\rangle \quad \Leftrightarrow \quad \lambda_\psi \prec \lambda_\varphi. \tag{3.53}$$

The theorem above certainly is of major importance in QM, as it gives necessary and sufficient conditions for entanglement transformation to be possible. It would be interesting to see if such characterisation is attainable also in our context, namely we wonder if the vector of the OSC of states transformed under LO – rather than

[5] Notice that majorization is not a partial order since $\mathbf{a} \prec \mathbf{b}$ and $\mathbf{b} \prec \mathbf{a}$ do not imply that the two vectors are equal. We can only imply that $\mathbf{a}$ and $\mathbf{b}$ have the same components, but not necessarily in the same order.

[6] The reader is referred to [113] (and reference therein) for an extensive review of the various applications of majorization in QM.

LOCC – satisfies similar conditions to Theorem 3.7. In particular, we are looking for majorization relations between certain normalized versions of the vector of the OSC. If we denote by $\{r_i\}$ and $\{p_i\}$ the OSC of an arbitrary–dimensional bipartite state ρ_{AB} and of its square root $\sqrt{\rho_{AB}}$, respectively, the normalized vectors we are considering are given by

$$\mathbf{v_1}(\rho_{AB}) = \left(\frac{r_1}{\sum_i r_i}, \frac{r_2}{\sum_i r_i}, \ldots, \frac{r_{d^2}}{\sum_i r_i}\right) \tag{3.54}$$

$$\mathbf{v_2}(\rho_{AB}) = \left(\frac{r_1^2}{\sum_i r_i^2}, \frac{r_2^2}{\sum_i r_i^2}, \ldots, \frac{r_{d^2}}{\sum_i r_i^2}\right) \tag{3.55}$$

$$\mathbf{v_3}(\rho_{AB}) = (p_1, p_2, \ldots, p_{d^2}), \tag{3.56}$$

where the usual convention about the dimensions of the subsystems is in force.

We can finally formulate our question rigorously. Given two quantum states related by LO, i.e. $\sigma_{AB} = (\Lambda_A \otimes \Lambda_B)[\rho_{AB}]$, is it true that $\mathbf{v_i}(\rho_{AB})$ is majorized by $\mathbf{v_i}(\sigma_{AB})$, for some $i = 1, 2, 3$?

Here we show that for $i = 1, 2$ we can find two quantum states connected by LO but such that no majorization relation between the corresponding vectors $\mathbf{v_i}$ is in place. To see this, we reason as follows. First, for simplicity let us consider the action of one local map only, e.g $\sigma_{AB} = (\Lambda_A \otimes \mathrm{id}_B)[\rho_{AB}]$, since the extension to two–sided operations follows easily from their independence. Thanks to Stinespring dilation theorem 1.4, the action of any channel Λ_A can be characterized by introducing an auxiliary Hilbert space $\mathcal{H}_E$ and a particular protocol. In our context such protocol reads as follows. We say that for any channel $\Lambda_A \in \mathcal{C}(\mathcal{H}, \mathcal{H}')$ there exists a Hilbert space $\mathcal{H}_E$ and a unitary operation U on $\mathcal{L}(\mathcal{H} \otimes \mathcal{H}_E)$ such that $\sigma_{AB} = (\Lambda_A \otimes \mathrm{id}_B)[\rho_{AB}] = \mathrm{Tr}_E[U(\rho_{AB} \otimes |0\rangle\langle 0|_E)U^\dagger]$, for all $\rho_{AB} \in \mathcal{S}_{AB}$. In other words, the action of any channel can be simulated through the operations of tensorization, unitary evolution and partial trace. Now notice that since both the tensorization with the pure state in $\mathcal{H}_E$ and the unitary operation are performed locally, the OSC of ρ_{AB} are left unchanged by them. Then, if the OSC of σ_{AB} are different from the OSC of ρ_{AB}, it can only depend upon the application of the partial trace. In turn, looking for a majorization relation between the OSC (or between the vectors $\mathbf{v_i}$) of two quantum states related by LO translates into looking for a majorization relation between the OSC (or between the vectors $\mathbf{v_i}$) of two quantum states related by the partial trace. We are now in a position to introduce the sought example. Consider the following state vector

$$|\psi\rangle = \frac{1}{\sqrt{3}} \left(|001\rangle + |110\rangle + |111\rangle \right), \tag{3.57}$$

and write $\rho_{ABC} = |\psi\rangle\langle\psi|$ and $\rho_{AB} = \mathrm{Tr}_C(\rho_{ABC})$. Moreover, consider the OSD of ρ_{ABC} with respect to the bipartition $A : BC$ (and, of course, decompose ρ_{AB} with respect to the bipartition $A : B$). Then, one easily verifies that

$$\mathbf{v_i}(\rho_{ABC}) \nprec \mathbf{v_i}(\rho_{AB}), \quad \text{for } \mathbf{i} = 1, 2. \tag{3.58}$$

We can conclude that the fact that two states are related by LO does not imply that the respective vectors defined by Eqs. (3.55)–(3.56) satisfy the supposed majorization relation. On the other hand, we have run several numerical tests which corroborate the plausibility of the majorization assumption for $i = 3$, i.e. no counterexample has been found yet to the conjecture

$$\rho \underset{LO}{\longrightarrow} \sigma \quad \Rightarrow \quad \mathbf{v_1}(\rho) \prec \mathbf{v_1}(\rho). \tag{3.59}$$

The absence of counterexample compels us to look for an analytic way to prove the conjecture above. Moreover, other kinds of vectors based on the OSC can be taken into consideration. These possibilities remain open and a represent an interesting starting point for future research.

3.5 A fidelity–based measure

The measure of total correlation introduced in this section is based on a measure of proximity between quantum states. The latter is defined as follows.

Definition 3.7. (Holevo's Fidelity [114, 115]) Given any two quantum states ρ and σ, the function

$$F_H(\rho,\sigma) := [\mathrm{Tr}(\sqrt{\rho}\sqrt{\sigma})]^2 \tag{3.60}$$

is known as Holevo's fidelity, also called Holevo"just-as-good fidelity".

The function above was first introduced in [114] (without the square), where it was proved that $\sqrt{F_H}$ is a measure of proximity for quantum states just as good as the trace distance (from which the name), since

$$1-\sqrt{F_H(\rho,\sigma)} \leq \frac{1}{2}\|\rho-\sigma\|_1 \leq \sqrt{1-F_H(\rho,\sigma)}. \tag{3.61}$$

Proposition 3.3. *F_H satisfies the following properties:*

i) $0 \leq F_H(\rho,\sigma) \leq 1$*;*

ii) $F_H(\rho,\sigma) = 1$ *iff* $\rho = \sigma$*;*

iii) $F_H(\rho,\sigma) = 0$ *iff* $\rho \perp \sigma$*;*

iv) *F_H is jointly concave. Using Uhlmann [116] and Stinespring theorems [59] (see Theorem 1.4), one shows that the joint concavity of F_H is equivalent to its monotonicity under quantum operation, i.e.*

$$F_H(\rho,\sigma) \leq F_H(\Lambda[\rho]\Lambda[\sigma]), \quad \textit{for any channel } \Lambda. \tag{3.62}$$

By using the Holevo's fidelity and taking advantage of its properties, we define a new measure of total correlations as a function of the proximity between a quantum states and the state given by the products of its marginals.

Definition 3.8. For any bipartite state ρ_{AB} with marginals ρ_A and ρ_B, we define the function

$$\mathbf{F}(\rho_{AB}) := -\log[F_H(\rho_{AB},\rho_A\otimes\rho_B)]. \tag{3.63}$$

We argue that $\mathbf{F}$ is a meaningful measure of total correlations.[7]

[7] The quantity $\mathbf{F}$ is equal to a particular instance of the quantum Rényi relative entropy. This fact is proven in Sec. 3.7.

Theorem 3.8. $\mathbf{F}(\rho_{AB})$ *attains its minimum for product states, in particular* $\mathbf{F}(\rho_{AB}) = 0$ *if and only if* ρ_{AB} *is a product state. Moreover,* $\mathbf{F}(\rho_{AB})$ *does not increase under local operations, i.e.* $\mathbf{F}(\rho_{AB}) \geq \mathbf{F}(\Lambda_A \otimes \Lambda_B[\rho_{AB}])$*.*

Proof. To see the first claim, notice that $\mathbf{F}(\rho_{AB}) = 0$ if and only if $F_H(\rho_{AB}, \rho_A \otimes \rho_B) = 1$, which happens if and only if ρ_{AB} is the product of its own marginals (cf. properties (ii) of Proposition 3.3).

The see the second claim observe that

$$\begin{aligned} \mathbf{F}(\Lambda_A \otimes \Lambda_B[\rho_{AB}]) &= -\log[F_H(\Lambda_A \otimes \Lambda_B[\rho_{AB}], \Lambda_A[\rho_A] \otimes \Lambda_B[\rho_B])] \\ &\geq -\log[F_H(\rho_{AB}, \rho_A \otimes \rho_B)] \\ &= \mathbf{F}([\rho_{AB}]), \end{aligned} \tag{3.64}$$

where the inequality is a consequence of properties (iv) of Proposition 3.3 together with the monotonicity of the (minus) logarithm function. □

3.5.1 Properties of F

Here we list some facts about $\mathbf{F}(\rho_{AB})$ and its application to entanglement detection.

Proposition 3.4. *Let us consider the state space* $\mathcal{S}_{AB}$ *and denote* $d = \min\{d_A, d_B\}$*.* $\mathbf{F}(\rho_{AB})$ *satisfies the following properties:*

i) for any quantum state $\rho_{AB} \in \mathcal{S}_{AB}$*,* $\mathbf{F}(\rho_{AB}) \leq \log(d)$*;*

ii) $\mathbf{F}(\rho_{AB})$ *attains its maximum for the maximally entangled state* $\rho_{AB} = |\psi^+\rangle\langle\psi^+|$*, i.e.* $\mathbf{F}(|\psi^+\rangle\langle\psi^+|) = \log(d)$*.*

Proof. $i)$ Given a quantum state ρ_{AB}, let us consider the Schmidt decomposition of the purification ρ_{ABC}, which is given by

$$|\psi\rangle_{ABC} = \sum_i \sqrt{p_i}\, |a_i\rangle_A\, |b_i\rangle_{BC}\,. \tag{3.65}$$

Notice that the marginals ρ_A and ρ_{BC} are diagonal in the respective local bases, then we can easily consider their square roots (as for pure states). By using the monotonicity of F_H one has

$$\begin{aligned}
F_H\left(\rho_{AB},\rho_A\otimes\rho_B\right) &\geq F_H\left(|\psi\rangle\langle\psi|_{ABC},\rho_A\otimes\rho_{BC}\right) \\
&= F_H\Bigg(\sum_{kl}\sqrt{p_kp_l}\,|a_k\rangle\langle a_l|_A\otimes|a_k\rangle\langle a_l|_{BC}, \\
&\quad \sum_i p_i\,|a_i\rangle\langle a_i|_A\otimes\sum_j p_j\,|b_j\rangle\langle b_j|_{BC}\Bigg) \\
&= \sum_{ijkl}\sqrt{p_ip_jp_kp_l}\,\mathrm{Tr}\left(|a_k\rangle\langle a_l|_A\otimes|b_k\rangle\langle b_l|_{BC}\,|a_i\rangle\langle a_i|_A\otimes|b_j\rangle\langle b_j|_{BC}\right) \\
&= \sum_i p_i^2 \\
&= \mathrm{Tr}\left(\rho_A^2\right). \qquad (3.66)
\end{aligned}$$

On the other hand, if we were to purify ρ_{AB} by attaching the ancilla on the subsystem A, we would get $F_H\left(\rho_{AB},\rho_A\otimes\rho_B\right)\geq \mathrm{Tr}(\rho_B^2)$. Then, in general we can conclude that

$$F_H\left(\rho_{AB},\rho_A\otimes\rho_B\right)\geq \max\left\{\mathrm{Tr}\left(\rho_A^2\right),\mathrm{Tr}\left(\rho_B^2\right)\right\}=\frac{1}{d}. \qquad (3.67)$$

ii) By definition of $\mathbf{F}(\rho_{AB})$, its maximum correspond to the minimum of the Holevo fidelity between ρ_{AB} and the product of the reduced states which, because of the joint concavity of F_H, is attained for pure states. By following similar steps as in Eq. (3.66) one finds that for a pure state $|\psi\rangle\langle\psi|_{AB}\in\mathcal{S}_{AB}$ with marginals ρ_A and ρ_B it holds that

$$\begin{aligned}
F_H\left(|\psi\rangle\langle\psi|_{AB},\rho_A\otimes\rho_B\right) &= F_H\Bigg(\sum_{kl}\sqrt{p_kp_l}\,|a_k\rangle\langle a_l|_A\otimes|a_k\rangle\langle a_l|_B, \\
&\quad \sum_i p_i\,|a_i\rangle\langle a_i|_A\otimes\sum_j p_j\,|b_j\rangle\langle b_j|_B\Bigg) \\
&= \sum_{ijkl}\sqrt{p_ip_jp_kp_l}\,\mathrm{Tr}\left(|a_k\rangle\langle a_l|_A\otimes|b_k\rangle\langle b_l|_B\,|a_i\rangle\langle a_i|_A\otimes|b_j\rangle\langle b_j|_B\right) \\
&= \sum_i p_i^2 \\
&= \mathrm{Tr}\left(\rho_{A(B)}^2\right). \qquad (3.68)
\end{aligned}$$

The theorem follows by recalling that the purity of the reduced density matrices is minimal for the maximally entangled state, which has completely mixed marginals □

To conclude, here we find an upper bound for $\mathbf{F}$ on the set of separable states. In turn, this can be used for entanglement detection.

Proposition 3.5. *For any separable state $\rho_{AB} \in \mathcal{S}_{AB}$, one has $\mathbf{F}(\rho_{AB}) \leq \log(\sqrt{d})$.*

Proof. Consider a generic separable state $\rho_{AB} = \sum_i p_i |a_i\rangle\langle a_i| \otimes |b_i\rangle\langle b_i|$. As before, we study the Holevo's fidelity between ρ_{AB} and the product of its reduced density matrices.

$$\begin{aligned} F_H(\rho_{AB}, \rho_A \otimes \rho_B) &= F_H\Bigg(\sum_i p_i |\alpha_i\rangle\langle\alpha_i|_A \otimes |\beta_i\rangle\langle\beta_i|_B\,, \\ &\qquad \sum_i p_i |\alpha_i\rangle\langle\alpha_i|_A \otimes \sum_j p_j |\beta_j\rangle\langle\beta_j|_B\Bigg) \qquad (3.69) \\ &\geq \sum_i p_i F_H(|\alpha_i\rangle\langle\alpha_i|_A \otimes |\beta_i\rangle\langle\beta_i|_B\,, |\alpha_i\rangle\langle\alpha_i|_A \otimes \rho_B) \\ &= \sum_i p_i F_H(|\beta_i\rangle\langle\beta_i|_B\,, \rho_B) \\ &= \mathrm{Tr}\left(\sum_i p_i |\beta_i\rangle\langle\beta_i|_B \sqrt{\rho_B}\right) \\ &= \mathrm{Tr}\left(\rho_B^{3/2}\right) \\ &= \sum_k \lambda_k^{3/2} \\ &\geq d\left(\frac{1}{d}\right)^{3/2} \\ &= \frac{1}{\sqrt{d}}. \end{aligned} \qquad (3.70)$$

The first inequality above comes from the joint concavity of the Holevo's fidelity, while the λ_i's of the second to last line are the eigenvalues of ρ_B, so $\sum_k \lambda_k = 1$. The equality in the third line is justified by the fact that adding or removing a fixed state (in this case removing $|\alpha\rangle\langle\alpha|_A$) form the argument of F_H does not change its value. Finally, the last inequality follow from the fact that a uniform probability distribution is majorized by any other, and the sum of power 3/2 is Schur convex. To conclude, we apply the definition of $\mathbf{F}$ and get the claim. □

3.6 Decomposing the square root of a quantum state

The measure $\mathbf{F}(\rho_{AB})$ introduced in the last section employs the square root of ρ_{AB}, as well as of its marginals, via the application of the Holevo's fidelity. Here we take a closer look at the square root of a quantum state by making use of its OSD. We write

$$\sqrt{\rho_{AB}} = \sum_{i=1}^{d^2} \sqrt{p_i}\, A_i \otimes B_i, \tag{3.71}$$

where $d = \min\{d_A, d_B\}$, $\sum_{i=1}^{d^2} p_i = 1$ and $\mathrm{Tr}(A_i^\dagger A_j) = \mathrm{Tr}(B_i^\dagger B_j) = \delta_{ij}$ and the coefficients are sorted in decreasing order. Also, we consider the square root of the reduced density matrices $\sqrt{\rho_A}, \sqrt{\rho_B}$ and expand them over the bases $\{A_i\}, \{B_i\}$, respectively. We obtain

$$\sqrt{\rho_A} = \sum_{i=1}^{d^2} \mathrm{Tr}(\sqrt{\rho_A} A_i) A_i \tag{3.72}$$

$$\sqrt{\rho_B} = \sum_{i=1}^{d^2} \mathrm{Tr}(\sqrt{\rho_B} B_i) B_i. \tag{3.73}$$

Notice that the OSC in Eq. (3.71) are denoted with p_i instead of r_i to highlight the fact that they form a probability distribution. Moreover, the OSC of ρ_{AB} and $\sqrt{\rho_{AB}}$ are generally different, and the relationship between the two sets of coefficients is not entirely clear.

The interesting feature of the decomposition (3.71) is that the largest OSC $\sqrt{p_{\max}}$ of $\sqrt{\rho_{AB}}$ – which can be expressed as the uniform norm of the correlation matrix $\|\mathcal{C}(\sqrt{\rho_{AB}})\|_\infty$ – is monotonic, as we are going to see in Theorem 3.9. This allows for the definition of an interesting measure of total correlations.

Definition 3.9. For any bipartite state ρ_{AB}, we define the function

$$\mathbf{C}(\rho_{AB}) := -\log \|\mathcal{C}(\sqrt{\rho_{AB}})\|_\infty^2. \tag{3.74}$$

$\mathbf{C}(\rho_{AB})$ satisfies the desired requirements:

Theorem 3.9. $\mathbf{C}(\rho_{AB})$ *attains its minimum for product states, in particular* $\mathbf{C}(\rho_{AB}) = 0$ *if and only if* ρ_{AB} *is a product state. Moreover,* $\mathbf{C}(\rho_{AB})$ *does not increase under local operations, i.e.* $\mathbf{C}(\rho_{AB}) \geq \mathbf{C}(\Lambda_A \otimes \Lambda_B[\rho_{AB}])$.

Proof. Let us start by observing observing that the largest OSC $\sqrt{p_1}$ of $\sqrt{\rho_{AB}}$ is equivalent to $\mathrm{Tr}(A_1^\dagger \otimes B_1^\dagger \sqrt{\rho_{AB}})$, according to Eq. (3.71). Moreover, since $\|A_1\|_2^2 = \|B_1\|_2^2 = 1$, we can characterize $\|\mathcal{C}(\rho_{AB})\|_\infty \equiv \sqrt{p_1}$ as

$$\|\mathcal{C}(\sqrt{\rho_{AB}})\|_\infty = \max_{\substack{M,N \\ \|M\|_2^2 = \|N\|_2^2 = 1}} \left| \mathrm{Tr}(M \otimes N \sqrt{\rho_{AB}}) \right|. \tag{3.75}$$

We already know that A_i and B_i, thus M and N, can be chosen to be Hermitian. This implies

$$\begin{aligned} \mathrm{Tr}(M^2) &= \mathrm{Tr}[(M^+ - M^-)^2] \\ &= [\mathrm{Tr}(M^+)^2] + [\mathrm{Tr}(M^-)^2] \\ &= \mathrm{Tr}(|M|^2). \end{aligned} \tag{3.76}$$

In other words, M and N can be chosen to be square roots of quantum states (they are positive semidefinite trace one operators), say $M = \sqrt{\tilde{\sigma}_A}, N = \sqrt{\tilde{\sigma}_B}$. Then, using Eq. (3.75) and invoking the monotonicity of F_H, one gets

$$\begin{aligned} \|\mathcal{C}(\sqrt{\rho_{AB}})\|_\infty &= \mathrm{Tr}\left(\sqrt{\tilde{\sigma}_A} \otimes \sqrt{\tilde{\sigma}_B} \sqrt{\rho_{AB}}\right) \\ &\leq \mathrm{Tr}\left(\sqrt{\Lambda_A[\tilde{\sigma}_A]} \otimes \sqrt{\Lambda_B[\tilde{\sigma}_B]} \sqrt{\Lambda_A \otimes \Lambda_B[\rho_{AB}]}\right) \\ &\leq \max_{\substack{M',N' \\ \|M'\|_2^2 = \|N'\|_2^2 = 1}} \left| \mathrm{Tr}\left(M' \otimes N' \sqrt{\Lambda_A \otimes \Lambda_B[\rho_{AB}]}\right) \right| \\ &= \|\mathcal{C}\left(\sqrt{\Lambda_A \otimes \Lambda_B[\rho_{AB}]}\right)\|_\infty. \end{aligned} \tag{3.77}$$

□

3.7 Comparison between measures of total correlations

In this section we look for a relation between the measure of total correlations introduced in the previous sections, i.e. **F** and **C**, and the quantum mutual information. Before that, here we prove that the fidelity–based measure considered in Sec. 3.5 corresponds to a particular instance of the quantum Rényi relative entropy [117]. To see this, we recall that the latter is defined as

$$D_\alpha(\rho||\sigma) := \frac{1}{\alpha-1}\log \mathrm{Tr}\left(\rho^\alpha \sigma^{1-\alpha}\right), \tag{3.78}$$

for $\alpha \in (0,1)\cup(1,\infty)$. Among the several properties satisfied by the quantum Rényi relative entropy, it is convenient to recall that $\alpha \mapsto D_\alpha(\rho||\sigma)$ is a monotonically increasing function of α and, in particular, for $\alpha \to 1$ one recovers the quantum relative entropy defined in Sec. 3.1.

In order to compare the quantum mutual information with our fidelity–based measure of total correlations introduced in Sec. 3.5, we first notice that the latter is nothing but the quantum Rényi relative entropy $D_\alpha(\rho||\sigma)$ for $\alpha = 1/2$:

$$\begin{aligned} D_{1/2}(\rho||\sigma) &= \frac{1}{1/2-1}\log \mathrm{Tr}(\rho^{1/2}\sigma^{1/2}) \\ &= -2\log \mathrm{Tr}(\sqrt{\rho}\sqrt{\sigma}) \\ &= -\log \mathrm{Tr}(\sqrt{\rho}\sqrt{\sigma})^2 \\ &= -\log F_H(\rho,\sigma). \end{aligned} \tag{3.79}$$

Then, thanks to the monotonicity of $D_\alpha(\rho||\sigma)$ with respect to α, we are able to connect our fidelity–based measure $\mathbf{F}(\rho_{AB})$ to the quantum mutual information:

$$\begin{aligned} \mathbf{F}(\rho_{AB}) &:= D_{1/2}(\rho_{AB}||\rho_A\otimes\rho_B) \\ &\leq \lim_{\alpha\to 1} D_\alpha(\rho_{AB}||\rho_A\otimes\rho_B) \\ &= S(\rho_{AB}||\rho_A\otimes\rho_B) \\ &= \mathbf{I}(\rho_{AB}) \end{aligned} \tag{3.80}$$

For what the comparison between $\mathbf{C}$ and $\mathbf{F}$ is concerned instead, notice that

$$\begin{aligned}\mathbf{C}(\rho_{AB}) &:= -\log \|\mathcal{C}(\rho_{AB})\|_\infty^2 \\ &= \min_{\sigma_A,\sigma_B} D_{1/2}(\rho_{AB}||\sigma_A \otimes \sigma_B) \\ &\leq D_{1/2}(\rho_{AB}||\rho_A \otimes \rho_B) \\ &= \mathbf{F}(\rho_{AB}).\end{aligned} \tag{3.81}$$

So, from Eqs. (3.80) and (3.81), we conclude that

$$\mathbf{C} \leq \mathbf{F} \leq \mathbf{I}. \tag{3.82}$$

The relational expression above is the final result of this chapter, and it validates one more time the hypothesis that the OSD can offer useful tools in the study of total correlations. It would be interesting to check if $\mathbf{C}$ and $\mathbf{F}$ are lower bounded by some function of the QMI as well. In fact, this would imply that our measures are, in a sense, equivalent to the QMI. We lastly notice that even if $\mathbf{C}$ and $\mathbf{F}$ might not be easier to compute than the QMI, they surely enrich our knowledge about the OSD, its role in the quantification of total correlations (which is the original motivation of this chapter) and its relation to the most relevant measure of TC in Quantum Information theory.

3.8 Conclusion

By taking advantage of the operator Schmidt decomposition of a bipartite operator and of its square root, we have introduced in this chapter several measures of correlations, some of which have been shown to be faithful in the sense of Definition 3.1. Our results validate the initial intuition that, in a similar manner to what happens at the pure states level, the tensor product expression of bipartite density matrices carries valuable information on the amount of correlation exhibited by quantum states. Furthermore, we have also seen how such measures can be used to detect entanglement and discord. Besides that, it is also remarkable that – thanks to the result of the last section –

the measures of total correlations from Sec. 3.5 and 3.6 are in a sense equivalent to the quantum mutual information. In order to give meaning to this statement, it looks worthwhile to investigate further the properties of **F** and **C**, trying to learn to what extent they can be employed alternatively to the QMI. However, other questions are still open. In particular, it would be interesting to understand the behaviour of the l^{th}−order elementary symmetric polynomial in the OSC under the action of local maps for $1 < l < d^2$, as well as to check the conjectured average monotonicity of the Operator G-concurrence studied in Sec. 3.4. These issues are still under consideration, their clarification being beneficial to the completion of [3].

This concludes the first part of our research project. In what follows, we change our perspective and show how the degree of correlations possessed by a quantum state, estimated through the OSD and the theory presented so far, can be exploited to allow, and in certain cases enhance, some specific protocols actualizing the tasks of channel discrimination and tomography.

Chapter 4

Channel discrimination

Channel discrimination is a fundamental task that falls under the umbrella of quantum metrology [118, 119] and consists in the attempt to tell apart two or more *known* channels; think of the situation where we want to probe the presence or absence of a magnetic field. In the prototypical and simplest case, one of two channels is applied once to a probe, and we try to identify the channel by performing a measurement on the output probe. Channel discrimination is typically performed by tailoring the state of the input probe to the channels to be discriminated. The wrong choice of input might make the probability of correct identification less than optimal or even not better than a random guess.

There can be advantages in channel discrimination by making use of correlations between the probe and a reference ancilla. One possible advantage is that correlations may lead to a probability of success in the discrimination that is higher than what possible without the use of an ancillary system [18, 62, 118, 120–129]. In general, achieving such a higher probability of success requires (i) to tailor the probe–ancilla input state to the specific channels to be discriminated and (ii) input entanglement between probe and ancilla. Another advantage provided by probe–ancilla correlations, on which we focus in this thesis, is that they may allow to discriminate between an arbitrary pair of known channels, without the need to tailor the input probe–ancilla state to avoid 'being blind' to the difference between the channels. This fact is at the basis of the celebrated Choi-Jamiołkowski isomorphism [52, 53] between linear maps and linear operators presented in Sec. 1.3.2. . The use of an ancilla allows one to perform channel tomography – that is, to identify an unknown channel with many uses of the unknown channel – with a fixed input state [15]. Such a feat can be achieved even in the absence of entanglement, and Ref. [15] already identified the operator Schmidt rank (OSR) of the probe–ancilla input state as the key property determining whether such state makes ancilla-assisted tomography possible. An equivalent result was independently derived in [130], where the *faithfulnees* of bipartite quantum states was introduced. A bipartite state used in ancilla-assisted channel tomography is faithful if the action of the channel on the probe leads to an output probe–ancilla state that is uniquely associated with the specific channel. Nonetheless, the study of the usefulness of correlations in fixed-input ancilla-assisted channel discrimination and channel tomography has been limited [131]. In this chapter we shed light on ancilla-assisted channel discrimination, providing an analysis of how the operator Schmidt decomposition (OSD) of the probe–ancilla input state affects the quality of the discrimination. In particular, we introduce a worst-case quantifier for the performance of a probe–ancilla state in channel discrimination, the *Channel Discrimination Power* (CDP). We provide general upper and lower bounds to the CDP of a state in terms of the OSD of the state. We compute the exact CDP of pure states. Remarkably, we show that, while correlated but unentangled states can have non-zero CDP, and allow the discrimination of any pair of channels as long as they have maximal OSR, they cannot have maximal CDP. More in general, we provide a non-trivial bound on the channel discrimination power of any state – entangled or unentangled – that passes the realignment criterion for separability (see Sec. 2.2). Furthermore, we prove that quantum discord provides a bound for the channel discrimination power of a bipartite state.

4.1 Channel discrimination power

Channel discrimination is a generalization of state discrimination, where the objects to tell apart are now channels. One can define a physically meaningful notion of distance between two channels Λ_0 and Λ_1 via [128] $D(\Lambda_0, \Lambda_1) := \max_{\rho\in\mathcal{D}(\mathcal{H}_S)} D\left(\Lambda_0[\rho], \Lambda_1[\rho]\right)$, that is by considering the trace distance of the output states of a probe upon acting on the same input state of the probe. One fundamental—and relevant for applications—way in which quantum physics differs from classical physics, is that the distinguishability of two channels, as captured by $D(\Lambda_0, \Lambda_1)$, can be enhanced by the use of entanglement between the input probe and an ancilla [18, 62, 118, 120–129]. One can prove that the best ancilla system can be chosen to be a copy S' of the input probe system S, so that we can introduce the diamond distance between Λ_0 and Λ_1 as $D_\diamond(\Lambda_0, \Lambda_1) := D(\Lambda_{0,S} \otimes \mathrm{id}_{S'}, \Lambda_{1,S} \otimes \mathrm{id}_{S'})$, where id_X indicates the identity map on system X (cf. Sec. 1.4). The diamond distance formalizes the notion of best possible one-shot distinguishability of two quantum channels.

Figure 4.1: Two strategies for distinguishing channels. (a) No ancilla is used: a probe undergoes one of many possible quantum evolutions described by channels $\{\Lambda_\alpha\}$, and is later measured (box M). Many different input states $\{\rho_k\}$ are in general needed to discriminate between arbitrary channels, if one cannot tailor the input to the channels. (b) ancilla–assisted: the probe A is correlated with an ancilla B; the output probe and the ancilla are jointly measured. Depending on the initial probe–ancilla correlations, it might be possible to distinguish between arbitrary evolutions, without modifying the input.

In general, it is not possible to distinguish arbitrary quantum channels in $\mathcal{T}(\mathcal{H}_X, \mathcal{H}_Y)$ by means of their action on an input state $\rho \in \mathscr{D}(\mathcal{H}_X)$ of the probe alone that is independent of the channels considered [1]. Nonetheless, it is always possible to tell two arbitrary channels in $\mathcal{T}(\mathcal{H}_X, \mathcal{H}_Y)$ apart by 'feeding' them with many different input states ρ_k. As long as $\{\rho_k\}$ constitutes a basis for $\mathcal{L}(\mathcal{H}_X)$, and as long as an arbitrary number of uses of the channel are allowed, one can even perform a tomographic reconstruction of a channel Λ [see Figure 4.1a] [13].

[1]For example, consider the case where Λ_0 is the identity channel, so that $\Lambda_0[\sigma] = \sigma$ for all σ, and Λ_1 is the channel with fixed output ρ. Then, obviously, $D(\Lambda_0[\rho], \Lambda_1[\rho]) = 0$, even if the two channels are very different, and even having many copies of $\Lambda_i[\rho]$ we cannot tell the two channels apart.

Remarkably, it is possible to perform tomography of the channel, or the non-trivial discrimination of an arbitrary number of channels, even with just a fixed input state, as long as one uses an ancilla: this constitutes the framework of ancilla-assisted channel discrimination and channel tomography [see Figure 4.1b]. Ref. [15] proves both theoretically and experimentally that channel tomography is possible also when the state ρ_{AB} of probe A and ancilla B is separable. The key condition that permits channel tomography on A with ρ_{AB} is that $\mathrm{OSR}(\rho_{AB}) = d_A^2$. Indeed, one has $\Lambda_A[\rho_{AB}] = \sum_{i=1}^{\mathrm{OSR}(\rho)} r_i \Lambda[A_i] \otimes B_i$, and, as long as the state has $\mathrm{OSR}(\rho) = d_A^2$, one can reconstruct the action of the map Λ on an arbitrary state $\sigma \in D(\mathcal{H}_A)$ as $\Lambda[\sigma] = \sum_{i=1}^{d_A^2} \frac{1}{r_i} \langle\langle A_i|\sigma\rangle\rangle \mathrm{Tr}_B(\mathbb{1}_A \otimes B_{i,B}^\dagger \Lambda_A[\rho_{AB}])$. We improve on this basic observation, by introducing and studying a simple and meaningful measure of merit for the usefulness of a fixed probe–ancilla state in channel discrimination.

For any quantum state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, we define the channel discrimination power (CDP) of ρ_{AB} on A (and similarly on B) as

$$\mathrm{CDP}_A(\rho_{AB}) := \inf_{\Lambda_0, \Lambda_1} \frac{D(\Lambda_{0,A}[\rho_{AB}], \Lambda_{1,A}[\rho_{AB}])}{D_\diamond(\Lambda_0, \Lambda_1).} \tag{4.1}$$

The infimum is taken over all pairs Λ_0, Λ_1 of quantum channels with input in $\mathcal{L}(\mathcal{H}_A)$, and we have used the notation $\Lambda_{i,A} := \Lambda_{i,A} \otimes \mathrm{id}_B$. The parameter $\mathrm{CDP}_A(\rho_{AB})$ captures how suitable ρ_{AB} is for ancilla-assisted channel discrimination as compared with the optimal distinguishability of those two channels, in a worst-case scenario approach.

In the following we report a number of results about the channel discrimination power. As notation goes, we will indicate the difference of two channels Λ_0 and Λ_1 as $\Delta = \Lambda_0 - \Lambda_1$. The channel discrimination power can then be expressed as

$$\mathrm{CDP}_A(\rho_{AB}) = \inf_{\Delta} \frac{\|\Delta_A \otimes \mathrm{id}_B[\rho_{AB}]\|_1}{\|\Delta\|_\diamond}. \tag{4.2}$$

4.2 Properties of the CDP

One can easily prove that $\mathrm{CDP}_A(\rho_{AB})$ is continuous in its argument. To show that, first notice that by definition of (Hermitian) super-operator 1-norm we have the following.

Proposition 4.1. *Let* Γ *be any Hermiticity preserving map, and* X_{AB} *Hermitian. Then*

$$\left\|\Gamma_A \otimes \mathrm{id}_B[X_{AB}]\right\|_1 \leq \|\Gamma\|_\diamond \left\|X_{AB}\right\|_1 . \tag{4.3}$$

Proposition 4.2. $\mathrm{CDP}_A(\rho)$ *is continuous:*

$$\left|\mathrm{CDP}_A(\rho_{AB}) - \mathrm{CDP}_A(\sigma_{AB})\right| \leq \left\|\rho_{AB} - \sigma_{AB}\right\|_1 , \tag{4.4}$$

for any two states ρ_{AB} *and* σ_{AB}*.*

Proof. Because of the triangle inequality and Proposition 4.1, one has

$$\begin{aligned}\left\|\Delta \otimes \mathrm{id}[\rho_{AB}]\right\|_1 &= \left\|\Delta \otimes \mathrm{id}[\sigma_{AB}] + \Delta \otimes \mathrm{id}[\rho_{AB} - \sigma_{AB}]\right\|_1 \\ &\leq \left\|\Delta \otimes \mathrm{id}[\sigma_{AB}]\right\|_1 + \|\Delta\|_\diamond \left\|\rho_{AB} - \sigma_{AB}\right\|_1 ,\end{aligned} \tag{4.5}$$

that is

$$\frac{\left\|\Delta \otimes \mathrm{id}[\rho_{AB}]\right\|_1 - \left\|\Delta \otimes \mathrm{id}[\sigma_{AB}]\right\|_1}{\|\Delta\|_\diamond} \leq \left\|\rho_{AB} - \sigma_{AB}\right\|_1 . \tag{4.6}$$

The claim follows immediately.

□

Proposition 4.3. $\mathrm{CDP}_A(\rho)$ *is monotone under local channels on* B*:*

$$\mathrm{CDP}_A(\mathrm{id}_A \otimes \Lambda_B[\rho_{AB}]) \leq \mathrm{CDP}_A(\rho_{AB}). \tag{4.7}$$

Proof. This comes directly from the monotonicity of the trace norm of Hermitian operators under channels, i.e. from $\|\Lambda[X]\|_1 \leq \|X\|_1$. One has

$$\begin{aligned}\left\|\Delta_A \otimes \mathrm{id}_B[\mathrm{id}_A \otimes \Lambda_B[\rho_{AB}]]\right\|_1 &= \left\|\Delta_A \otimes \mathrm{id}_B[\mathrm{id}_A \otimes \Lambda_B[\rho_{AB}]]\right\|_1 \\ &= \left\|\mathrm{id}_A \otimes \Lambda_B[\Delta_A \otimes \mathrm{id}_B[\rho_{AB}]]\right\|_1 \\ &\leq \left\|\Delta_A \otimes \mathrm{id}_B[\rho_{AB}]\right\|_1 ,\end{aligned} \tag{4.8}$$

for any $\Delta = \Lambda_0 - \Lambda_1$, and the claim follows. □

Proposition 4.4. *The channel discrimination power* CDP_A *is invariant under local unitaries on* A.

Proof. For any map Λ on A and any unitary U on A we can consider the map $\Lambda'[\cdot] = \Lambda[U^\dagger \cdot U]$ such that $(\Lambda_A \otimes \mathrm{id}_B)[\rho_{AB}] = (\Lambda'_A \otimes \mathrm{id}_B)[U_A \rho_{AB} U_A^\dagger]$. Given the freedom in the minimization through which CDP_A is defined, the claim follows immediately. □

Notice that Proposition 4.3 immediately implies that, for fixed dimension of A, the CDP assumes maximal value for pure states, as any bipartite state ρ_{AB} can be seen as the reduced state of a pure state $\psi_{ABB'}$, with B' a purifying system, and BB' considered together as one ancilla. Furthermore, this fact together with Proposition 4.7 imply that the CDP of a pure state only depends on its Schmidt coefficients.

4.3 CDP of pure states

We find that for pure states the CDP can be computed exactly. We will need the following lemma, which is a slight generalization of observations in, e.g., Ref. [132].

Lemma 4.1. *Let* $|\psi\rangle_{AA'} = \sum_{k=1}^{d} \sqrt{p_k}\, |a_k\rangle_A \otimes |b_k\rangle_{A'}$ *be a pure state with* $d = d_A = d_{A'}$*, and the Schmidt coefficients ordered as* $p_1 \geq p_2 \geq \ldots \geq p_d$*. Then*

$$p_d \,\|\Delta\|_\diamond \leq \|\Delta \otimes \mathrm{id}[|\psi\rangle\langle\psi|]\|_1 \,. \tag{4.9}$$

Proof. We use the fact that any pure state $|\psi\rangle_{AA'}$ can be expressed as

$$|\psi\rangle_{AA'} = (\mathbb{1} \otimes C) \left|\tilde{\psi}^+\right\rangle_{AA'}, \tag{4.10}$$

with $\left|\tilde{\psi}^+\right\rangle_{AA'} = \sum_{k=1}^{d} |k\rangle_A \otimes |k\rangle_{A'}$, and $C = \sum_{l=1}^{d} \sqrt{p_l}\, |b_l\rangle \langle a_l^*|$, where $|a_l^*\rangle$ is the basis state whose coefficients in the basis $|k\rangle$ are the complex conjugates of those of $|a_l\rangle$. Notice that the singular values of C coincide with the Schmidt coefficients of $|\psi\rangle$, and the fact that $|\psi\rangle$ is normalized implies $\|C\|_2 = 1$, hence $\|C\|_\infty \leq 1$.

The claim is trivial if $p_d = 0$. If $p_d > 0$, then C is invertible, and we can express any other state $|\phi\rangle_{AA'} = (\mathbb{1} \otimes D) \left|\tilde{\psi}^{+}\right\rangle_{AA'}$ as

$$|\phi\rangle_{AA'} = (\mathbb{1} \otimes DC^{-1}) \, |\psi\rangle_{AA'} \,. \tag{4.11}$$

Let $|\phi\rangle_{AA'}$ be the state that achieves the diamond norm $\|\Delta\|_\diamond$, that is $\|\Delta\|_\diamond = \|\Delta_A \otimes \mathrm{id}_{A'}[|\phi\rangle\langle\phi|_{AA'}]\|_1$. Then

$$\begin{aligned}
\|\Delta\|_\diamond &= \|\Delta_A \otimes \mathrm{id}_{A'}[|\phi\rangle\langle\phi|_{AA'}]\|_1 \\
&= \|(\mathbb{1} \otimes DC^{-1})\big(\Delta_A \otimes \mathrm{id}_{A'}[|\psi\rangle\langle\psi|_{AA'}]\big)(\mathbb{1} \otimes DC^{-1})^\dagger\|_1 \\
&\le \|\mathbb{1} \otimes DC^{-1}\|_\infty^2 \|\Delta_A \otimes \mathrm{id}_{A'}[|\psi\rangle\langle\psi|_{AA'}]\|_1 \\
&\le \|D\|_\infty^2 \|C^{-1}\|_\infty^2 \|\Delta_A \otimes \mathrm{id}_{A'}[|\psi\rangle\langle\psi|_{AA'}]\|_1 \\
&= p_d^{-1} \|\Delta_A \otimes \mathrm{id}_{A'}[|\psi\rangle\langle\psi|_{AA'}]\|_1,
\end{aligned} \tag{4.12}$$

where in the first inequality we have used Hölder's inequality, $|\operatorname{Tr}(XY)| \le \|X\|_\infty \|Y\|_1$, twice. For the last line, just observe that the largest singular value of C^{-1} is the reciprocal of the smallest singular value of C. □

Theorem 4.1. *Let $|\psi\rangle_{AB}$ be a pure state with Schmidt decomposition as in* (1.10). *Then, if $d_{\min} = d_A = d_B$, $\mathrm{CDP}_A(\psi_{AB}) = \mathrm{CDP}_B(\psi_{AB}) = p_{d_{\min}}$, while, if $d_{\min} = d_A < d_B$, $\mathrm{CDP}_A(\psi_{AB}) = p_{d_{\min}}$ and $\mathrm{CDP}_B(\psi_{AB}) = 0$.*

Notice that it might be that $p_{d_{\min}} = 0$, in which case both $\mathrm{CDP}_A(\psi_{AB})$ and $\mathrm{CDP}_B(\psi_{AB})$ vanish. We remark that $p_{d_{\min}}$ is a quantifier of the entanglement of $|\psi\rangle_{AB}$ [133]. Having already established that CDP_A is maximal for pure states, we find that it achieves its maximum, $1/d_A$, for maximally entangled states, e.g., for $|\psi^+\rangle_{AB} = \frac{1}{\sqrt{d_A}} \sum_{i=1}^{d_A} |i\rangle_A |i\rangle_B$. Note that it is reasonable that the maximum of the channel discrimination power, being defined as in Eq. (4.1), decreases with d_A, since the number of parameters describing an arbitrary channel with input in A increases with the size of A.

Proof. (of Theorem 4.1) Lemma 4.1 implies immediately $\mathrm{CDP}_A(|\psi\rangle\langle\psi|) \ge p_{d_A}$. We will prove the inequality in the other direction, that is, $\mathrm{CDP}_A(|\psi\rangle\langle\psi|) \le p_{d_A}$, by constructing a pair of perfectly distinguishable channels that are hard to distinguish by means of $|\psi\rangle$. We observe that, because in the case of pure states CDP_A only depends on the Schmidt coefficients, we can assume $|a_k\rangle = |b_k\rangle = |k\rangle$, without loss of generality. Let us introduce the channels

$$\Lambda_0[X] = \mathrm{Tr}[PX]\,|2\rangle\langle 2| + \mathrm{Tr}[(\mathbb{1}-P)X]\,|0\rangle\langle 0|\,, \tag{4.13}$$

$$\Lambda_1[X] = \mathrm{Tr}[PX]\,|2\rangle\langle 2| + \mathrm{Tr}[(\mathbb{1}-P)X]\,|1\rangle\langle 1|\,, \tag{4.14}$$

with $P = \sum_{i=1}^{d_A-1} |i\rangle\langle i|$ and $\mathbb{1}-P = |d_A\rangle\langle d_A|$. Then, $\Delta[X] = \langle d_A|\,X\,|d_A\rangle\,(|0\rangle\langle 0| - |1\rangle\langle 1|)$. It is clear by their definition that the two channels are perfectly distinguishable, even without the use of an ancilla, since

$$\Lambda_0[|d_A\rangle\langle d_A|] = |0\rangle\langle 0|\,, \quad \Lambda_1[|d_A\rangle\langle d_A|] = |1\rangle\langle 1|\,, \tag{4.15}$$

so that $\|\Lambda_0 - \Lambda_1\|_\diamond = \|\Lambda_0 - \Lambda_1\|_1 = 2$. On the other hand,

$$\begin{aligned}\|(\Lambda_0 - \Lambda_1)\otimes \mathrm{id}\,|\psi\rangle\langle\psi|\|_1 &= \|(|0\rangle\langle 0| - |1\rangle\langle 1|)\otimes \mathrm{Tr}_A(|d_A\rangle\langle d_A|_A\,|\psi\rangle\langle\psi|_{AB})\|_1\\ &= p_{d_A}\,\|(|0\rangle\langle 0| - |1\rangle\langle 1|)\otimes |d_A\rangle\langle d_A|)\|_1\\ &= 2p_{d_A}.\end{aligned} \tag{4.16}$$

Thus, we have proven that it must be $\mathrm{CDP}_A(|\psi\rangle\langle\psi|) \le p_{d_A}$. □

We now show that the CDP attains its maximum for the maximally entangles states.

Theorem 4.2. *The channel discrimination power* CDP_A *is maximal for maximally entangled states, for which it is equal to* $1/d_A$.

Proof. Given Propositions 4.3, it is clear that the maximum of the channel discrimination power is achieved by pure states. On the other hand, Theorem 4.1 tells us that the CDP of a pure state is equivalent to the (square) of the last Schmidt coefficient. The latter cannot be bigger than $1/d_A$, which is achieved for a maximally entangled state. □

4.4 Bounds for mixed states

We now present general bounds for the CDP.

Theorem 4.3. *Let $\rho_{AB} \in \mathcal{S}_{AB}$, with $\{A_i\}$, $\{B_i\}$ the Hermitian local orthonormal bases appearing in its OSD, cf. Eq.* (2.2)*. Then*

$$\frac{r_{d_A^2}}{d_A^{5/2}} \leq \mathrm{CDP}_A(\rho_{AB}) \leq \min_i \left\{ r_i \frac{\|B_i\|_1}{\|A_i\|_\infty} \right\} \leq r_{d_A^2} \sqrt{d_A d_B}. \tag{4.17}$$

Proof. We first prove $r_{d_A^2}/d_A^{5/2} \leq \mathrm{CDP}_A(\rho_{AB})$.

We start by finding a lower bound for the numerator in the definition of the $\mathrm{CDP}_A(\rho_{AB})$ [cf. Eq. (4.2)]. First, observe that

$$\begin{aligned} \|\Delta \otimes \mathrm{id}[\rho_{AB}]\|_1 &= \left\| \sum_i r_i \Delta(A_i) \otimes B_i \right\|_1 \\ &= \max_{-\mathbb{1} \leq M_{AB} \leq \mathbb{1}} \left| \mathrm{Tr} \left(M_{AB} \sum_i r_i \Delta(A_i) \otimes B_i \right) \right| \\ &\geq \max_{\substack{-\mathbb{1} \leq M_A \leq \mathbb{1} \\ -\mathbb{1} \leq M_B \leq \mathbb{1}}} \left| \mathrm{Tr} \left(M_A \otimes M_B \sum_i r_i \Delta(A_i) \otimes B_i \right) \right| \\ &\geq \max_i \left\{ r_i \frac{\|\Delta[A_i]\|_1}{\|B_i\|_\infty} \right\} \\ &\geq r_{d_A^2} \max_i \|\Delta[A_i]\|_1. \end{aligned} \tag{4.18}$$

The first inequality is due to restricting the class of operators M_{AB} to be product. The second inequality is due to further choosing M_A such that $\|\Delta[A_k]\|_1 = |\operatorname{Tr}(M_A \Delta[A_k])|$ and $M_B = B_k/\|B_k\|_\infty$, with k the index such that the maximum over i in the last line is achieved. Notice that, because of the orthonormality of the $B_i's$, this choice for M_B selects only one term in the OSD of ρ_{AB}. The last inequality is due to the fact that $\|B_i\|_\infty \leq \|B_i\|_2 = 1$, and that $r_i \geq r_{d_A^2}$ by assumption.

The maximally entangled state can be expressed as $|\psi^+\rangle\langle\psi^+| = \frac{1}{d_A} \sum_{i=1}^{d_A^2} C_i \otimes C_i^*$ for any orthonormal operator basis $\{C_k\} \subset \mathcal{L}(\mathcal{H}_A)$, in particular for the one appearing in the OSD of ρ_{AB}. Thus, using Lemma 4.1,

$$\begin{aligned}
\|\Delta\|_\diamond &\leq d_A \left\|\Delta \otimes \mathrm{id}[|\psi^+\rangle\langle\psi^+|]\right\|_1 \\
&= d_A \left\| \frac{1}{d_A} \sum_i \Delta\left[A_i\right] \otimes A_i^* \right\|_1 \\
&\leq \sum_i \|\Delta\left[A_i\right]\|_1 \|A_i^*\|_1 \\
&\leq d_A^{5/2} \max_i \|\Delta[A_i]\|_1 ,
\end{aligned}$$

having used the triangle inequality, the fact that there are d_A^2 terms in the sum, and that $\|A_i^*\|_1 = \|A_i\|_1 \leq \sqrt{d_A}\|A_i\|_2 = \sqrt{d_A}$. Thus, combining the above,

$$\mathrm{CDP}_A(\rho_{AB}) = \inf_\Delta \frac{\|\Delta \otimes \mathrm{id}[\rho_{AB}]\|_1}{\|\Delta\|_\diamond} \geq \frac{r_{d_A^2}}{d_A^{5/2}},$$

which completes the first part of the theorem.

We now show how to upper bound the CDP. To do that, let us consider the following channels:

$$\Lambda_i[X] = \mathrm{Tr}(X)\frac{\mathbb{1}}{d_A} + \epsilon \, \mathrm{Tr}(A_l X) Y_i, \tag{4.19}$$

for $i = 0, 1$, with traceless Hermitian operators Y_0 and Y_1, and where A_l is the local basis operator of the OSD of ρ_{AB} corresponding to the l^{th} OSC r_l. Such maps are trace-preserving by construction, and completely positive for ϵ small enough, e.g. for $\epsilon \leq 1/(d_A \|A_l\|_\infty \| \max\{\|Y_0\|_\infty, \|Y_1\|_\infty\})$. Then,

$$\Delta[X] = \epsilon \, \mathrm{Tr}(A_l X)(Y_0 - Y_1), \tag{4.20}$$

and

$$\begin{aligned}
\|\Delta \otimes \mathrm{id}[\rho_{AB}]\|_1 &= \epsilon \left\| \sum_i r_i \, \mathrm{Tr}(A_l A_i)(Y_0 - Y_1) \otimes B_i \right\|_1 \\
&= \epsilon \, \|r_l (Y_0 - Y_1) \otimes B_l\|_1 \\
&= r_l \epsilon \, \|Y_0 - Y_1\|_1 \, \|B_l\|_1 .
\end{aligned} \tag{4.21}$$

On the other hand, we claim that

$$\|\Delta\|_\diamond = \epsilon \|Y_0 - Y_1\|_1 \|A_l\|_\infty . \tag{4.22}$$

Before proving such claim, let us notice that Eqs. (4.21) and (4.22) complete the proof of the theorem. Indeed, by recalling the definition of the CDP and using Eqs. (4.21) and (4.22), one gets

$$\mathrm{CDP}_A(\rho_{AB}) \leq r_l \frac{\|B_l\|_1}{\|A_l\|_\infty}, \tag{4.23}$$

for any l, that is

$$\mathrm{CDP}_A(\rho_{AB}) \leq \min_i \left\{ r_i \frac{\|B_i\|_1}{\|A_i\|_\infty} \right\}.$$

We observe that the right-hand side can be itself upper bounded:

$$\begin{aligned}\min_i \left\{ r_i \frac{\|B_i\|_1}{\|A_i\|_\infty} \right\} &\leq r_{d_A^2} \frac{\|B_d\|_1}{\|A_d\|_\infty} \\ &\leq r_{d_A^2} \frac{d_B^{1/2}\,\|B_d\|_2}{d_A^{-1/2}\,\|A_d\|_2} \\ &= r_{d_A^2} (d_A d_B)^{1/2}, \end{aligned} \tag{4.24}$$

where we have used properties of the p-norms in the second inequality.

We now prove Eq. (4.22). To do so, let us consider an arbitrary

$$\begin{aligned}|\psi\rangle &= \sum_i \sqrt{p_i}\,|a_i\rangle\,|b_i\rangle \\ &= (\mathbb{1} \otimes C) \left|\tilde{\psi}^+\right\rangle \end{aligned} \tag{4.25}$$

where $\|C\|_2 = 1$ for $|\psi\rangle$ to be normalized (see the proof of Lemma 4.1). Notice that

$$\begin{aligned}\|\Delta \otimes \mathrm{id}[|\psi\rangle\langle\psi|]\|_1 &= \|(\mathbb{1} \otimes C)(\Delta \otimes \mathrm{id}[\left|\tilde{\psi}^+\right\rangle\left\langle\tilde{\psi}^+\right|])(\mathbb{1} \otimes C)^\dagger\|_1 \\ &= \|\epsilon(Y_0 - Y_1) \otimes C A_l^T C^\dagger\|_1 \\ &= \epsilon\|Y_0 - Y_1\|_1 \|C A_l^T C^\dagger\|_1. \end{aligned} \tag{4.26}$$

Thus, it is sufficient to prove that, for a given $X = X^\dagger$,

$$\max_{\|C\|_2=1} \|CXC^\dagger\|_1 = \|X\|_\infty. \tag{4.27}$$

Notice that $\|X\|_\infty = \|X^T\|_\infty$. Let $|x\rangle$ be the eigenvector of X corresponding to the largest eigenvalue (in modulus) $\|X\|_\infty$. Choosing $C = |x\rangle\langle x|$ we have $\|CXC^\dagger\|_1 = \| \, |x\rangle\langle x| \, X \, |x\rangle\langle x| \, \|_1 = \|X\|_\infty$, thus $\max_{\|C\|_2=1} \|CXC^\dagger\|_1 \geq \|X\|_\infty$.

To prove the other direction it is enough to prove that

$$\|CXC^\dagger\|_1 \leq \|X\|_\infty \operatorname{Tr}(C^\dagger C) = \|X\|_\infty \|C\|_2^2 = \|X\|_\infty, \tag{4.28}$$

for $X = X^\dagger$ and C satisfying $\|C\|_2 = 1$. The inequality can be seen as a trivial consequence of the fact that, for any vector $|\psi\rangle$, one has

$$\begin{aligned} | \langle\psi| \, CXC^\dagger \, |\psi\rangle \, | &= | \langle\psi| \, C(X_+ - X_-)C^\dagger \, |\psi\rangle \, | \\ &\leq \langle\psi| \, C(X_+ + X_-)C^\dagger \, |\psi\rangle \\ &\leq \|X\|_\infty \, \langle\psi| \, CC^\dagger \, |\psi\rangle \, , \end{aligned} \tag{4.29}$$

where we have used that any Hermitian matrix can be expressed as the difference of two positive semidefinite matrices with orthogonal support,

$$X = X^+ - X^-, \tag{4.30}$$

with $X^\pm \geq 0$, $X^+X^- = X^-X^+ = 0$, and that

$$X^+ + X^- \leq \|X\|_\infty \mathbb{1}. \tag{4.31}$$

We have proved the claim in Eq. (4.22), hence the theorem. □

The bounds above are not tight in general, as proven by the results about pure states. Nonetheless, they capture quantitatively, rather than purely qualitatively, the fact that the necessary and sufficient condition for ρ_{AB} to always enable ancilla-assisted discrimination and tomography of an arbitrary channel on A is that $\mathrm{OSR}(\rho) = d_A^2$.

4.5 Bound for separable states

We recall that mixed unentangled states may have maximal OSR, that is $\mathrm{OSR}(\rho_{AB}) = d_A^2$, so that, according to Eq. (4.17), they have non-zero CDP. This is the case, for example, of isotropic states, considered more in detail below.

We now focus on the case $d_A = d_B = d$. As we have seen, CDP can be as high as $1/d$. We prove that such a value cannot be achieved by states passing the realignment criterion for separability [63, 134] (cf. Sec. 2.2). The proof makes use of the following bounds, which characterize the correlations present in a state in terms of its purity, and may be of independent interest.

Lemma 4.2. *For any ρ_{AB} and any product state $\sigma_A \otimes \sigma_B$, one has $\sum_{i\geq 2} r_i^2(\rho_{AB}) = \mathrm{Tr}(\rho^2) - r_1^2 \leq \|\rho_{AB} - \sigma_A \otimes \sigma_B\|_2^2$.*

Proof. We recall that the OSC $r_i(\rho_{AB})$ are the singular values of the correlation matrix $[C_{ij}(\rho_{AB})]_{ij}$, with

$$C_{ij}(\rho_{AB}) := \langle\langle F_i \otimes G_j | \rho_{AB} \rangle\rangle, \tag{4.32}$$

where $\{F_i\}$ and $\{G_j\}$ are arbitrary local orthonormal bases for operators. We will use that, for any two matrices M and N, with ordered singular values $\sigma_i(M)$ and $\sigma_i(N)$, respectively, it holds (see Corollary 7.3.5 in [60]),

$$\sum_i (\sigma_i(M) - \sigma_i(N))^2 \leq \|M - N\|_2^2. \tag{4.33}$$

Notice that $r_i(\sigma_A \otimes \sigma_B) = 0$, for $i \geq 2$. Thus,

$$\begin{aligned}
\sum_{i\geq 2} r_i^2(\rho_{AB}) &= \sum_{i\geq 2} (r_i(\rho_{AB}) - r_i(\sigma_A \otimes \sigma_B))^2 \\
&\leq \sum_i (r_i(\rho_{AB}) - r_i(\sigma_A \otimes \sigma_B))^2 \\
&\leq \|C(\rho_{AB}) - C(\sigma_A \otimes \sigma_B)\|_2^2
\end{aligned}$$

$$
\begin{aligned}
&= \|C(\rho_{AB} - \sigma_A \otimes \sigma_B)\|_2^2 \\
&= \|\rho_{AB} - \sigma_A \otimes \sigma_B\|_2^2,
\end{aligned} \tag{4.34}
$$

having used that $\|C(X)\|_2 = \|X\|_2$ for any X. $\square$

Proposition 4.5. *For any state ρ_{AB} on $\mathbb{C}^d \otimes \mathbb{C}^d$, the smallest operator Schmidt coefficient obeys*

$$
r_{d^2} \leq \sqrt{\mathrm{Tr}(\rho^2) - \frac{1}{d^2}}. \tag{4.35}
$$

Proof. Immediate, by using Lemma 4.2 in the case $\sigma_A \otimes \sigma_B = \frac{\mathbb{1}}{d} \otimes \frac{\mathbb{1}}{d}$, and the fact that

$$
\begin{aligned}
\left\| \rho_{AB} - \frac{\mathbb{1}}{d} \otimes \frac{\mathbb{1}}{d} \right\|_2^2 &= \mathrm{Tr}\left(\left(\rho_{AB} - \frac{\mathbb{1}}{d} \otimes \frac{\mathbb{1}}{d} \right)^2 \right) \\
&= \mathrm{Tr}(\rho^2) - \frac{1}{d^2}.
\end{aligned} \tag{4.36}
$$

$\square$

Applying these bounds, we obtain the following.

Theorem 4.4. *If the OSC of ρ_{AB} satisfy $\sum_i r_i \leq 1$, then $r_d \leq r_{\mathrm{CN}}$ with*

$$
r_{\mathrm{CN}} = \frac{d(d^2-1) - \sqrt{d^2-1}}{d(d^2-1)^2 + d^3} < \frac{1}{d^2}. \tag{4.37}
$$

Proof. We want to find the maximal value r_{d^2} can assume under the condition

$$
\sum_i r_i \leq 1. \tag{4.38}
$$

We notice that Proposition 4.5 implies that the OSC of every state respect

$$
r_{d^2}^2 \leq \sum_i r_i^2 - \frac{1}{d^2} \tag{4.39}
$$

(recall that $\mathrm{Tr}(\rho^2) = \sum_i r_i^2$). We want aim to find the maximum of r_{d^2} under conditions (4.38) and (4.39), irrespectively of the physicality of the choice coefficient—as long as they respect (4.38) and (4.39). Notice that, by definition, $r_i \geq 0$, and $r_1 \geq r_2 \geq \ldots \geq r_d^2$.

It it clear that the maximum r_{d^2} will be found for the condition (4.38) being satisfied with equality, since, if the left-hand side of (4.38) was smaller than 1, then we could increase all the OSC, including r_{d^2}, to make it equal to 1. Moreover, for fixed r_{d^2}, the largest value of $\sum_i r_i^2$ is achieved for $r_2 = r_3 = \ldots = r_{d^2} = r$ and $r_1 = 1-r$. This is due to the fact that $\sum_i r_i^2$ is Schur convex. Thus, we can find the maximal r_{d^2} compatible with the constraints, by finding the largest r such that

$$r^2 \leq (d^2-1)r^2 + (1-(d^2-1)r)^2 - \frac{1}{d^2}. \tag{4.40}$$

One finds that such a value is given by

$$r_{CN} = \frac{d(d^2-1) - \sqrt{d^2-1}}{d(d^2-1)^2 + d^3} < \frac{1}{d^2}. \tag{4.41}$$

$\square$

By combining Theorem 4.4 with Theorem 4.3 we prove that, if the OSC of ρ_{AB} satisfy $\sum_i r_i \leq 1$, then $\mathrm{CDP}(\rho_{AB}) \leq r_{\mathrm{CN}} d < 1/d$. We remark that the realignment criterion for separability is satisfied by all separable states, and by many (weakly) entangled states [11, 63, 134].

4.6 Relation with discord

As we have just seen, entanglement is needed to achieve the maximal possible CDP. Nonetheless, separable states can have non-vanishing CDP, when they have maximal OSR. As pointed out in Ref. [94], this is not possible for states that do not exhibit quantum discord. As seen in Sec.1.2.2, a bipartite state is classical on A if it can be expressed as $\rho_{AB} = \sum_i p_i |a_i\rangle\langle a_i|_A \otimes \rho_i^B$, for some orthonormal basis $\{|a_i\rangle\}$. Such states manifestly have OSR $\leq d_A$. On the contrary, states which are not classical on A may be detected as discordant by looking at their OSR [94, 96].

In this section we shed light on the role of discord in channel discrimination. To do so, it will be convenient to first study the behaviour of the CDP under the action of maps that reduce the OSR.

Theorem 4.5. *We have*

$$\mathrm{CDP}_A(\rho_{AB}) \leq \min_{\substack{\Lambda\ s.t.\\ \mathrm{OSR}(\Lambda\otimes\mathrm{id}[\rho_{AB}])<d_A^2}} \|\rho_{AB} - \Lambda\otimes\mathrm{id}[\rho_{AB}]\|_1 \tag{4.42}$$

where the minimization is over all channels that acting on A reduce the OSR of ρ_{AB} to less than maximal.

Proof. It holds

$$\begin{aligned}\|\Delta\otimes\mathrm{id}[\rho_{AB}]\|_1 &\leq \|\Delta\otimes\mathrm{id}[\rho_{AB} - \Lambda\otimes\mathrm{id}[\rho_{AB}]]\|_1 + \|(\Delta\circ\Lambda)\otimes\mathrm{id}[\rho_{AB}]\|_1\\ &\leq \|\Delta\|_\diamond \|\rho_{AB} - \Lambda\otimes\mathrm{id}[\rho_{AB}]\|_1 + \|\Delta\otimes\mathrm{id}[\Lambda\otimes\mathrm{id}[\rho_{AB}]]\|_1\,, \end{aligned}\tag{4.43}$$

having used Proposition 4.1. Then,

$$\begin{aligned}\inf_\Delta \frac{\|\Delta\otimes\mathrm{id}[\rho_{AB}]\|_1}{\|\Delta\|_\diamond} &\leq \|\rho_{AB} - \Lambda\otimes\mathrm{id}[\rho_{AB}]\|_1 + \inf_\Delta \frac{\|\Delta\otimes\mathrm{id}[\Lambda\otimes\mathrm{id}[\rho_{AB}]]\|_1}{\|\Delta\|_\diamond}\\ &= \|\rho_{AB} - \Lambda\otimes\mathrm{id}[\rho_{AB}]\|_1\,, \qquad (4.44)\\ & \qquad (4.45)\end{aligned}$$

where we have used that the CDP of $\Lambda\otimes\mathrm{id}[\rho_{AB}]$ (the second term on the right-hand side of the inequality) vanishes under the assumption $\mathrm{OSR}\left(\Lambda\otimes\mathrm{id}\left[\rho_{AB}\right]\right) < d_A^2$. The claim then follows. □

As a particular example involving the last theorem, let $\Pi[X] = \sum_{i=1}^d |i\rangle\langle i|\, X\, |i\rangle\langle i|$ be the channel which dephases in an arbitrary basis. Then

$$\mathrm{CDP}_A(\rho_{AB}) \;\; \leq \min_{\Pi_A\otimes\mathrm{id}_B} \|\rho_{AB} - \Pi_A\otimes\mathrm{id}_B[\rho_{AB}]\|_1\,. \tag{4.46}$$

In the light of the last theorem, we find that

$$\begin{aligned}\mathrm{CDP}_A(\rho_{AB}) &\leq \min_{\substack{\Lambda_A \text{ s.t.} \\ OSR(\Lambda_A[\rho_{AB}])<d_A^2}} 2D(\rho_{AB}, \Lambda_A[\rho_{AB}]) \\ &\leq \min_{\Pi_A} 2D(\rho_{AB}, \Pi_A[\rho_{AB}]).\end{aligned} \tag{4.47}$$

The first minimization is over channels that reduce the OSR of ρ_{AB} to less than maximal. The second minimization is over projective measurements of the form $\Pi[L] = \sum_i |a_i\rangle\langle a_i| L |a_i\rangle\langle a_i|$, for a choice of basis $\{|a_i\rangle\}$ to be optimized over. The quantity on the second line is a known geometric discord quantifier [135]. Thus, we see that the bipartite state ρ_{AB} must be contain a large amount of discord in order for ρ_{AB} to be useful in one-shot, worst-case ancilla-assisted channel discrimination.

4.7 CDP of isotropic states

As an example that goes beyond pure states, here we consider the class of isotropic states [99] defined in Eq. (3.13). As already observed, this is a paradigmatic class of noisy states that interpolates between an uncorrelated state (for $p = 0$) and a maximally entangled state (for $p = 1$). Isotropic states are separable for $0 \leq p \leq \frac{1}{d+1}$ and entangled for $\frac{1}{d+1} < p \leq 1$. This is also the class of states used in Ref. [15] in the context of ancilla-assisted channel tomography, where it was already observed that this class of states enables channel tomography as soon as $p > 0$. It is known and immediate to check that

$$\left|\psi^+\right\rangle\!\left\langle\psi^+\right| = \frac{1}{d}\sum_{k=1}^{d^2} A_k \otimes A_k^* \tag{4.48}$$

for any orthonormal operator basis $\{A_k\}$, with complex conjugation taken in the local Schmidt basis of the maximally entangled state. We can choose $A_1 = \frac{\mathbb{1}}{\sqrt{d}}$, and find immediately

$$\rho_{AB}(p) = \frac{1}{d}\frac{\mathbb{1}}{\sqrt{d}} \otimes \frac{\mathbb{1}}{\sqrt{d}} + \frac{p}{d}\sum_{k=2}^{d^2} A_k \otimes A_k^*, \tag{4.49}$$

where $\{A_k\}$ is any collection of d^2-1 traceless orthonormal operators. Thus, the OSC of $\rho_{AB}(p)$ are evidently $(1/d, p/d, \dots, p/d)$. Notice that $r_{d^2} = p/d$, so that the general bounds (4.17) become $p/d^{7/2} \leq \mathrm{CDP}(\rho_{AB}(p)) \leq p$; we are able to prove the following bounds, which reproduce the correct value for CDP in the limit in which the isotropic states become maximally entangled.

Theorem 4.6. *For the isotropic state it holds*

$$\frac{p}{d+1-p} \leq \mathrm{CDP}_A(\rho_{AB}(p)) \leq \min\left\{2\frac{p}{d}, \frac{1}{d}\right\}. \tag{4.50}$$

Proof. We start by proving the upper bound. That $\mathrm{CDP}_A(\rho_{AB}(p)) \leq 1/d$ can be straightforwardly be verified by using the same two maps (4.13) and (4.14) that were used to prove the upper bound for pure states. In order to prove $\mathrm{CDP}_A(\rho_{AB}(p)) \leq 2p/d$, we will use the bound $\mathrm{CDP}_A(\rho_{AB}) \leq \min_i \left\{ r_i \frac{\|B_i\|_1}{\|A_i\|_\infty} \right\}$ from Theorem 4.3, exploiting the freedom in choosing the decomposition (4.49). E.g., we can choose $A_2 = (|1\rangle\langle 2| + |2\rangle\langle 1|)/\sqrt{2}$, with $B_2 = A_2^* = A_2$, so that $\|A_2\|_\infty = 1/\sqrt{2}$ and $\|B_2\|_1 = \sqrt{2}$. Thus,

$$\mathrm{CDP}_A(\rho_{AB}(p)) \leq r_2 \frac{\|B_2\|_1}{\|A_2\|_\infty} = \frac{p}{d}\frac{\|A_2\|_1}{\|A_2\|_\infty} = \frac{p}{d}2.$$

For the lower bound, we generalize the approach of Lemma 4.1. Given two arbitrary channels, let $|\psi\rangle\langle\psi|$ be optimal for the diamond norm of their difference, i.e.

$$\|\Delta\|_\diamond = \sup_\rho \|\Delta \otimes \mathrm{id}[\rho]\|_1 = \|\Delta \otimes \mathrm{id}[|\psi\rangle\langle\psi|]\|_1 \tag{4.51}$$

and let us consider C such that

$$|\psi\rangle_{AA'} = (\mathbb{1} \otimes C) \left|\tilde{\psi}^+\right\rangle_{AA'} \tag{4.52}$$

Notice that $\mathrm{Tr}_A(|\psi\rangle\langle\psi|) = CC^\dagger$, with $CC^\dagger \geq 0$ a normalized state.

Let us define the state

$$\sigma(p) := (1-p)\frac{\mathbb{1}}{d} \otimes CC^\dagger + p\,|\psi\rangle\langle\psi|$$

$$
\begin{aligned}
&= d(\mathbb{1} \otimes C) \left[(1-p)\frac{\mathbb{1}}{d} \otimes \frac{\mathbb{1}}{d} + p \left|\psi^+\right\rangle\!\left\langle\psi^+\right| \right] (\mathbb{1} \otimes C^\dagger) \\
&= d(\mathbb{1} \otimes C)\ \rho_{AB}(p)\ (\mathbb{1} \otimes C^\dagger).
\end{aligned} \tag{4.53}
$$

Then,

$$
|\psi\rangle\langle\psi| = \frac{1}{p}\left[\sigma(p) - (1-p)\frac{\mathbb{1}}{d} \otimes CC^\dagger\right], \tag{4.54}
$$

and

$$
\begin{aligned}
\|\Delta\|_\diamond &= \|\Delta \otimes \mathrm{id}[|\psi\rangle\langle\psi|]\|_1 \\
&= \left\| \frac{1}{p}\left[\Delta \otimes \mathrm{id}[\sigma(p)] - (1-p)\Delta\left[\frac{\mathbb{1}}{d}\right] \otimes CC^\dagger\right] \right\|_1 \\
&\leq \frac{1}{p}\|\Delta \otimes \mathrm{id}[\sigma(p)]\|_1 + \frac{1-p}{p}\left\|\Delta\left[\frac{\mathbb{1}}{d}\right]\right\|_1 \\
&= \frac{d}{p}\left\|(\mathbb{1} \otimes C)\ \Delta \otimes \mathrm{id}[\rho_{AB}(p)]\ (\mathbb{1} \otimes C^\dagger)\right\|_1 + \frac{1-p}{p}\left\|\Delta\left[\frac{\mathbb{1}}{d}\right]\right\|_1 \\
&\leq \frac{d}{p}\|C\|_\infty^2 \|\Delta \otimes \mathrm{id}[\rho_{AB}(p)]\|_1 + \frac{1-p}{p}\left\|\Delta\left[\frac{\mathbb{1}}{d}\right]\right\|_1 \\
&\leq \frac{d}{p}\|\Delta \otimes \mathrm{id}[\rho_{AB}(p)\|_1 + \frac{1-p}{p}\left\|\Delta\left[\frac{\mathbb{1}}{d}\right]\right\|_1.
\end{aligned} \tag{4.55}
$$

Lastly, since $\frac{\mathbb{1}}{d} = \mathrm{Tr}_B(\rho_{AB}(p))$ and the partial trace is a channel, the monotonicity of the trace distance implies

$$
\begin{aligned}
\left\|\Delta\left[\frac{\mathbb{1}}{d}\right]\right\|_1 &= \|\Delta_A\left[\mathrm{Tr}_B(\rho_{AB}(p))\right]\|_1 \\
&= \|\mathrm{Tr}_B\left(\Delta_A[\rho_{AB}(p)]\right)\|_1 \\
&\leq \|\Delta \otimes \mathrm{id}[\rho_{AB}(p)]\|_1.
\end{aligned} \tag{4.56}
$$

Thus,

$$
\|\Delta\|_\diamond \leq \left(\frac{d+1-p}{p}\right)\|\Delta \otimes \mathrm{id}[\rho_{AB}(p)]\|_1, \tag{4.57}
$$

from which we obtain

$$
\mathrm{CDP}_A(\rho_{AB}(p)) \geq \frac{p}{d+1-p}. \tag{4.58}
$$

□

4.8 Conclusion

In this chapter we focused on the usefulness of quantum correlations for ancilla-assisted channel discrimination with fixed input, introducing the channel discrimination power (CDP) of the input state. We argued that the key factor that dictates the CDP of a state is its smallest operator Schmidt coefficient. We proved that the CDP is maximal for maximally entangled states. This can be seen as an argument to consider the Choi-Jamiołkowski isomorphism [52, 53] as the best possible one-to-one mapping between states and maps. We derived general bounds for the CDP that allowed us to prove that highly entangled states outperform—in the sense of having a larger CDP—all states that pass the so-called realignment criterion of separability [63, 134]. We added to the list of quantum information processing tasks for which the quantum discord provides a bound on the performance: we proved that a disturbance-based discord quantifier bounds the CDP. Several questions remain open, like whether the CDP is equal to the lowest operator-Schmidt-coefficient of the state, and which channels are the hardest to discriminate for a given input state. Finally, while the CDP is defined in terms of optimal probability of channel discrimination, it would be interesting to consider more in general how a probe–ancilla state induces a mapping between a metric on the space of channels and a metric in the space of output probe–ancilla states (see Chapter 6).

Chapter 5

Correlation–assisted process tomography

One critical element in accomplishing the progress promised by quantum information is the ability to completely and accurately characterize quantum physical processes. In the extensive literature upon the subject, one recognises two major, extreme examples of quantum process tomography: standard quantum process tomography (SPT) and ancilla–assisted quantum process tomography (AAPT). In the general case of SPT, an unknown quantum channel Λ acting on a d-level system (also called a qudit) can be reconstructed through its action on an ensemble of linearly independent input states [13, 16, 17]. In particular, a probe is prepared in a fixed set of d^2 input states $\{\rho_i\}$, which form a basis for the space of qudit linear operators. Each of the ρ_i states goes through the process Λ to be characterized, and the outputs $\Lambda[\rho_i]$ are determined using quantum state tomography [13,136,137] (see Figure 5.1). Once the outputs are known,

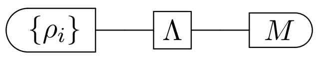


Figure 5.1: Standard quantum process tomography. To reconstruct the action of a channel Λ acting on a d-dimensional system A, d^2 linearly independent input states $\{\rho_i\}$ are needed, with state tomography done on the outputs by means of measurement(s) M. Time goes from left to right. Single lines represent quantum systems, and boxes represent operations: a square box has quantum input and quantum output, while a D-shaped (reverse-D-shaped) box has only quantum input (quantum output).

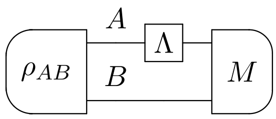


Figure 5.2: Ancilla-assisted quantum process tomography. One input state ρ_{AB} suffices, as long as it has Operator Schmidt Rank equal to the square of the input dimension of the channel.

the evolution under Λ of any arbitrary operator can be determined uniquely by linearity, thus characterizing the channel. An alternative tomographic technique is offered by the renowned AAPT [15,130,138] which, in contrast to SPT, needs only one single bipartite input state. As seen in the previous chapter, the possibility of performing AAPT can be seen as a consequence of the correspondence between linear maps and linear operators established by the well-known Choi-Jamiołkowski isomorphism [52, 53]. In general, an ancillary system B is prepared in a correlated state ρ_{AB} with the quantum system subject to the channel to be determined, the probe A. Complete information about the channel can be imprinted on the global state by the action of the process on the probe alone, and then extracted by state tomography on the bipartite output state (see Figure 5.2).

An input enabling enabling AAPT is the maximally entangled state $|\Phi\rangle_{AB} = \sum_{i=1}^{d} d^{-1/2}\,|i\rangle \otimes |i\rangle$, with the output $\rho_\Lambda = (\Lambda \otimes \mathrm{id})[|\Phi\rangle\langle\Phi|_{AB}]$ simply being the Choi-Jamiołkowski state isomorphic to Λ [52, 53, 138].

However, it was observed [15,130] that the key property for a bipartite input to enable AAPT is that of having maximal Operator Schmidt Rank (OSR), with a refining of this observation being that the channel discrimination power of a bipartite state is dictated by its smallest operator Schmidt coefficient [1]. It follows that, in principle, also non-entangled but correlated states can be used to perform AAPT. Bipartite states carrying a complete imprinting of a channel acting on one of the two subsystems were defined as *faithful* in Ref. [130]. Nonetheless, non-faithful states can still be used to obtain substantial albeit partial information on the action of a channel. This observation suggests that the property of being faithful can be associated with a *set* of bipartite states, the latter being faithful when any unknown channel can be fully retrieved from the tomographic reconstruction of the corresponding output states [130]. Indeed, SPT can be seen as an extreme case of such a situation, where the presence of an ancilla is actually irrelevant, and one just uses a faithful set of probe states. We remark that the correlations present in one or more of the bipartite states of a faithful set can be deemed as effectively assisting process tomography as long as the faithful set comprises less than d^2 states. The results in this chapter lie between the two archetypical techniques sketched above, and focus on the exploitation of correlations to reduce the number of distinct inputs needed for what we could call in general *correlation-assisted process tomography* (CAPT).

5.1 Correlation-assisted process tomography

The tomographic scheme proposed in this section arises from the question of whether and how a faithful set can be generated by means of local actions $\{\Gamma_i\}$ on a fixed input state (see Figure 5.3). In the case where there is no ancilla (or, if there is an ancilla, where there are no probe–ancilla correlations), a local action is not very different from simply considering d^2 inputs, but one may need strictly less than d^2 local operations on the input if correlations are present between probe and ancilla. That is, our results may be interpreted as an interpolation between the use of fully uncorrelated or fully correlated (that is, having maximum OSR) input states.

We show in general that a faithful set can always be generated via $\left\lceil \frac{d^2}{\mathrm{OSR}(\rho_{AB})} \right\rceil$ local transformations on a *fixed bipartite state*. Notice that this is optimal, as it is clearly impossible to generate a faithful set with less local operations. We also consider the

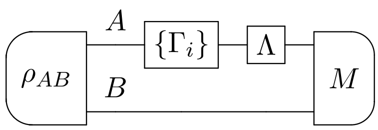


Figure 5.3: Correlation-assisted quantum process tomography. Any bipartite state ρ_{AB} can be used in this scheme. The presence of correlations in the state may substantially reduce the number of known channels $\{\Gamma_i\}$ that need to be applied so that $\{\Gamma_{i,A}[\rho_{AB}]\}$ is a faithful set. Standard process tomography and ancilla-assisted process tomography are extreme cases of this more general scenario.

case where such local transformations are constrained to be unitary. For pure fixed states, we find that such a constraint does not change the result: any pure bipartite state of Schmidt rank k (hence with OSR equal to k^2) can be used to generate a faithful set with $\left\lceil \frac{d^2}{k^2} \right\rceil$ unitaries. For mixed states, the constraint can actually be limiting: we exhibit a class of qudit-qudit states with OSR equal to two but such that one still needs d^2 local unitaries to generate a faithful set. We conjecture that in general one may need $\left\lceil \frac{d^2}{\mathrm{OSR}(\rho_{AB})-1} \right\rceil$ local unitaries to generate a faithful set. On the other hand, the mixed-state case can display a highly "efficient" (in terms of local unitaries employed) generation of a faithful set; specifically, we exhibit a family of qudit-qudit states with OSR $\approx d^2/2$ where only two unitaries are needed to achieve faithfulness; notice that such a case is impossible in the pure-state case. Finally, by exploiting the relation between $SO(3)$ and $SU(2)$ (that is, in a sense, the Bloch ball qubit representation), we fully characterize the qubit-qudit case for qubit channels, once more highlighting the importance of discord in the issue of correlation-assisted channel tomography/discrimination: a two-qubit state gives rise to a faithful set with at most two local unitaries if an only if it exhibits discord on the probe side.

5.2 Generating a faithful set of inputs with general local channels

As anticipated, AAPT requires the preparation of a bipartite system in a single bipartite state ρ_{AB}. One subsystem (the probe) is sent through the channel Λ to be characterized. Using Eq. (2.2), the output $\rho_\Lambda := (\Lambda \otimes \mathrm{id})[\rho_{AB}]$ reads

$$\rho_\Lambda = \sum_{l=1}^{OSR(\rho_{AB})} r_l \Lambda[A_l] \otimes B_l. \tag{5.1}$$

Then, by reconstructing ρ_Λ, one recovers the action of the channel on the basis element A_l via $\Lambda[A_l] = \mathrm{Tr}_B((\mathbb{1} \otimes B_l^\dagger)\rho_\Lambda)/r_l$ (for $r_l > 0$). It follows that inputs with maximal OSR enable complete characterization of the channel, since its action on a complete operator basis of $\mathcal{L}(\mathcal{H}_A)$ can be reconstructed [15, 130]. It is clear that input states defined as faithful are states with maximal OSR, more precisely with OSR $= d_A^2$. Correspondingly, a set of (potentially unfaithful, when considered individually) bipartite states $\{\rho_{AB,i}\}$ is called faithful if the local operators $\{A_{l,i}\}_{l=1}^{\mathrm{OSR}(\rho_{AB,i})}$ of OSD($\rho_{AB,i}$), when considered together, generate the whole $\mathcal{L}(\mathcal{H}_A)$, i.e., if $\mathrm{span}(\{A_{l,i}\}_{l,i}) = \mathcal{L}(\mathcal{H}_A)$.

The core idea of this chapter is to show that the correlations of a fixed bipartite state, of whatever degree, can in principle be exploited to allow "more efficient" process tomography. Such correlations can be measured through the logarithm of the OSR (see Sec. 3.3). Indeed, the OSR is minimal for and only for non correlated (product) states, and it is monotone under local channels. Not relying on correlations, like in SPT, is the same as considering minimal OSR – that is, OSR equal to one – for the inputs. On the other hand, fully relying on correlations, like in AAPT, means requiring maximal OSR for a single bipartite input. These considerations legitimate the intuition that intermediate values for the OSR should be consistent with the use of an intermediate number of inputs.

In this section we analyze how we can achieve the condition $\mathrm{span}(\{A_{l,i}\}) = \mathcal{L}(\mathcal{H}_A)$ indicated earlier, where each set $\{A_{l,i}\}_l$ is a local orthonormal basis for $(\Gamma_i \otimes \mathrm{id})[\rho_{AB}]$, for Γ_i the local operations applied on the probe before it is subject to the channel Λ (see Figure 5.3).

Let us first remark why channel tomography is certainly possible in this setup. The reason is simple: each channel Γ_i may simply be taken to have constant output corresponding to one of the input states ρ_A^i used in standard channel tomography (Figure 5.1). With this "trivial" strategy, we do not make use of correlations at all, but we certainly achieve the task at hand. Having established this, let us move to the issue of "optimizing" the Γ_i's, at least with respect to their number.

Let $\{A_l\}_{l=1}^{d^2}$ be a Hermitian local orthonormal basis for the operator Schmidt decomposition for ρ_{AB}, comprising the OSR(ρ_{AB}) elements corresponding to non-zero OSCs. It is clear that $\mathrm{span}(\{A_{l,i}\}_l) = \mathrm{span}(\{\Gamma_i[A_l]\}_l)$, so that $\mathrm{span}(\{A_{l,i}\}_{l,i}) = \mathrm{span}$ $(\{\Gamma_i[A_l]\}_{l,i})$. Thus, our goal is the following: given $\{A_l\}_{l=1}^{\mathrm{OSR}(\rho_{AB})}$, find a (minimal) way of choosing the local maps Γ_i so that $\mathrm{span}(\{\Gamma_i[A_l]\}_{l,i}) = \mathcal{L}(\mathcal{H}_A)$. By minimal, we mean that we want to identify the smallest possible number of local channels Γ_i that are needed to achieve such a condition. In the following we provide a construction to achieve this.

Let us consider the following family of maps,

$$\Gamma_i[X] := (1-\epsilon)\,\mathrm{Tr}(X)\frac{\mathbb{1}}{d} + \epsilon\,\tilde{\Gamma}_i[X], \tag{5.2}$$

which are each a convex combination (for $0 \leq \epsilon \leq 1$) of the totally depolarizing channel $X \mapsto \mathrm{Tr}(X)\frac{\mathbb{1}}{d}$ and of

$$\tilde{\Gamma}_i[X] := \sum_j \mathrm{Tr}(A_j X) A_{\gamma_{j,i}} + \left[\mathrm{Tr}(X) - \sum_j \mathrm{Tr}(A_j X)\,\mathrm{Tr}(A_{\gamma_{j,i}})\right]\frac{\mathbb{1}}{d}. \tag{5.3}$$

Notice that $\tilde{\Gamma}_i$ is not necessarily a channel, but it is a linear map that is trace preserving by construction. Here we denote $\gamma_{j,i} := j \oplus (i \cdot \mathrm{OSR}(\rho_{AB}))$, and the A_i's form a local orthonormal basis that is a superset of the local operators of the OSD of ρ_{AB}. The symbol $\oplus$ indicates addition modulo d^2, and we let $i = 1, 2, 3 \ldots$. By construction, the maps $\Gamma_i[X]$ are trace preserving, and, for $0 < \epsilon < 1$ small enough, completely positive. This is because, within the set of linear trace-preserving maps, there is a ball of completely positive maps around the totally depolarizing channel. Notice that, in principle, we could consider any other channel with full-rank fixed output, at the "cost" of considering some other ϵ. Such full-rank fixed output (as well as ϵ) could even be made to depend on i.

It is easy to recognize the action of Γ_i on the generic basis element A_n:

$$\Gamma_i[A_n] = \big(\,\mathrm{Tr}(A_n - \epsilon A_{\gamma_{n,i}})\big)\frac{\mathbb{1}}{d} + \epsilon\, A_{\gamma_{n,i}}, \tag{5.4}$$

i.e. the $i-$th map, acting upon the n-th element of the local basis, returns a linear combination of the basis element indexed as $n \oplus (i \cdot \mathrm{OSR}(\rho_{AB}))$ and of the identity. To make the action of the channels clearer, let us consider the action of, e.g., Γ_1. The latter would map the set $\{A_n\}_{n=1}^{\mathrm{OSR}(\rho_{AB})}$ to $\{\big(\,\mathrm{Tr}(A_n - \epsilon A_{n\oplus\mathrm{OSR}(\rho_{AB})})\big)\frac{\mathbb{1}}{d} + \epsilon\, A_{n\oplus\mathrm{OSR}(\rho_{AB})}\}_{n=1}^{\mathrm{OSR}(\rho_{AB})}$. Thus, it should be clear that – up to the detail of whether we need to choose a fixed state different from the maximally mixed one in (5.2) and (5.3) in order to certainly obtain a set whose elements are all linearly independent from the other generated sets – given an incomplete set of basis elements $\{A_n\}_{n=1}^{\mathrm{OSR}(\rho_{AB})}$, we are able to obtain operators spanning the same space as the remaining $d_A^2 - \mathrm{OSR}(\rho_{AB})$ ones through the application of $\lceil \frac{d_A^2}{\mathrm{OSR}(\rho_{AB})} \rceil - 1$ channels. This leads directly to the following theorem.

Theorem 5.1. *Let ρ_{AB} be a bipartite state with $OSR(\rho_{AB}) = k$. Then there is a set of quantum channels Γ_i, with $i = 1, \dots, \lceil \frac{d_A^2}{k} \rceil$, such that $\{(\Gamma_i \otimes \mathrm{id})[\rho_{AB}]\}_i$ is faithful. Without loss of generality, one of the channels can be taken to be identity channel.*

5.3 Generating a faithful set of inputs with unitary local channels

In this section we consider constraining the local channels Γ_i that act on the probe in Fig. 5.3 to be unitary. The question we address is that of determining how many unitary rotations U_i are needed in order to obtain a faithful set of input states $\{U_i \otimes \mathbb{1}\rho_{AB}U_i^\dagger \otimes \mathbb{1}\}_i$. As discussed in the case of general local operations, this corresponds to finding out how many unitary rotations are needed so that $\{U_i A_l U_i^\dagger\}_{i,l}$ spans the entire space $\mathcal{L}(\mathcal{H}_A)$, where $\{A_l\}_l$ is a set of orthonormal local OSD operators corresponding to non-zero operator Schmidt coefficients.

We remark that in the case where we impose the constraint that the channels be unitary, the fact that process tomography is possible at all is not immediate. Indeed, it is not anymore the case that this is possible for *all* input states ρ_{AB}. Nonetheless, we prove that it is possible for all states that are not of the form $\frac{\mathbb{1}_A}{d} \otimes \rho_B$: notice that the latter states are not only uncorrelated, but such that the state of the probe A is maximally mixed. Notice also that this means that any form of correlations is enough to make process tomography by local unitaries possible.

Let us first establish this result.

5.4 Process tomography via local unitary rotation of almost any input

It is convenient to recall the definition of frame [139]. Such a concept is generally defined for families of vectors in inner product spaces. In our framework we exploit the inner product structure of $\mathcal{L}(\mathcal{H})$ and define a frame as a collection $\{P_k\}$ of operators such that there are real numbers $0 < a \leq b < \infty$ satisfying

$$a\|X\|_2^2 \leq \sum_k |\operatorname{Tr}(P_k X)|^2 \leq b\|X\|_2^2 \tag{5.5}$$

for any $X \in \mathcal{L}(\mathcal{H})$. A frame generalizes the notion of basis. Notice in particular that, if the frame is actually an orthonormal basis, that is, if $\operatorname{Tr}(P_k^\dagger P_l) = \delta_{kl}$, then the frame condition (5.5) is satisfied with $a = b = 1$. In finite dimensions, a finite collection $\{P_k\}$ is a frame for $\mathcal{L}(\mathcal{H})$ if and only if it is a spanning set for $\mathcal{L}(\mathcal{H})$, while an infinite collection $\{P_k\}$, even when a spanning set, may not constitute a frame, as there might not be a finite b that satisfies (5.5). The lower bound in (5.5), for $a > 0$, ensures that X can be reconstructed from the values $\operatorname{Tr}(P_k X)$. It should be clear that, given a frame, one can always consider a subset of the elements of the frame, so that such subset forms a basis, that is, a spanning set of linearly independent operators.

What we will prove is that it is possible to choose d^2 unitaries $\{U_i\}_{i=1}^{d^2}$ (with $U_1 = \mathbb{1}$ without loss of generality) so that $\{U_i \rho U_i^\dagger\}_i$ is a frame and a basis for the space of operators of the input ancilla, initially prepared in the state ρ, as long as $\rho \not\propto \mathbb{1}$. To prove this, we will need the notion of twirling, or twirl operation [31]. The latter is the linear projection $\mathscr{T}$ on bipartite operators $Y \in \mathcal{L}(\mathbb{C}^d \otimes \mathbb{C}^d)$ defined as

$$\mathscr{T}(Y) = \int_U (U \otimes U) Y (U^\dagger \otimes U^\dagger) dU, \tag{5.6}$$

where the integral is taken with respect to the Haar measure of the unitary group in $\mathbb{C}^d$. Since any operator commuting with all unitaries of the form $U \otimes U$ can be written as a linear combination of $\mathbb{1}$ and V (where V is the flip operator, defined implicitly by its action $V|\psi\rangle \otimes |\varphi\rangle = |\varphi\rangle \otimes |\psi\rangle$, for all $|\psi\rangle, |\varphi\rangle \in \mathbb{C}^d$), it follows that [31]

$$\mathscr{T}(Y) = \alpha(Y)\mathbb{1} + \beta(Y)V.$$

The coefficients $\alpha(Y)$ and $\beta(Y)$ are fixed by the conditions

$$\operatorname{Tr}(\mathscr{T}(Y)) = \operatorname{Tr}(Y) \tag{5.7}$$

$$\operatorname{Tr}(\mathscr{T}(Y)V) = \operatorname{Tr}(YV), \tag{5.8}$$

solved by

$$\alpha(Y) = \frac{d\operatorname{Tr}(Y) - \operatorname{Tr}(YV)}{d^3 - d}, \tag{5.9}$$

$$\beta(Y) = \frac{d\operatorname{Tr}(YV) - \operatorname{Tr}(Y)}{d^3 - d}. \tag{5.10}$$

We will use the fact that the twirling can be approximated by a unitary 2-design, that is by a finite set of n unitaries $\{U_i\}_{i=1}^n$ such that [31]

$$\mathscr{T}(Y) = \frac{1}{n}\sum_{i=1}^{n} U_i \otimes U_i Y U_i^\dagger \otimes U_i^\dagger. \tag{5.11}$$

We will obtain our frame by taking d^2 of such unitaries.

Let $\{U_i\}_{i=1}^n$ be a unitary 2-design (without loss of generality, one of the unitaries can be taken to be the identity). Let us check the frame conditions (5.5) of $\{U_i A U_i^\dagger\}_{i=1}^n$, for an arbitrary $A \in \mathcal{L}(\mathbb{C}^d)$. One has

$$\begin{aligned}
&\sum_{i=1}^{n} |\operatorname{Tr}(U_i A U_i^\dagger X)|^2 \\
&= \sum_{i=1}^{n} \operatorname{Tr}(U_i A U_i^\dagger X)\operatorname{Tr}(U_i A^\dagger U_i^\dagger X^\dagger) \\
&= \sum_{i=1}^{n} \operatorname{Tr}\Big((U_i \otimes U_i)(A \otimes A^\dagger)(U_i^\dagger \otimes U_i^\dagger)(X \otimes X^\dagger)\Big) \\
&\propto \operatorname{Tr}(\mathscr{T}(A \otimes A^\dagger)X \otimes X^\dagger)
\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{Tr}((\alpha(A \otimes A^\dagger)\mathbb{1} + \beta(A \otimes A^\dagger)V)X \otimes X^\dagger) \\
&= \alpha(A \otimes A^\dagger)\,\mathrm{Tr}(X \otimes X^\dagger) + \beta(A \otimes A^\dagger)\,\mathrm{Tr}(VX \otimes X^\dagger) \\
&= \alpha(A \otimes A^\dagger)|\,\mathrm{Tr}(X)|^2 + \beta(A \otimes A^\dagger)\|X\|_2^2, \qquad (5.12)
\end{aligned}
$$

where $\alpha(A \otimes A^\dagger)$ and $\beta(A \otimes A^\dagger)$ are given by Eqs. (5.9) and (5.10) applied to the case $Y = A \otimes A^\dagger$, so that

$$
\alpha(A \otimes A) = \frac{d|\,\mathrm{Tr}(A)|^2 - \|A\|_2^2}{d^3 - d}, \qquad (5.13)
$$

$$
\beta(A \otimes A) = \frac{d\|A\|_2^2 - |\,\mathrm{Tr}(A)|^2}{d^3 - d}. \qquad (5.14)
$$

Working in finite dimensions, we see that the frame condition (5.5) is achieved as long as $\alpha(A \otimes A^\dagger) > 0$, $\beta(A \otimes A^\dagger) > 0$, which means as long as

$$
d|\,\mathrm{Tr}(A)|^2 \geq \|A\|_2^2 \qquad (5.15)
$$

and

$$
|\,\mathrm{Tr}(A)|^2 < d\|A\|_2^2. \qquad (5.16)
$$

Let us assume that A is the state ρ_A, specifically the reduced state ρ_A of the probe. Then, the first inequality is automatically satisfied. Moreover, the Cauchy-Schwartz inequality implies that $|\,\mathrm{Tr}(A)|^2 \leq |\,\mathrm{Tr}(\mathbb{1})|\|A\|_2^2 = d\|A\|_2^2$, with equality if and only if $A \propto \mathbb{1}$. Having assumed that A is the state ρ_A, this is the condition that ρ_A is not maximally mixed.

Thus, we have found that, independently of the presence of an ancilla, as long as the reduced state ρ_A of the probe is not maximally mixed, we can find d^2 unitaries, one of which is the identity, such that $\{U_i \rho_A U_i^\dagger\}_{i=1}^n$ is a tomographically faithful set. If we assume that A is a state, specifically the reduced state ρ_A of the probe, this is the condition that ρ_A is not maximally mixed.

We can extend this result to the case where there are non-vanishing correlations. The operator A for which we want that $\{U_i A U_i^\dagger\}_{i=1}^{d^2}$ be a tomographically complete set can be taken to be any linear combination of the Hermitian operator Schmidt operators $\{A_l\}_l$ corresponding to non-zero operator Schmidt coefficients (one example being the reduced state ρ_A). Suppose ρ_{AB} is not product. Then there are at least two terms in its OSD, and at least one between one of the A_i's and the reduced state ρ_A is not proportional to the identity; we can then consider A in the above construction of the frame some linear combination of the latter two operators that respect conditions (5.15)–(5.16).

On the other hand, since a state of the form $\frac{\mathbb{1}_A}{d} \otimes \rho_B$ is invariant under local unitary transformation on A, we have proven our statement:

Theorem 5.2. *For all non product bipartite states and for all product states ρ_{AB} such that $\rho_A \neq \mathbb{1}_A/d$, there always exist d^2 unitary operators $U_i \in SU(d)$ such that the set $\{(U_i \otimes \mathrm{id})\rho_{AB}(U_i \otimes \mathrm{id})^\dagger\}_{i=1}^{d^2}$ is faithful.*

5.5 Pure probe–ancilla state

For pure states we are able to find the optimal number of local unitaries needed to construct a faithful set starting from a fixed pure state of Schmidt rank k:

$$|\psi\rangle\langle\psi|_{AB} = \sum_{i,j=1}^{k} \sqrt{p_i p_j}\, |i\rangle\langle j|_A \otimes |i\rangle\langle j|_B \,. \tag{5.17}$$

Theorem 5.3. *Let $|\psi\rangle\langle\psi|_{AB}$ be as in Eq. (5.17). Then, there are $n := \left\lceil \frac{d}{k} \right\rceil^2$ local unitaries U_i such that the set given by $\{(U_i \otimes \mathbb{1})\, |\psi\rangle\langle\psi|_{AB}\, (U_i \otimes \mathbb{1})^\dagger\}_{i=0}^{n-1}$, with $U_0 = \mathbb{1}$, is faithful.*

Proof. Let $|\psi_0\rangle = |\psi\rangle$. State tomography of the output $(\Lambda \otimes \mathrm{id})\,|\psi_0\rangle\langle\psi_0|$ determines the channel Λ partially, i.e., its action on $\{|i\rangle\langle j|\}$ only for $i, j = 1, \ldots, k$. To obtain the image under Λ of the remaining $|i\rangle\langle j|$ elements, it is convenient to consider the case when k divides d.

WE will start by analyzing how it is possible to reconstruct the action of Λ on all of $\{|i\rangle\langle j|\}$ for $i, j?1, \ldots, 2k$. Let us define the set of operators

$$A_{ij} = |i\rangle\langle j| \,, \tag{5.18}$$

$$B_{ij} = |i+k\rangle\langle j+k| \,, \tag{5.19}$$

$$C_{ij} = |i+k\rangle\langle j| \,, \tag{5.20}$$

$$D_{ij} = |i\rangle\langle j+k| \tag{5.21}$$

for $i, j = 1, \ldots, k$, and where sums within kets should be in general understood modulus d. Also, let us introduce unitary operators whose action restricted to the vectors $|n\rangle$, for $n = 1, \ldots, k$, is given by

$$\begin{aligned} X\,|n\rangle &= |n+k\rangle \,, \\ U\,|n\rangle &= 2^{-1/2}(|n\rangle + |n+k\rangle), \\ V\,|n\rangle &= 2^{-1/2}(|n\rangle + i\,|n+k\rangle). \end{aligned}$$

Acting locally on $|\psi\rangle\langle\psi|$, such operators produce the following states

$$\begin{aligned} |\psi_1\rangle\langle\psi_1| &:= (X \otimes \mathbb{1})\,|\psi\rangle\langle\psi|\,(X \otimes \mathbb{1})^\dagger \\ |\psi_2\rangle\langle\psi_2| &:= (U \otimes \mathbb{1})\,|\psi\rangle\langle\psi|\,(U \otimes \mathbb{1})^\dagger \\ |\psi_3\rangle\langle\psi_3| &:= (V \otimes \mathbb{1})\,|\psi\rangle\langle\psi|\,(V \otimes \mathbb{1})^\dagger. \end{aligned}$$

Define $\Lambda[Y] = [\Lambda[Y_{ij}]]_{i,j=1}^k$, for $X = A, B, C, D$. Then $\Lambda[A]$ is reconstructed through tomography of $(\Lambda \otimes \mathrm{id})\,|\psi_0\rangle\langle\psi_0|$ (as already noticed), while $\Lambda[B]$ is obtained from $(\Lambda \otimes \mathrm{id})\,|\psi_1\rangle\langle\psi_1|$. On the other hand, $\Lambda[C]$ and $\Lambda[D]$ can be reconstructed by measuring the four outputs (i.e. $(\Lambda \otimes \mathrm{id})\,|\psi_l\rangle\langle\psi_l|$ for $l = 0, \ldots, 3$) and then combining the results. To be more precise, since

$$C_{ij} = UA_{ij}U^\dagger + iVA_{ij}V^\dagger - \frac{1+i}{2}(A_{ij} + XA_{ij}X^\dagger) \tag{5.22}$$

$$D_{ij} = iVA_{ij}V^\dagger - UA_{ij}U^\dagger - \frac{i-1}{2}(A_{ij} + XA_{ij}X^\dagger), \tag{5.23}$$

linearity implies

$$\Lambda[C_{ij}] = \Lambda[UA_{ij}U^\dagger] + i\Lambda[VA_{ij}V^\dagger] - \frac{1+i}{2}(\Lambda[A_{ij}] + \Lambda[XA_{ij}X^\dagger]) \tag{5.24}$$
$$\Lambda[D_{ij}] = i\Lambda[VA_{ij}V^\dagger] - \Lambda[UA_{ij}U^\dagger] - \frac{i-1}{2}(\Lambda[A_{ij}] + \Lambda[XA_{ij}X^\dagger]). \tag{5.25}$$

Thus, we see that we have reconstructed $\Lambda[|i\rangle\langle j|]$ for $i,j = 1,\ldots,2k$ with four local unitaries.

Information on the remaining $\Lambda[|i\rangle\langle j|]$ can be reconstructed similarly. The theorem follows by reiterating this procedure, until recovering the action of Λ on all the blocks. More explicitly, it is possible to reconstruct $\Lambda[|i\rangle\langle j|]$ for $i,j \in \{1+p\cdot k,\ldots,k+p\cdot k\}$ and $p = 0,\ldots,d/k-1$ by considering the action of $p-$labelled d/k unitaries (one being the identity) each performing one of the transformations

$$|n\rangle \mapsto |n+p\cdot k\rangle\,. \tag{5.26}$$

Once these 'on-diagonal blocks' have been reconstructed, it is then possible to further reconstruct the 'off-diagonal blocks' $\Lambda[|i\rangle\langle j|]$ for $i \in \{1+p\cdot k,\ldots,k+p\cdot k\}$ and $j \in \{1+q\cdot k,\ldots k+q\cdot k\}$, $p \neq q$ by the use of $(d/k(d/k-1))/2$ pairs of unitaries that perform the transformations

$$|n\rangle = 2^{-1/2}(|n+p\cdot k\rangle + |n+q\cdot k\rangle) \tag{5.27}$$
$$|n\rangle = 2^{-1/2}(|n+p\cdot k\rangle + i\,|n+q\cdot k\rangle). \tag{5.28}$$

This gives a total of $d/k + 2\cdot(d/k(d/k-1))/2 = (d/k)^2$ unitaries.

If k does not divide exactly d, then one needs to consider an additional set of unitaries, but obviously the cost (in terms of unitaries) cannot be larger than in the case where we imagine the A system embedded in a d' dimensional system, with $d' = \lceil\frac{d}{k}\rceil\cdot k$. □

In the light of the last theorem we see that the higher the correlations (in terms of the OSR) of the fixed pure state, the less U_i are required. As expected, when the fixed pure state has maximal OSR, one recovers completely the AAPT scenario. For pure states with OSR $= 1$, the number of experimental settings to perform channel tomography is again the one of SPT. As a final remark we observe, by looking at the proof of Theorem 5.3, that one can derive the specific form of a particular set of U_i, besides establishing their existence.

Contrary to the pure state case, for the case where the fixed state is mixed, we have not derived a formula which directly links the OSR of the input to the number of unitaries needed to reach faithfulness. However, in the following we gives specif examples that show that also when the fixed state is mixed, the presence of correlations dramatically reduces the number of local unitaries required to perform channel tomography.

5.6 Mixed probe-ancilla state: qubit-qudit inputs

The first example involves a qubit-qudit system, for qubit channel tomography. We show that reducing the cardinality of the faithful set created by local unitaries on the qubit depends strongly on the quantumness of correlations on the qubit side. Before going into the details it is convenient to recall that a bipartite state is called classical on A if it can be expressed as $\rho_{AB} = \sum_i p_i \left|a_i\right\rangle\!\left\langle a_i\right|_A \otimes \rho_i^B$ for some orthonormal basis $\{|a_i\rangle\}$, and that states that are not classical on A are said to have non-zero quantum discord, cf. Sec. 1.2.2. Also, we will make use of the following Lemma, in which we use the notion of Bloch vector for a generic Hermitian operator $L = L^\dagger$, given by $\vec{l} = (l_1, l_2, l_3)$, with $l_i = \mathrm{Tr}(\sigma_i L)$ and σ_i, $i = 1, 2, 3$, the Pauli operators.

Lemma 5.1. *Consider Hermitian operators $A, B \in \mathcal{L}(\mathbb{C}^2)$. Then, A and B commute if and only if their Bloch vectors are proportional.*

Proof. Let $\sigma_0 = \mathbb{1}$ and denote $a_0 = \mathrm{Tr}(A)$, $b_0 = \mathrm{Tr}(B)$. Let also $\vec{a} = (a_1, a_2, a_3)$ and $\vec{b} = (b_1, b_2, b_3)$ be the Bloch vectors of A and B, respectively, so that $A = \frac{1}{2}\sum_{i=0}^{3} a_i \sigma_i$

and $B = \frac{1}{2}\sum_{i=0}^{3} b_i\sigma_i$. Observe that

$$\begin{aligned}
[A, B] &= \left[\frac{1}{2}\sum_{i=0}^{3} a_i\sigma_i, \frac{1}{2}\sum_{i=0}^{3} b_i\sigma_i\right] \\
&= \frac{1}{4}\sum_{i,j=0}^{3} a_i b_j \,[\sigma_i, \sigma_j] \\
&= \frac{1}{4}\sum_{i,j,k=1}^{3} a_i b_j \, 2i\epsilon_{ijk}\sigma_k \\
&= \frac{i}{2}\sum_{k=1}^{3}\left(\sum_{i,j=1}^{3} a_i b_j \epsilon_{ijk}\right)\sigma_k \\
&= \frac{i}{2}(\vec{a}\times\vec{b})\cdot\vec{\sigma},
\end{aligned} \tag{5.29}$$

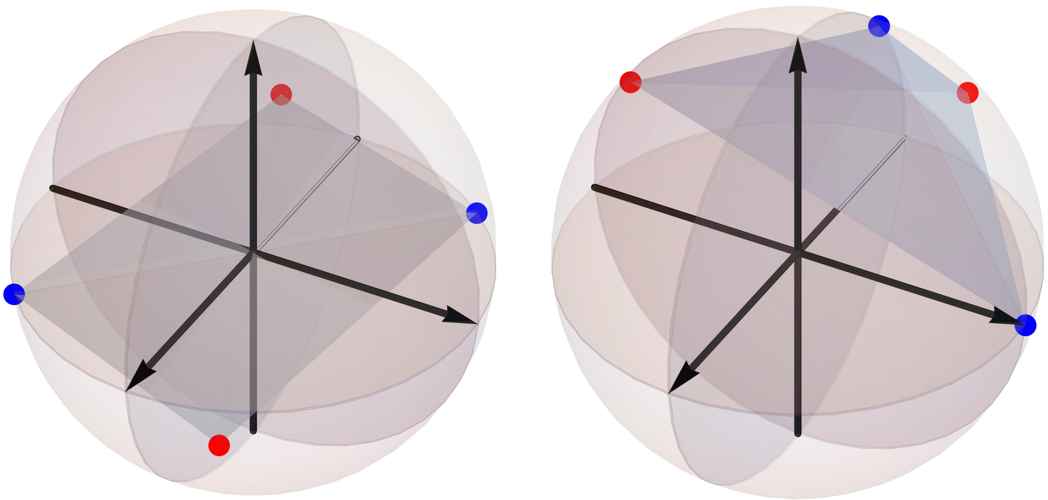
where we used the Levi-Civita symbol ϵ_{ijk}, and $\times$ indicates the standard cross product between three-dimensional vectors. Since $\sigma_1, \sigma_2, \sigma_3$ are linearly independent, the expression in the last line above is zero if and only if the cross product $\vec{a}\times\vec{b}$ vanishes, which happens if and only if $\vec{a} = \lambda\vec{b}$, with $\lambda \in \mathbb{R}$. □

We are now in the position to state as follows.

Theorem 5.4. *Let A be a qubit. Then, ρ_{AB} has quantum discord on A if and only if ρ_{AB} allows correlation-assisted process tomography on A with at most two unitary rotations.*

Proof. We recall that a qubit-qudit state has zero discord on the qubit side if and only if $\rho_{AB} = p\,|a_1\rangle\langle a_1|_A \otimes \rho_1^B + (1-p)\,|a_2\rangle\langle a_2|_A \otimes \rho_2^B$, with $\{|a_1\rangle, |a_2\rangle\}$ some orthonormal basis for A. While this is not necessarily the operator Schmidt decomposition, it is clear that the state only allows to reconstruct the action of a channel Λ on $\mathrm{span}(\{|a_1\rangle\langle a_1|, |a_2\rangle\langle a_2|\}) = \mathrm{span}(\{\mathbb{1}, |a_1\rangle\langle a_1|\})$. Overall, a single additional unitary rotation U allows us to reconstruct only the action of the same map on $\mathrm{span}(\{\mathbb{1}, |a_1\rangle\langle a_1|, U\,|a_1\rangle\langle a_1|\,U^\dagger\})$, which is not enough to tomographically reconstruct the channel. A geometric way of thinking about this is that the resulting four Bloch vectors are necessarily coplanar, and do not span affinely $\mathbb{R}^3$ (see Fig. 5.4a).

On the other hand, assume that ρ_{AB} has non-zero discord on A. This implies that there are some correlations, that is, that $\mathrm{OSR}(\rho_{AB}) \geq 2$. Without loss of generality, we can assume that the OSD of ρ must nontrivially contain A_1 and A_2 that do not commute, since, if all the non-trivial A_i's that appear in the OSD of ρ_{AB} commuted pair-wise, they would all commute, and there would not be any discord. From Lemma 5.1, the Bloch vectors of A_1 and A_2 are not collinear. This means that there is a rotation R of such vectors such that the resulting four vectors identify affinely independent points which span $\mathbb{R}^3$, (see Fig. 5.4b). Via the homomorphism between $SO(2)$ and $SU(3)$, the rotation R corresponds to unitary rotation U such that $\{\rho_{AB}, U_A \otimes \mathbb{1}_B \rho_{AB} U_A^\dagger \otimes \mathbb{1}_B\}$ is faithful.



(a) With no discord

(b) With discord

Figure 5.4: Bloch representation of two local operators for system A before (blue dots) and after (red dots) a local unitary rotation, for the case of a two-qubit state ρ_{AB}. (a) In the case of no discord, the blue dots correspond to the representation of two orthogonal pure states; red dots and blue dots are necessarily coplanar, independently of the unitary transformation, and hence do not span the entire three-dimensional (Bloch) space: channel tomography is not possible. (b) In the case with discord, the blue dots represent the (rescaled) Bloch component of two orthonormal (with respect to the Hilbert-Schmidt inner product) operators A_1 and A_2 that enter the OSD decomposition of ρ_{AB} not trivially, and that do not commute; there is a unitary such that red dots and blue dots are not coplanar, and hence span the entire three-dimensional space: channel tomography is possible.

□

5.7 Mixed probe–ancilla state: examples of efficient generation of faithful sets

In this example we present a family of mixed states of two qudits which generate faithful sets with even only two local unitaries, one being the identity. In order to construct the example we will make use of the Weyl (or generalized Pauli) basis for the space of $d \times d$ linear operators, which is given by $X^k Z^l$ with $k, l = 0, \ldots, d-1$, where $X = \sum_{p=0}^{d-1} |p+1\rangle\langle p|$, $Z = \sum_{q=0}^{d-1} \omega^q |q\rangle\langle q|$ and $\omega = e^{2\pi i/d}$ is a root of unity. Both X and Z are unitary, so that $X^\dagger = X^{-1}$ (similarly for Z). Since $X^d = Z^d = \mathbb{1}$, the sets $\{X^k\}_k$ and $\{Z^l\}_l$ form cyclic groups under multiplication, and we can think that the exponent is taken modulus d. Let $F = \frac{1}{\sqrt{d}} \sum_{k,l=0}^{d-1} w^{kl} |k\rangle\langle l|$ be the discrete Fourier transform unitary. One has $FXF^\dagger = Z$, $FZF^\dagger = X^\dagger$, and the braiding relation $ZX = \omega XZ$, from which one deduces $FX^kZ^lF^\dagger = Z^kX^{-l} = \omega^{-kl}X^{-l}Z^k$, and that the action of $F \cdot F^\dagger$ on the basis elements X^kZ^l induces closed and disjoint *orbits* within $\{X^kZ^l\}_{k,l}$ (up to irrelevant phases), defined as $O(k,l) = \{F^n X^k Z^l F^{\dagger n} \mid n = 0, \ldots, 3\}$. Specifically, one has

$$FX^kZ^lF^\dagger = \omega^{-kl}X^{-l}Z^k, \tag{5.30}$$

$$F^2X^kZ^lF^{\dagger 2} = X^{-k}Z^{-l}, \tag{5.31}$$

$$F^3X^kZ^lF^{\dagger 3} = \omega^{-kl}X^{l}Z^{-k}, \tag{5.32}$$

while $F^4 = \mathbb{1}$ so that obviously $F^4X^kZ^lF^{\dagger 4} = X^kZ^l$. Such orbits contains either one, two, or four distinct elements. The only orbits with a single element are the one including the identity, corresponding to $(k,l) = (0,0)$, for both d even and odd, and the one corresponding to the element $X^{d/2}Z^{d/2}$ for d even. No orbit can contain exactly three distinct elements, as this would require that one of such elements is invariant under $F \cdot F^\dagger$, which is a contradiction for an orbit that contains more than one element and that is known to close necessarily under four repeated actions of $F \cdot F^\dagger$.

Consider the set $\mathscr{O} = \{O(k,l)\}$ of orbits (up to irrelevant phases). We identify $O(k,l) = O(k',l')$ if X^kZ^l is in the same orbit as $X^{k'}Z^{l'}$ (up to irrelevant phases). Consider the set $\mathscr{P}^{(1)} = \{W^{(1)}(O)|O \in \mathscr{O}\}$ composed of one Weyl-operator representative $W^{(1)}(O)$ per orbit O (the exact choice of representative is irrelevant in this case). It is clear that $\cup_{i=0}^{3} F^i \mathscr{P}^{(1)} F^{\dagger i} = \{X^k Z^l\}_{k,l}$ up to irrelevant phases, where we have have used the shorthand notation $\Lambda\{K_m\}$ for the image $\{\Lambda[K_m]\}$ of a set $\{K_m\}$ under the action of a map Λ. In particular, $\mathrm{span}(\cup_{i=0}^{3} F^i \mathscr{P}^{(1)} F^{\dagger i}) = \mathrm{span}(\{X^k Z^l\}_{k,l})$. Furthermore, it is also clear that there is a choice of *pairs* of representatives (the two representatives may be chosen to coincide, in the case of 1-element and 2-element orbits) per orbit, forming a set $\mathscr{P}^{(2)} = \{W_1^{(2)}(O), W_2^{(2)}(O)|O \in \mathscr{O}\}$, such that $\cup_{i=0}^{1} F^i \mathscr{P}^{(2)} F^{\dagger i} = \{X^k Z^l\}_{k,l}$ (up to irrelevant phases).

Let us be more concrete, providing a specific choice for the set $\mathscr{P}^{(1)}$. We consider two cases, according to the parity of d. First, let $d = 2m+1$ be odd. It is easy to verify that (up to irrelevant phases) we can pick

$$\mathscr{P}^{(1)} = \{X^k Z^l \mid k = 0, \ldots, m;\ l = 1, \ldots, m\} \cup \{\mathbb{1}\}. \tag{5.33}$$

If $d = 2m$ is instead even, then we can choose (again, up to irrelevant phases)

$$\mathscr{P}^{(1)} = \{X^k Z^l \mid k = 0, \ldots, m-1;\ l = 1, \ldots, m\} \cup \{\mathbb{1}, X^m Z^m\}. \tag{5.34}$$

Notice the the above sets $\mathscr{P}^{(1)}$ contain $(d^2+3)/4$ and $(d^2+8)/4$ elements, for odd and even dimension respectively, thus scaling approximately as $d^2/4$.

One can similarly construct a set $\mathscr{P}^{(2)} \subset \{X^k Z^l\}_{k,l}$, with approximately $\approx d^2/2$ elements, such that $\mathscr{P}^{(2)} \cup F\mathscr{P}^{(2)}F^\dagger$ spans the entire operator space. For example, one can take $\mathscr{P}^{(2)} = \mathscr{P}^{(1)} \cup F\mathscr{P}^{(1)}F^\dagger$, where elements that are identical up to a phase factor are equated.

The issue we still have to face is how to use the facts above to construct, e.g., a state $\sigma_{AB} \in \mathcal{L}(\mathbb{C}^d \otimes \mathbb{C}^d)$ with $\mathrm{OSD}(\sigma_{AB}) \approx d^2/2$ such that $\{\sigma_{AB}, U_A \otimes \mathbb{1}_A \sigma_{AB} U_A^\dagger \otimes \mathbb{1}_A\}$ is a faithful set for channel tomography on qudit A. This is easily done by considering the Hermitian operators

$$H_{k,l} = \frac{1}{\sqrt{2}}\left(X^k Z^l + (X^k Z^l)^\dagger\right) = \frac{1}{\sqrt{2}}\left(X^k Z^l + \omega^{kl} F^2 (X^k Z^l) F^{\dagger 2}\right), \tag{5.35}$$

$$J_{k,l} = \frac{1}{i\sqrt{2}}\left(X^k Z^l - (X^k Z^l)^\dagger\right) = \frac{1}{i\sqrt{2}}\left(X^k Z^l - \omega^{kl} F^2 (X^k Z^l) F^{\dagger 2}\right) \tag{5.36}$$

where we have used the relation $F^2 X^k Z^l F^{\dagger 2} = X^{-k} Z^{-l}$ and the braiding relation of X and Z. The operators $H_{k,l}$ and $J_{k,l}$ obviously span the same operator subspace as $X^k Z^l$ and $F^2(X^k Z^l)F^{\dagger 2}$. Furthermore, we will use the fact that, given any Hermitian operator H in finite dimensions, there is $\epsilon > 0$ small enough (more precisely, it is enough that $\epsilon|\lambda_-(H)| \leq 1$, with $\lambda_-(H)$ the most negative eigenvalue of H) such that $\mathbb{1} + \epsilon H$ is positive semidefinite. Thus, for example, given our choice of $\mathscr{P}^{(1)}$ above, it is clear that, for, say, d odd (the even case is handled similarly), we can take

$$\sigma_{AB} \propto \frac{\mathbb{1}_A}{d} \otimes \frac{\mathbb{1}_B}{d} + \epsilon \sum_{k=0}^{(d-1)/2} \sum_{l=1}^{(d-1)/2} \left(H_{k,l} \otimes H_{k,l} + J_{k,l} \otimes J_{k,l}\right). \tag{5.37}$$

with $\epsilon > 0$ small enough. Then, by construction, $\left\{\sigma_{AB}\,,\,(F_A \otimes \mathbb{1}_B)\sigma_{AB}(F_A \otimes \mathbb{1}_B)^\dagger\right\}$ is a faithful set. Notice that σ_{AB} has OSR less or equal to $2\frac{d-1}{2}\left(\frac{d-1}{2}+1\right)+1 = \frac{d^2+1}{2}$.

5.8 Conclusion

In this chapter we have introduced and analyzed some properties of a framework for process tomography assisted by correlations. Our framework interpolates between standard process tomography and ancilla-assisted process tomography, and it is based on applying local transformations on the input probe—part of an probe-ancilla bipartite system—before the probe undergoes the process to be reconstructed. In particular, we focused on determining how the correlation properties of the starting probe-ancilla state ρ_{AB} affect such a number. We proved that essentially all correlations can be helpful, in the sense of reducing such a number from d^2 for standard process tomography to roughly $d^2/\mathrm{OSR}(\rho_{AB})$, with $\mathrm{OSR}(\rho_{AB})$ the operator Schmidt rank of ρ_{AB}, which is necessarily optimal. We proved that this holds true in the pure-input case even if the local transformations are restricted to be unitary. In the mixed-state case, we pointed out the role of discord in the case of qubit probes and unitary local transformations, and gave "extreme" examples where just one additional initial local unitary rotation suffices for process tomography, even if the initial state has operator Schmidt rank approximately $d^2/2$. It would be interesting to fully understand the mixed-state case for unitary local rotations, which appears to be related to studying and applying the adjoint representation of the unitary group, and will be investigated in future work.

Chapter 6

Metric and pseudometric spaces of quantum channels

In this chapter we collect a number of results about a possible generalisation of the Choi-Jamiołkowski isomorphism [52, 53, 57]. As it is clear from Chapters 4 and 5, the output of a channel acting on one part of a bipartite state having maximal OSR contains all the information about the channel itself, allowing e.g. process tomography. On the other hand, when the input states has OSR strictly less than maximal, the channel can be reconstructed only partially. As a consequence, there exist channels which cannot be distinguished if we probe their action only on a OSR-deficient input state (or on a non faithful set of inputs; cf. Chapter 4). In particular, we will show that for any channel Λ_1, Λ_2 with difference $\Delta := \Lambda_0 - \Lambda_1$ the quantity $\|(\Delta_A \otimes \mathrm{id}_B)[\rho_{AB}]\|_1$ defines a metric on the space of quantum channels when $OSR(\rho_{AB}) = d_A^2$, while it defines a *pseudometric* (to be defined later) when $OSR(\rho_{AB}) < d_A^2$. The latter give rise to pseudometric spaces which can be completed to ordinary metric spaces if we regard as equivalent all the channels which cannot be distinguished, e.g., only by looking at the outputs of the ancilla–assisted quantum process tomography (AAPT) scheme.

6.1 Metrics and pseudometrics

In order to articulate more rigorously what introduced so far, we need some preliminary definitions and results.

Definition 6.1 (Metric and pseudometric)**.** Let X be a set. A map $D : X \times X \to [0, +\infty)$ is called a pseudometric if, for any $x, y, z \in X$,

i) $D(x, x) = 0$

ii) $D(x, y) = D(y, x)$

iii) $D(x, z) \leq D(x, y) + D(y, z)$

Moreover, the map D defines a metric if the first condition is restricted to $D(x, y) = 0 \Leftrightarrow x = y$.

Theorem 6.1 (Every pseudometric induces a metric [140])**.** *Let X be a set a and $D : X \times X \to [0, +\infty)$ a pseudometric. For any $x, y \in X$, consider the equivalence relation $x \sim y$ if and only if $D(x, y) = 0$.*[1] *Denote by $[x] := \{y \in X | y \sim x\}$ the equivalence class of x under $\sim$ and by $\bar{X}$ the quotient space of X with respect to $\sim$. Then, the map $\bar{D} : \bar{X} \times \bar{X} \to [0, +\infty)$ defined as*

$$\bar{D}([x], [y]) = D(x, y) \tag{6.1}$$

is a metric, hence $(\bar{X}, \bar{D})$ a metric space.

Proposition 6.1 (State-dependent pseudometric)**.** *Let $D(\rho, \sigma)$ be any metric on the space of quantum states. In particular, we are assuming that $D(\rho, \sigma)$ is the metric induced by the one–norm $D(\rho, \sigma) := \|\rho - \sigma\|_1$, but the construction can be generalized to any metric on the space of (bipartite) states. Then the mapping*

$$D_{\rho_{AB}}(\Lambda_1, \Lambda_2) := D(J_{\rho_{AB}}(\Lambda_1), J_{\rho_{AB}}(\Lambda_2)) \tag{6.2}$$

[1]Notice that the relation $\sim$ introduced in the definition is a proper equivalence relation, i.e. it is reflexive, symmetric and transitive.

with

$$J_{\rho_{AB}}(\Lambda) := (\Lambda \otimes \mathrm{id})[\rho_{AB}]. \tag{6.3}$$

is a pseudometric, and it is a metric if and only if $\mathrm{OSR}(\rho) = d_A^2$.

Proof. Let us start by considering a bipartite state $\rho_{AB} \in \mathcal{S}_{AB}$ with $\mathrm{OSR}(\rho_{AB}) < d^2$, where $d := \min\{d_A, d_B\}$. It is easy to see that the properties satisfied by the $1-$norm are inherited by the induced map $D_{\rho_{AB}}(\Lambda_1, \Lambda_2)$, which in turn is nonnegative, satisfies the triangle inequality and, by direct inspection, it is zero when $\Lambda_1 \equiv \Lambda_2$. To convince ourselves that $D_{\rho_{AB}}$ is only a pseudometric when ρ_{AB} is OSR–deficient it remains to prove that there are at least two channels Γ_1, Γ_2 such that $\Gamma_1 \neq \Gamma_2$ but $D_{\rho_{AB}}(\Gamma_1, \Gamma_2)$ vanishes. To see that, assume without loss of generality that $\mathrm{OSR}(\rho_{AB}) = d^2 - 1$, i.e.

$$\sum_{i=1}^{d^2-1} r_i A_i \otimes B_i. \tag{6.4}$$

Moreover, similarly to what we have defined in Eq. (4.19), let us consider the following maps:

$$\Gamma_i[X] = \mathrm{Tr}(X)\frac{\mathbb{1}}{d} + \epsilon\,\mathrm{Tr}(A_d X) Y_i, \quad i = 0, 1, \tag{6.5}$$

where Y_0, Y_1 are traceless Hermitian operators with $Y_0 \neq Y_2$.[2] To conclude, it is enough to see that the two maps act the same on the first $d^2 - 1$ local operators of ρ_{AB}, while they produce different outputs only when acting on A_{d^2}, i.e. on the local operator corresponding to the only vanishing OSC of ρ_{AB}. Then it is easy to calculate

$$\begin{aligned} D_{\rho_{AB}}(\Gamma_0, \Gamma_1) &= \|(\Gamma_0 - \Gamma_1)_A \otimes \mathrm{id}_B[\rho_{AB}]\|_1 \\ &= r_{d^2} \|(\Gamma_0 - \Gamma_1)[A_{d^2}] \otimes B_{d^2}\|_1 \\ &= 0 \end{aligned} \tag{6.6}$$

Observe that in general two maps which differ only by their action on those subspaces spanned by local channels corresponding to null coefficients cannot be distinguished.

[2]We recall that the maps defined through Eq. (6.5) are trace-preserving by construction, and completely positive for ϵ small enough, e.g. for $\epsilon \leq (d\|A_{d^2}\|_\infty \| \max\{\|Y_0\|_\infty, \|Y_1\|_\infty\})^{-1}$.

On the other hand, if ρ_{AB} has maximal OSR, one has $\|\Gamma_0[A_i] - \Gamma_1[A_i]\|_1 \neq 0$ for at least one $i = 1, \ldots, d^2$ (in this case $i = d^2$), since the two maps are different by assumption. $\square$

Remark. *Notice that when the input state in Proposition 6.1 is maximally entangled, one recovers the so-called Jamiołkowski distance [18, 141].*

6.2 Equivalence classes of quantum channels

In order to obtain a proper metric starting from the pseudometric induced by OSR–deficient states, we introduce the following equivalence classes:

Definition 6.2. For any ρ_{AB} and any $\Lambda \in T(H)$ we define the equivalence classes

$$[\Lambda]_{\rho_{AB}} := \{\Gamma \in T(H) \mid D_{\rho_{AB}}(\Lambda, \Gamma) := \|(\Lambda - \Gamma)_A \otimes \mathrm{id}_B\ [\rho_{AB}]\|_1 = 0\}, \tag{6.7}$$

i.e. the set of all channels which, under the pseudometric induced by ρ_{AB}, are at zero distance from Λ).

In the following we provide a series of results characterising the equivalence classes just introduced.

Proposition 6.2. *Let ρ_{AB}, σ_{AB} be two bipartite quantum states with $\alpha := \mathrm{OSR}(\rho_{AB}) \leq \mathrm{OSR}(\sigma_{AB}) =: \beta$ and OSD given, respectively, by*

$$\rho_{AB} = \sum_{i=1}^{\alpha} r_i A_i \otimes B_i, \quad \sigma_{AB} = \sum_{i=1}^{\beta} s_i A_i \otimes C_i, \tag{6.8}$$

i.e. the local basis of ρ_{AB} on A equals the first α elements of the local basis of σ_{AB} on A. Then $[\Lambda]_{\sigma_{AB}} \subseteq [\Lambda]_{\rho_{AB}}$ for any Λ. Moreover, the inclusion is strict if $\alpha < \beta$.

Proof. Let Λ_1, Λ_2 be two arbitrary channels such that $\Lambda_2 \in [\Lambda_1]_{\sigma_{AB}}$ and denote $\Delta := \Lambda_1 - \Lambda_2$. Then, since

$$0 = \|\Delta \otimes \mathrm{id}[\sigma_{AB}]\|_1$$

$$
\begin{aligned}
&\geq \|\Delta \otimes \mathrm{id}[\sigma_{AB}]\|_2 \\
&= \left\| \sum_{i=1}^{\beta} s_i \Delta(A_i) \otimes C_i \right\|_2 \\
&= \left(\sum_{i=1}^{\beta} s_i^2 \|\Delta(A_i)\|_2^2 \right)^{1/2} ,
\end{aligned} \tag{6.9}
$$

and since the terms in the last parenthesis are non-negative, one concludes that all the $\|\Delta(A_i)\|_2$ must be zero. This fact implies that all the singular values of $\Delta(A_i)$, for any $i \in \{1, \dots, \beta\}$, must be zero. Moreover, from the monotonicity of the Schatten p-norms follows that also $\|\Delta(A_i)\|_1 = 0$ for any $1 \leq i \leq \alpha \leq \beta$, hence

$$
\begin{aligned}
\|\Delta \otimes \mathrm{id}[\rho_{AB}]\|_1 &= \left\| \sum_{i=1}^{\alpha} r_i \Delta(A_i) \otimes B_i \right\|_1 \\
&\leq \sum_{i=1}^{\alpha} r_i \|\Delta(A_i)\|_1 \|B_i\|_1 = 0.
\end{aligned} \tag{6.10}
$$

We have proved that if $\Lambda_2 \in [\Lambda_1]_{\sigma_{AB}}$ then $\Lambda_2 \in [\Lambda_1]_{\rho_{AB}}$, so that $[\Lambda_1]_{\sigma_{AB}} \subseteq [\Lambda_1]_{\rho_{AB}}$.

To prove that the inclusion is strict, we show that $\exists\, \Gamma_1, \Gamma_2$ such that $\Gamma_1 \in [\Gamma_2]_{\rho_{AB}}$ but $\Gamma_1 \neq [\Gamma_2]_{\sigma_{AB}}$. Let

$$
\Gamma_i[X] = \mathrm{Tr}(X)\frac{\mathbb{1}}{d_A} + \epsilon\, \mathrm{Tr}(A_\beta X) Y_i \quad (i = 1, 2), \tag{6.11}
$$

where $\epsilon > 0$ is small enough in order for $\Gamma_i[\,\cdot\,]$ to be a channel, and Y_i are traceless Hermitian operators (see the proof of Proposition 6.1). Define, as $\Xi := \Gamma_1 - \Gamma_2$ and observe that

$$
D_{\sigma_{AB}}(\Gamma_1 \Gamma_2) = \|\Xi \otimes \mathrm{id}\sigma_{AB}\|_1 = \|Y_1 - Y_2\| \|C_\beta\| \epsilon s_\beta > 0, \tag{6.12}
$$

while

$$
D_{\rho_{AB}}(\Gamma_1, \Gamma_2) = \|\Xi \otimes \mathrm{id}[\rho_{AB}]\|_1 = 0. \tag{6.13}
$$

□

Proposition 6.3. *There are states which are not comparable, in the sense that there exist ρ_{AB}, σ_{AB} and channels Λ_1, Λ_2 which are distinguished by $D_{\rho_{AB}}$ but not by $D_{\sigma_{AB}}$, and there are channels Γ_1, Γ_2 such that the opposite holds.*

Proof. The possible states and the channels that prove the claim are the following:

$$\rho_{AB} = |0\rangle\langle 0| \otimes |1\rangle\langle 1| \tag{6.14}$$
$$\sigma_{AB} = |x_+\rangle\langle x_+| \otimes |x_+\rangle\langle x_+| \tag{6.15}$$
$$\Lambda_1[X] = \sigma_x X \sigma_x \tag{6.16}$$
$$\Lambda_2[X] = \sigma_y X \sigma_y \tag{6.17}$$
$$\Gamma_1[X] = \sigma_y X \sigma_y \tag{6.18}$$
$$\Gamma_2[X] = \sigma_z X \sigma_z, \tag{6.19}$$

where $|x_+\rangle := 2^{-1/2}(|0\rangle + |1\rangle)$ and σ_i are the usual Pauli matrices. Then, it is straightforward to see that

$$\begin{aligned} D_{\rho_{AB}}(\Lambda_1, \Lambda_2) &= \|(\sigma_x |0\rangle\langle 0| \sigma_x - \sigma_y |0\rangle\langle 0| \sigma_y) \otimes |0\rangle\langle 0| \|_1 \\ &= \|(|1\rangle\langle 1| - |1\rangle\langle 1|) \otimes |0\rangle\langle 0| \|_1 = 0 \end{aligned} \tag{6.20}$$
$$\begin{aligned} D_{\sigma_{AB}}(\Lambda_1, \Lambda_2) &= \|(\sigma_x |x_+\rangle\langle x_+| \sigma_x - \sigma_y |x_+\rangle\langle x_+| \sigma_y) \otimes |x_+\rangle\langle x_+| \|_1 \\ &= \|(|x_+\rangle\langle x_+| - |x_-\rangle\langle x_-|) \otimes |x_+\rangle\langle x_+| \|_1 = 2 \end{aligned} \tag{6.21}$$
$$\begin{aligned} D_{\rho_{AB}}(\Gamma_1, \Gamma_2) &= \|(\sigma_y |0\rangle\langle 0| \sigma_y - \sigma_z |0\rangle\langle 0| \sigma_z) \otimes |0\rangle\langle 0| \|_1 \\ &= \|(|1\rangle\langle 1| - |0\rangle\langle 0|) \otimes |0\rangle\langle 0| \|_1 = 2 \end{aligned} \tag{6.22}$$
$$\begin{aligned} D_{\sigma_{AB}}(\Gamma_1, \Gamma_2) &= \|(\sigma_y |x_+\rangle\langle x_+| \sigma_y - \sigma_z |x_+\rangle\langle x_+| \sigma_z) \otimes |x_+\rangle\langle x_+| \|_1 \\ &= \|(|x_-\rangle\langle x_-| - |x_-\rangle\langle x_-|) \otimes |x_+\rangle\langle x_+| \|_1 = 0. \end{aligned} \tag{6.23}$$

□

In the next two propositions we provide two results proving some properties satisfied by all the equivalence classes.

Proposition 6.4. *Given two different channels Λ_1, Λ_2, the distance between arbitrary* *elements of the classes (generated by the same input state)* $[\Lambda_1]_{\rho_{AB}}, [\Lambda_2]_{\rho_{AB}}$ *is constant and equals the distance between representatives:* $D_{\rho_{AB}}(\bar{\Gamma}_1, \bar{\Gamma}_2) = D_{\rho_{AB}}(\Lambda_1, \Lambda_2)$.

Proof. We first prove the $\leq$. Let $\bar{\Gamma}_1 \in [\Lambda_1]_{\rho_{AB}} := \{\Gamma_1 \mid D_{\rho_{AB}}(\Lambda_1, \Gamma_1) = 0\}$ and $\bar{\Gamma}_2 \in [\Lambda_2]_{\rho_{AB}} := \{\Gamma_2 \mid D_{\rho_{AB}}(\Lambda_2, \Gamma_2) = 0\}$; then, using the triangle inequality, $D\rho_{AB}(\bar{\Gamma}_1, \bar{\Gamma}_2) \leq D_{\rho_{AB}}(\bar{\Gamma}_1, \Lambda_1) + D_{\rho_{AB}}(\Lambda_1, \Lambda_2) + D_{\rho_{AB}}(\Lambda_2, \bar{\Gamma}_2) = D_{\rho_{AB}}(\Lambda_1, \Lambda_2)$. Finally, since to say that $\bar{\Gamma}_i \in [\Lambda_i]_{\rho_{AB}}$ is equivalent to say that $\Lambda_i \in [\bar{\Gamma}_i]_{\rho_{AB}}$, for $i = 1, 2$, we can invert the previous inequality and obtain the desired result. $\square$

Proposition 6.5. *For any channel* Λ *and any input state* ρ_{AB}*,* $[\Lambda]_{\rho_{AB}}$ *is a convex set, i.e.* $\forall\, \Lambda_1, \Lambda_2 \in [\Lambda]_{\rho_{AB}}, \forall\, t \in [0,1]$*, the channel* $\Lambda_c := t\Lambda_1 + (1-t)\Lambda_2 \in [\Lambda]_{\rho_{AB}}$ *(equivalently,* $\exists\, \Gamma \in [\Lambda]_{\rho_{AB}}$ *s.t.* $D_{\rho_{AB}}(\Lambda_c, \Gamma) = 0$*).*

Proof. Let $\Gamma := \Lambda_1 \in [\Lambda]_{\rho_{AB}}$; then $D_{\rho_{AB}}(\Lambda_c, \Gamma) = \|(t\Lambda_1 + (1-t)\Lambda_2 - \Lambda_1) \otimes \mathrm{id}[\rho_{AB}]\|_1 = (t-1)\|(\Lambda_1 - \Lambda_2) \otimes \mathrm{id}[\rho_{AB}]\|_1 = 0$, by hypothesis. $\square$

In the light of the previous result we can now state and prove our main theorem about the characterisation of the equivalence classes of quantum channels.

Theorem 6.2 (Characterisation of equivalence classes)**.** *Let*

$$\rho_{AB} = \sum_{i=1}^{OSR(\rho_{AB})} r_i A_i \otimes B_i \tag{6.24}$$

$$\sigma_{AB} = \sum_{i=1}^{OSR(\sigma_{AB})} s_i C_i \otimes D_i \tag{6.25}$$

and

$$L_{A,\rho_{AB}} := \mathrm{span}\{A_i \mid i = 1, \ldots, OSR(\rho_{AB})\} \tag{6.26}$$

$$L_{A,\sigma_{AB}} := \mathrm{span}\{C_i \mid i = 1, \ldots, OSR(\sigma_{AB})\}. \tag{6.27}$$

Moreover, we assume without loss of generality that $\mathrm{OSR}(\rho_{AB}) \leq \mathrm{OSR}(\sigma_{AB})$*. Then,*

$$\forall\, \Lambda \in \mathcal{C}(\mathcal{H}_A, \mathcal{H}_A),\ [\Lambda]_{\rho_{AB}} \equiv [\Lambda]_{\sigma_{AB}} \ \Leftrightarrow\ L_{A,\rho_{AB}} \equiv L_{A,\sigma_{AB}}. \tag{6.28}$$

Proof. We first prove sufficiency. Notice that this is equivalent to prove

$$L_{A,\rho_{AB}} \neq L_{A,\sigma_{AB}} \;\Rightarrow\; \exists\,\bar{\Lambda} \in T(H) \text{ s.t. } [\bar{\Lambda}]_{\rho_{AB}} \neq [\bar{\Lambda}]_{\sigma_{AB}}. \tag{6.29}$$

The assumption $L_{A,\rho_{AB}} \neq L_{A,\sigma_{AB}}$ implies that there exists at least one C_j in the OSD of σ_{AB}, i.e. with $j \leq OSR(\sigma_{AB})$, which cannot be expressed as a linear combination of the A_i defining $L_{A,\rho_{AB}}$ only. That is (recall that $\{A_i\}$ is a basis),

$$C_j = \underbrace{\sum_{i=1}^{OSR(\rho_{AB})} \langle A_i, C_j\rangle A_i}_{=:C_j^{\parallel}} + \underbrace{\sum_{i=OSR(\rho_{AB})+1}^{d^2} \langle A_i, C_j\rangle A_i}_{=:C_j^{\perp}} \tag{6.30}$$

with $C_j^{\perp} \neq 0$. Then, there must exists an A_k with $k \in \{OSR(\rho_{AB})+1, \ldots, d^2\}$ such that $\langle A_k, C_j\rangle \neq 0$, i.e. $\mathrm{Tr}(A_k, C_j) \neq 0$ (assuming hermiticity). Now define

$$\Lambda_i[X] = \mathrm{Tr}(X)\frac{\mathbb{1}}{d_A} + \epsilon\,\mathrm{Tr}(A_k X) Y_i \quad (i = 1, 2), \tag{6.31}$$

which gives $\Delta[X] := (\Lambda_0 - \Lambda_1)[X] = \epsilon(Y_0 - Y_1)\,\mathrm{Tr}(A_k, X)$, and notice that

$$\begin{aligned}
D_{\rho_{AB}}(\Lambda_1, \Lambda_2) = \|\Delta \otimes \mathrm{id}\rho_{AB}\|_1 &= \left\| \sum_{i=1}^{OSR(\rho_{AB})} \Delta(A_i) \otimes r_i B_i \right\|_1 \\
&= \epsilon \|Y_0 - Y_1\|_1 \left\| \sum_{i=1}^{OSR(\rho_{AB})} \mathrm{Tr}(A_k A_i) r_i B_i \right\|_1 \\
&= \epsilon \|Y_0 - Y_1\|_1 \left\| \sum_{i=1}^{OSR(\rho_{AB})} \delta_{k,i}\, r_i B_i \right\|_1 = 0,
\end{aligned} \tag{6.32}$$

because $\delta_{k,i} = 0$ for any $i < OSR(\rho_{AB}) + 1$. On the other hand,

$$\begin{aligned}
D_{\sigma_{AB}}(\Lambda_1, \Lambda_2) = \|\Delta \otimes \mathrm{id}\sigma_{AB}\|_1 &= \left\| \sum_{i=1}^{OSR(\sigma_{AB})} \Delta(C_i) \otimes r_i D_i \right\|_1 \\
&= \epsilon \|Y_0 - Y_1\|_1 \left\| \sum_{i=1}^{OSR(\sigma_{AB})} \mathrm{Tr}(A_k C_i) r_i D_i \right\|_1
\end{aligned}$$

$$
\begin{aligned}
&\geq \epsilon \|Y_0 - Y_1\|_1 \left\| \sum_{i=1}^{OSR(\sigma_{AB})} \mathrm{Tr}(A_k C_i) r_i D_i \right\|_2 \\
&> 0,
\end{aligned}
\tag{6.33}
$$

where we have used the fact that the $r_i D_i$ are linear independent and, for at least one i, there is a nonzero coefficient $\mathrm{Tr}(A_k C_i)$ by hypothesis.

To prove necessity, one can observe that if the two spans generate the same subspace, then each C_i in the OSD of σ_{AB} can be expressed as a function of the first A_i, i.e. with $i = 1, \dots, OSR(\rho_{AB})$. Assume $\Lambda_1 \in [\Lambda_0]_{\rho_{AB}}$ and define $\Delta := \Lambda_0 - \Lambda_1$. If $D_{\rho_{AB}}(\Lambda_0, \Lambda_1) = 0$, i.e. if $\Delta(A_i) = 0$ for any $1 \leq i \leq OSR(\rho_{AB})$, then

$$
\begin{aligned}
D_{\sigma_{AB}}(\Lambda_0, \Lambda_1) = \|\Delta \otimes \mathrm{id}[\sigma_{AB}]\|_1 &= \left\| \sum_{i=1}^{OSR(\sigma_{AB})} \Delta(C_i) \otimes r_i D_i \right\|_1 \\
&= \left\| \sum_{i=1}^{OSR(\sigma_{AB})} \Delta \left(\sum_{j=1}^{OSR(\rho_{AB})} Tr(A_j C_i) A_j \right) \otimes r_i D_i \right\|_1 \\
&= \left\| \sum_{i=1}^{OSR(\sigma_{AB})} \sum_{j=1}^{OSR(\rho_{AB})} Tr(A_j C_i) \Delta(A_j) \otimes r_i D_i \right\|_1 = 0,
\end{aligned}
\tag{6.34}
$$

that is $[\Lambda_0]_{\rho_{AB}} \subseteq [\Lambda_0]_{\sigma_{AB}}$. To obtain the reverse inclusion and conclude the proof, it is enough to exploit the hypothesis in order to express the A_i as a function of the first C_i, i.e. $i \in \{1, \dots, \mathrm{OSR}(\sigma_{AB})\}$. □

6.3 Conclusion

The theory introduced in this chapter is preliminary and part of an ongoing research project [4]. The essence of the matter discussed hitherto is that the Jamiołkowski distance for quantum channels [18, 141] can be actually seen as a particular instance of a more general mapping, the latter giving a whole family of state-dependent metric functions. In the spirit of this entire thesis, another central result was to show how the metric space of quantum channels induced by a given quantum state can be identified with the operator Schmidt decomposition of the latter, specifically, by the span of its operator Schmidt local operators. These result are manifestly theoretical, but they do have an appealing practical consequence. Indeed, the content of this chapter can be employed in order to show that the discrimination of certain pairs of quantum channels can be facilitated by tailoring the input in an ancilla-assisted quantum discrimination protocol (as the AAPT of Chapter 4), where such tailoring is controlled according to the OSD of the input itself.[3]

[3] For a fully developed discussion on this result, we refer the reader to our soon-to-be submitted paper [4].

Conclusion

In this thesis we have examined in details the operator Schmidt decomposition (OSD) of bipartite operators. We have shown that the operator Schmidt rank (OSR), as well as the operator Schmidt coefficients (OSC) of quantum states and of its square root, can be exploited in order to devise measures of correlations and of total correlations. In particular, in Chapter 3 we have defined a measure of correlations and several measures of total correlations, we have analysed their properties and their application to entanglement and steering detection. Moreover, we have provided a relational expression between two of our measures of total correlations (which have been defined by taking advantage of the square root of quantum states and its OSD) and the quantum mutual information. However, our analysis leaves some unanswered questions, like, for example, the conjectured average monotonicity of the operator G–Concurrence from Sec. 3.4. Another open problem is whether the $l^{th}-$order elementary symmetric polynomial (ESN) in the OSC are monotonic for any $l \neq 1$ (cf. Sec. 3.4.2). Since the ESN are Schur–convex functions of the OSC, the monotonicity of the latter in the sense of Eq. (3.6), for OSR $\neq 1$, would be enough in order for the respective ESN to define a family of meaningful measures of correlations. Lastly, since the OSD of the square root of a quantum state looks like carrying valuable information about the degree of correlations of the original state (as seen in Sec. 3.5 and 3.6), it would be interesting to look for an explicit relationship between the OSD of a state and of its square root, and specifically between the two sets of operator Schmidt coefficients. Chapter 3 in its entirety, together with a possible answer to the open questions outlined above, might constitute the content of a future scientific publication.

The second part of the thesis has a more practical connotation, albeit still of theoretical nature. Here we have discussed the role played by the OSD in the tasks of channel discrimination and tomography. In particular, in Chapter 4 we have examined the scheme of ancilla–assisted quantum process discrimination and defined a worst–case scenario quantifier for the performance of the bipartite input states in such protocol: the channel discrimination power (CDP). We have computed general upper and lower bounds to the CDP of a state in terms of its OSD, and we have shown that the CDP of a pure state corresponds to its smallest OSC. Moreover, we have provided a lower bound on the CDP of any state that passes the realignment criterion for separability. The central observation of this chapter was to notice that also certain correlated but separable states can have non-zero CDP, as long as they posses a certain amount of discord. The main open problem remains to compute the CDP of a general mixed state, that we conjecture corresponds again to the smallest OSC. The general bounds to the CDP of mixed states given in Sec. 4.4 do not falsify such conjecture, making room for a possible generalization of the exact result obtained in the pure state scenario. The tasks of channel discrimination and channel tomography are closely related. For this reason some of the observations and results from Chapter 4 readily transition to Chapter 5, where we have proved that the correlations of a fixed bipartite state measured by the logarithm of the OSR can be employed to allow process tomography. We have shown that any single bipartite input state can be used to perform ancilla–assisted channel tomography. In fact, only states with non maximal OSR are inadequate to perform AAPT. Nevertheless, they can be transformed through local operations, in order to obtain a set of states which altogether provide a full tomographic reconstruction of any given channel. We have also considered the particular case of quantum channels given by local unitaries, and shown that the number of unitary operators allowing channel tomography depends upon the correlations of the initial input state. For pure states this number scales as the inverse of the OSR: the more correlated the input, the less unitaries are needed. For mixed states instead, examples showing how the presence of correlations efficiently reduces the number of local unitaries were provided, but the optimal number of such unitaries remains unknown, offering a hint for further research.

The analysis of the relationship between the OSD and the tasks of ancilla–assisted quantum process discrimination and tomography have led us to some interesting observations of purely theoretical character, which were summarised in Chapter 6. We have observed that any operator with maximal OSR gives rise to a state–dependent metric on the space of quantum channels, while OSR-deficient states induces pseudometrics. The latter generate pseudometric spaces of channels, which can be lifted to proper metric spaces by taking the quotient of the space of quantum channels with respect to a specific equivalence relation. Our contribution was to show that two equivalence classes of quantum channels (induced by two different bipartite states) are equal if and only if the spans of the local operators in the OSD of the two states coincide. The results of Chapter 6 are part of an ongoing research project that is unfolding on the application side as well, returning promising results which can be employed, again, in the context of channel discrimination, giving means by which improving the distinguishability of certain pairs of quantum maps [4].

In conclusion, this work validated our initial assumptions about the value of the OSD in the quantification and detection of classical and quantum correlations. This naturally brought us to face the more pragmatic side of the coin, with the analysis of the tasks of channel discrimination and tomography. Finally, we have initiated a geometrical theory of metric spaces of quantum channels which, albeit preliminary, endowed this research project with a renewed, wholehearted strength. What remains, together with an investigation of the open questions arisen during the last three years, is to study any potential usefulness of the OSD for the implementation of other specific tasks in the context, for example, of quantum communication and cryptography. Also, the closeness of the topics studied hitherto with the field of quantum metrology deserves a try, as well as the evident connection of our results with the study and the evaluation of correlations in quantum many-body systems.

All in all, by addressing several problems which may be of interests to the specialists of the subject, this thesis may be considered as a first step towards a deeper understanding of the role played by the OSD in Quantum Physics, but most importantly it hatched new, unforeseen questions, suggesting several routes for the continuation of the present research project.

Appendix A

Proof of Theorem 2.6

For the reader convenience, we first restate the theorem in question.

Theorem 2.6. *Let $\rho_{AB} \in \mathcal{S}_{AB}$ and consider, for $n \geq 1$, the set of $2\,n$ jointly linear (or jointly antilinear) superoperators*

$$\Gamma_i^A : \mathcal{H}_A \to \mathcal{H}_A, \quad \Gamma_i^B : \mathcal{H}_B \to \mathcal{H}_B, \quad \textit{for } i = 1, \ldots, n \tag{A.1}$$

such that for some $\epsilon_A, \epsilon_B \geq 0$ and for any $\sigma_A^i \in \mathcal{S}_A$ and $\sigma_B^i \in \mathcal{S}_B$, with $i = 1, \ldots, n$, one has

$$\sum_{i=1}^{n} \left\| \Gamma_i^A[\sigma_A^i] \right\|_2^2 \leq n\,\epsilon_A \quad \textit{and} \quad \sum_{i=1}^{n} \left\| \Gamma_i^B[\sigma_B^i] \right\|_2^2 \leq n\,\epsilon_B. \tag{A.2}$$

Moreover, define the following linear operator on $\mathcal{S}_{AB}$:

$$\rho_{AB}\left(\Gamma_{1,\ldots,n}^{(A,B)}\right) := \frac{1}{n}\left(\sum_{i=1}^{n} \Gamma_i^A \otimes \Gamma_i^B[\rho_{AB}] + \sum_{i \neq j}^{n} \Gamma_i^A \otimes \Gamma_j^B[\rho_A \otimes \rho_B],\right) \tag{A.3}$$

where ρ_A, ρ_B are the reduced density matrices of ρ_{AB}. If ρ_{AB} is separable, then

$$\begin{aligned} \left\| \left[\rho_{AB}\left(\Gamma_{1,\ldots,n}^{(A,B)}\right)\right]^R \right\|_1 \leq &\left(\epsilon_A + \frac{1}{n}\sum_{i<j}\left(\langle \Gamma_i^A[\rho_A], \Gamma_j^A[\rho_A]\rangle + c.c.\right)\right)^{1/2} \\ &\times \left(\epsilon_B + \frac{1}{n}\sum_{i<j}\left(\langle \Gamma_i^B[\rho_B], \Gamma_j^B[\rho_B]\rangle + c.c.\right)\right)^{1/2}, \end{aligned} \tag{A.4}$$

where c.c. stands for complex conjugate and $\langle \cdot\,,\cdot\rangle$ *is the HS inner product in* $\mathcal{L}(\mathcal{H}_{A(B)})$.

Proof. Let $\rho_{AB} \in \mathcal{S}_{AB}$ be a separable bipartite state with decomposition $\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i$. Then $\rho_{A(B)} = \sum_i p_i \rho_{A(B)}^i$. For notational convenience here we prove the theorem for the operator $\rho_{AB}\left(\Gamma_{1,2}^{(A,B)}\right)$; the extension to $\rho_{AB}\left(\Gamma_{1,\ldots,n}^{(A,B)}\right)$ is straightforward. Let us start by observing that

$$\begin{aligned}\left\|\left[\rho_{AB}\left(\Gamma_{1,\ldots,n}^{(A,B)}\right)\right]^R\right\|_1 &= \frac{1}{2}\left\|\sum_{i,j} p_i p_j \left[(\Gamma_1^A[\rho_A^i] + \Gamma_2^A[\rho_A^j]) \otimes (\Gamma_1^B[\rho_B^i] + \Gamma_2^B[\rho_B^j])\right]^R\right\|_1 \\ &\leq \frac{1}{2}\sum_{i,j} p_i p_j \left\|\left[(\Gamma_1^A[\rho_A^i] + \Gamma_2^A[\rho_A^j]) \otimes (\Gamma_1^B[\rho_B^i] + \Gamma_2^B[\rho_B^j])\right]^R\right\|_1 .\end{aligned} \tag{A.5}$$

where the second follows from the triangle inequality. Now notice that the linear operator $\tilde{\rho} := (\Gamma_1^A[\rho_A^i] + \Gamma_2^A[\rho_A^j]) \otimes (\Gamma_1^B[\rho_B^i] + \Gamma_2^B[\rho_B^j])$ has a single product term, hence a single nonzero OSC $\lambda_1 = \|\tilde{\rho}\|_2$, cf. Proposition 2.1. It follows that

$$\begin{aligned}\|\tilde{\rho}^R\|_1 &= \left\|\Gamma_1^A[\rho_A^i] + \Gamma_2^A[\rho_A^j]\right\|_2 \left\|\Gamma_1^B[\rho_B^i] + \Gamma_2^B[\rho_B^j]\right\|_2 \\ &= \left(\langle\Gamma_1^A[\rho_A^i], \Gamma_1^A[\rho_A^i]\rangle + \langle\Gamma_2^A[\rho_A^j], \Gamma_2^A[\rho_A^j] + \langle\Gamma_1^A[\rho_A^i], \Gamma_2^A[\rho_A^j] + \text{c.c.}\right)^{1/2} \\ &\quad\times \left(\langle\Gamma_1^B[\rho_B^i], \Gamma_1^B[\rho_B^i]\rangle + \langle\Gamma_2^B[\rho_B^j], \Gamma_2^B[\rho_B^j] + \langle\Gamma_1^B[\rho_B^i], \Gamma_2^B[\rho_B^j] + \text{c.c.}\right)^{1/2} .\end{aligned} \tag{A.6}$$

Then, if we assume that there exist $\epsilon_A, \epsilon_B \geq 0$ such that

$$\left\|\Gamma_1[\sigma_A^1]\right\|_2^2 + \left\|\Gamma_2[\sigma_A^2]\right\|_2^2 \leq 2\epsilon_A \quad\text{and}\quad \left\|\Gamma_1[\sigma_B^1]\right\|_2^2 + \left\|\Gamma_2[\sigma_B^2]\right\|_2^2 \leq 2\epsilon_B \tag{A.7}$$

for any $\sigma_A^{1(2)} \in \mathcal{S}_A$ and $\sigma_B^{1(2)} \in \mathcal{S}_A$, it follows from Eq. (A.6) that

$$\|\tilde{\rho}^R\|_1 \leq 2\sqrt{\left(\epsilon_A + \frac{1}{2}\left(\langle\Gamma_1^A[\rho_A^i], \Gamma_2^A[\rho_A^j] + \text{c.c.}\right)\right)\left(\epsilon_B + \frac{1}{2}\left(\langle\Gamma_1^B[\rho_B^i], \Gamma_2^B[\rho_B^j] + \text{c.c.}\right)\right)} . \tag{A.8}$$

Finally, from Eqs. (A.5) and (A.8) and using the Cauchy–Schwarz inequality one obtains

$$
\begin{aligned}
\left\| \left[\rho_{AB} \left(\Gamma^{(A,B)}_{1,\dots,n} \right) \right]^R \right\|_1 &\leq \sum_{ij} \sqrt{p_i p_j \left(\epsilon_A + \frac{1}{2} \left(\langle \Gamma_1^A[\rho_A^i], \Gamma_2^A[\rho_A^j] + \text{c.c.} \right) \right)} \\
&\quad \times \sqrt{p_i p_j \left(\epsilon_B + \frac{1}{2} \left(\langle \Gamma_1^A[\rho_B^i], \Gamma_2^B[\rho_B^j] + \text{c.c.} \right) \right)} \\
&\leq \sqrt{\left(\epsilon_A + \frac{1}{2} \left(\sum_{ij} p_i p_j \langle \Gamma_1^A[\rho_A^i], \Gamma_2^A[\rho_A^j] + \text{c.c.} \right) \right)} \\
&\quad \times \sqrt{\left(\epsilon_B + \frac{1}{2} \left(\sum_{ij} p_i p_j \langle \Gamma_1^A[\rho_B^i], \Gamma_2^B[\rho_B^j] + \text{c.c.} \right) \right)}. \qquad \text{(A.9)}
\end{aligned}
$$

Then, for every separable state one has

$$
\begin{aligned}
\left\| \left[\rho_{AB} \left(\Gamma^{(A,B)}_{1,\dots,n} \right) \right]^R \right\|_1 &\leq \left(\epsilon_A + \frac{1}{2} \left(\langle \Gamma_1^A[\rho_A], \Gamma_2^A[\rho_A] \rangle + \text{c.c.} \right) \right)^{1/2} \\
&\quad \times \left(\epsilon_B + \frac{1}{2} \left(\langle \Gamma_1^B[\rho_B], \Gamma_2^B[\rho_B] \rangle + \text{c.c.} \right) \right)^{1/2}. \qquad \text{(A.10)}
\end{aligned}
$$

The claim follows from generalising the above reasoning to a generic n. $\square$

Appendix B

Evaluating the Operator G-Concurrence

B.1 Operator G-Concurrence of isotropic states

In this section we compute the OGC of isotropic states, as defined in Eq. (3.13). Then we will prove that the bound Eq. (3.48) given in Theorem 3.6 is tight.

We have seen in Sec. 3.3 that the OSC of the generic isotropic state $\rho_{AB}(p)$ are gien by $\{1/d, p/d, \ldots, p/d\}$. Then, according to the definition of the OGC one finds

$$\mathbf{G}(\rho_{AB}(p)) = \left(\frac{1}{d}\right) p^{1-1/d^2}. \tag{B.1}$$

Now, since isotropic states are entangled if and only if $p > 1/(d+1)$ [99], it follows that

$$\sup\{\mathbf{G}(\rho_{AB}(p)) \mid (\rho_{AB}(p)) \text{ is separable}\} \leq \frac{1}{d}\left(\frac{1}{d+1}\right)^{1-1/d^2}. \tag{B.2}$$

By comparing Eq. (3.48) with Eq. (B.2) we observe that

$$\sup_{\text{sep}} \mathbf{G}(\rho_{AB}) - \sup_{\text{sep}} \mathbf{G}(\rho_{AB}(p)) \approx \mathscr{O}\left(d^{-3}\right), \tag{B.3}$$

hence we can conclude that the bound of Eq. (3.48) is tight (since the gap between the two upper bounds inEq. (B.3) shrinks faster than $\sup_{\text{sep}} \mathbf{G}(\rho_{AB})$ approaches zero).

The same argument does not help us to check if the upper bound of the OGC on the set of unsteerable states, expressed by Eq. (3.49), is tight too. Indeed, as shown in [36], isotropic states are steerable if and only if $p > (H_d - 1)/(d-1)$, where $H_d = \sum_{k=1}^{d}(1/k)$ is the Harmonic series. It follows that

$$\sup\{\mathbf{G}(\rho_{AB}(p)) \mid (\rho_{AB}(p)) \text{ is unsteerable}\} \leq \frac{1}{d}\left(\frac{H_d - 1}{d-1}\right)^{1-1/d^2} \tag{B.4}$$

and

$$\sup_{\text{non steer}} \mathbf{G}(\rho_{AB}) - \sup_{\text{non steer}} \mathbf{G}(\rho_{AB}(p)) \approx \mathscr{O}\left(d^{-3/2}\right), \tag{B.5}$$

i.e. the gap between the two upper bounds shrinks as fast as $\sup_{\text{non steer}} \mathbf{G}(\rho_{AB})$ approaches zero.

B.2 Operator G-Concurrence of Werner states

Firstly introduced in [31], Werner states in $\mathcal{S}_{AB}$ can be parametrized as

$$\rho^{w}_{AB}(p) = \left(\frac{d-1+p}{d-1}\right)\frac{\mathrm{I}}{d^2} - \left(\frac{p}{d-1}\right)\frac{\mathrm{V}}{d}, \tag{B.6}$$

with $0 \leq p \leq 1$ and $V(\varphi_1 \otimes \varphi_2) = \varphi_2 \otimes \varphi_1$ is the so-called flip operator. Let us define

$$\alpha_{d,p} = \frac{d-1+p}{d^2(d-1)}, \qquad \beta_{d,p} = \frac{p}{d(d-1)}$$

and recall that the flip operator can be expressed as $V = \sum_{a,b=1}^{d} |ab\rangle\langle ba|$. In analogy to what we did for isotropic states in Sec. 3.3, we compute the entries of the correlation matrix:

$$\begin{aligned}
\mathcal{C}(\rho^{w}_{AB}(p))_{ij,kl} &= \mathrm{Tr}\left(\, |j\rangle\langle i| \otimes |l\rangle\langle k| \left(\rho^{w}_{AB}(p)\right)\right) \\
&= \alpha_{d,p}\,\mathrm{Tr}\left(\, |j\rangle\langle i| \otimes |l\rangle\langle k|\, \mathrm{I}\right) \\
&\quad - \beta_{d,p}\,\mathrm{Tr}\left(\, |j\rangle\langle i| \otimes |l\rangle\langle k| \sum_{a,b=1}^{d} |a\rangle\langle b| \otimes |b\rangle\langle a|\,\right)
\end{aligned}$$

$$
\begin{aligned}
&= \alpha_{d,p}\, \delta_{ij}\, \delta_{kl} - \beta_{d,p} \sum_{a,b^1}^{d} \delta_{ia}\delta_{bj}\delta_{kb}\delta_{al} \\
&= \alpha_{d,p}\, \delta_{ij}\delta_{kl} - \beta_{d,p}\, \delta_{il}\delta_{jk}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\mathcal{C}(\rho^w_{AB}(p)) &= \sum_{ij,kl} \rho^w_{AB}(p)\, |ij\rangle\langle kl| \\
&= d\, \alpha_{d,p} |\psi^+\rangle\langle\psi^+| - \beta_{d,p}\, V,
\end{aligned}
$$

from which we conclude

$$
OSC(\rho^w_{AB}(p)) = \left(\frac{1}{d}, \frac{p}{d(d-1)}, \ldots, \frac{p}{d(d-1)} \right) \tag{B.7}
$$

and

$$
\mathbf{G}(\rho^w_{AB}(p)) = \frac{1}{d} \left(\frac{p}{d-1} \right)^{1-1/d^2}. \tag{B.8}
$$

Werner states are entangled if and only if $p > 1/(d+1)$ [31] and steerable if and only if $p > 1 - 1/d$ [36]. By substituting these limit values in Eq. (B.8), we found the upper bounds for $\mathbf{G}(\rho^w_{AB}(p))$ on the subset of separable and non-steerable states:

$$
\sup\{\mathbf{G}(\rho^w_{AB}(p)) \mid (\rho^w_{AB}(p)) \text{ is separable}\} \leq \frac{1}{d} \left(\frac{1}{d^2-1} \right)^{1-1/d^2}, \tag{B.9}
$$

$$
\sup\{\mathbf{G}(\rho^w_{AB}(p)) \mid (\rho^w_{AB}(p)) \text{ is unsteerable}\} \leq \left(\frac{1}{d} \right)^{2-1/d^2}. \tag{B.10}
$$

As for isotropic state, we would like to know if these bounds are tight in the sense of Eq. (B.3). However, the argument used in the previous section is not helpful in this case, in that

$$
\sup_{\text{non steer}} \mathbf{G}(\rho^w_{AB}) - \sup_{\text{non steer}} \mathbf{G}(\rho^w_{AB}(p)) \approx \mathscr{O}\left(d^{-2}\right), \tag{B.11}
$$

$$
\sup_{\text{non steer}} \mathbf{G}(\rho^w_{AB}) - \sup_{\text{non steer}} \mathbf{G}(\rho^w_{AB}(p)) \approx \mathscr{O}\left(d^{-3/2}\right). \tag{B.12}
$$

Bibliography

[1] M. Caiaffa and M. Piani, "Channel discrimination power of bipartite quantum states," *Physical Review A*, vol. 97, no. 3, p. 032334, 2018.

[2] M. Caiaffa and M. Piani, "Correlation-assisted process tomography," *arXiv preprint arXiv:1808.10835*, 2018.

[3] M. Caiaffa and M. Piani, "Measuring correlations via the operator Schmidt decomposition," *Provisional title, paper in preparation.*

[4] M. Caiaffa and M. Piani, "Metrics and pseudometrics in the space of quantum channels," *Provisional title, paper in preparation.*

[5] R. Jozsa and N. Linden, "On the role of entanglement in quantum-computational speed-up," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 459, pp. 2011–2032, The Royal Society, 2003.

[6] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM Review*, vol. 41, no. 2, pp. 303–332, 1999.

[7] D. E. Deutsch, A. Barenco, and A. Ekert, "Universality in quantum computation," *Proc. R. Soc. Lond. A*, vol. 449, no. 1937, pp. 669–677, 1995.

[8] D. P. DiVincenzo, "Quantum computation," *Science*, vol. 270, no. 5234, pp. 255–261, 1995.

[9] C. H. Bennett and S. J. Wiesner, "Communication via one-and two-particle operators on Einstein-Podolsky-Rosen states," *Physical Review Letters*, vol. 69, no. 20, p. 2881, 1992.

[10] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Physical Review Letters*, vol. 70, no. 13, p. 1895, 1993.

[11] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," *Rev. Mod. Phys.*, vol. 81, pp. 865–942, Jun 2009.

[12] J. P. Dowling and G. J. Milburn, "Quantum technology: the second quantum revolution," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 361, no. 1809, pp. 1655–1674, 2003.

[13] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.

[14] M. M. Wilde, *Quantum information theory*. Cambridge University Press, 2013.

[15] J. B. Altepeter, D. Branning, E. Jeffrey, T. Wei, P. G. Kwiat, R. T. Thew, J. L. O'Brien, M. A. Nielsen, and A. G. White, "Ancilla-assisted quantum process tomography," *Physical Review Letters*, vol. 90, no. 19, p. 193601, 2003.

[16] I. L. Chuang and M. A. Nielsen, "Prescription for experimental determination of the dynamics of a quantum black box," *Journal of Modern Optics*, vol. 44, no. 11-12, pp. 2455–2467, 1997.

[17] J. Poyatos, J. I. Cirac, and P. Zoller, "Complete characterization of a quantum process: the two-bit quantum gate," *Physical Review Letters*, vol. 78, no. 2, p. 390, 1997.

[18] A. Gilchrist, N. K. Langford, and M. A. Nielsen, "Distance measures to compare real and ideal quantum processes," *Physical Review A*, vol. 71, no. 6, p. 062310, 2005.

[19] C. Cohen-Tannoudji, F. Laloe, and B. Diu, *Mécanique quantique*, vol. 3. EDP Sciences, 2017.

[20] G. W. Mackey, "Quantum mechanics and Hilbert space," *American Mathematical Monthly*, vol. 64, pp. 45–57, 1957.

[21] A. Messiah, *Quantum Mechanics [Vol 1-2]*. Wiley & Sons, 1964.

[22] G. Ludwig, *Foundations of quantum mechanics I*. Springer Science & Business Media, 2012.

[23] A. S. Holevo, *Probabilistic and statistical aspects of quantum theory*, vol. 1. Springer Science & Business Media, 2011.

[24] K. Kraus, *States, effects and operations: fundamental notions of quantum theory*. Springer, 1983.

[25] L. E. Ballentine, "The statistical interpretation of quantum mechanics," *Reviews of Modern Physics*, vol. 42, no. 4, p. 358, 1970.

[26] L. E. Ballentine, *Quantum Mechanics: A Modern Development Second Edition*. World Scientific Publishing Company, 2014.

[27] V. Moretti, *Spectral theory and quantum mechanics*, vol. 64. Springer, 2013.

[28] J. Preskill, "Lecture notes for physics 229: Quantum information and computation," *California Institute of Technology*, vol. 16, 1998.

[29] L. A. Takhtadzhian, *Quantum mechanics for mathematicians*, vol. 95. American Mathematical Society, 2008.

[30] G. W. Mackey, *Mathematical foundations of quantum mechanics*. Courier Dover Publications, 2004.

[31] R. F. Werner, "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model," *Physical Review A*, vol. 40, no. 8, p. 4277, 1989.

[32] E. Schrödinger, "Discussion of probability relations between separated systems," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 31, pp. 555–563, 1935.

[33] J. Von Neumann, *Mathematical Foundations of Quantum Mechanics: New Edition*. Princeton University Press, 2018.

[34] I. Bengtsson and K. Życzkowski, *Geometry of quantum states: an introduction to quantum entanglement*. Cambridge University Press, 2017.

[35] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, "Quantifying entanglement," *Physical Review Letters*, vol. 78, no. 12, p. 2275, 1997.

[36] H. M. Wiseman, S. J. Jones, and A. C. Doherty, "Steering, entanglement, nonlocality, and the Einstein-Podolsky-Rosen paradox," *Physical Review Letters*, vol. 98, no. 14, p. 140402, 2007.

[37] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?," *Physical Review*, vol. 47, no. 10, p. 777, 1935.

[38] L. Gurvits, "Classical complexity and quantum entanglement," *Journal of Computer and System Sciences*, vol. 69, no. 3, pp. 448–484, 2004.

[39] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, "Bell nonlocality," *Rev. Mod. Phys.*, vol. 86, pp. 419–478, Apr 2014.

[40] J. S. Bell, "On the Einstein-Podolsky-Rosen paradox," in *John S Bell On The Foundations Of Quantum Mechanics*, pp. 7–12, World Scientific, 2001.

[41] H. Ollivier and W. H. Zurek, "Quantum discord: a measure of the quantumness of correlations," *Phys. Rev. Lett.*, vol. 88, no. 1, p. 017901, 2001.

[42] W. H. Zurek, "Decoherence, einselection, and the quantum origins of the classical," *Reviews of Modern Physics*, vol. 75, no. 3, p. 715, 2003.

[43] L. Henderson and V. Vedral, "Classical, quantum and total correlations," *J. Phys. A: Math. Gen.*, vol. 34, no. 35, p. 6899, 2001.

[44] K. Modi, A. Brodutch, H. Cable, T. Paterek, and V. Vedral, "The classical-quantum boundary for correlations: Discord and related measures," *Rev. Mod. Phys.*, vol. 84, pp. 1655–1707, Nov 2012.

[45] M. Piani, P. Horodecki, and R. Horodecki, "No-local-broadcasting theorem for multipartite quantum correlations," *Phys. Rev. Lett.*, vol. 100, no. 9, p. 090502, 2008.

[46] M. Piani, V. Narasimhachar, and J. Calsamiglia, "Quantumness of correlations, quantumness of ensembles and quantum data hiding," *New J. Phys.*, vol. 16, no. 11, p. 113001, 2014.

[47] S. Boixo, L. Aolita, D. Cavalcanti, K. Modi, and A. Winter, "Quantum locking of classical correlations and quantum discord of classical-quantum states," *International Journal of Quantum Information*, vol. 9, pp. 1643–1651, 2011.

[48] T. Chuan, J. Maillard, K. Modi, T. Paterek, M. Paternostro, and M. Piani, "Quantum discord bounds the amount of distributed entanglement," *Phys. Rev. Lett.*, vol. 109, no. 7, p. 070501, 2012.

[49] A. Streltsov, H. Kampermann, and D. Bruß, "Quantum cost for sending entanglement," *Phys. Rev. Lett.*, vol. 108, no. 25, p. 250501, 2012.

[50] D. Girolami, A. M. Souza, V. Giovannetti, T. Tufarelli, J. G. Filgueiras, R. S. Sarthour, D. O. Soares-Pinto, I. S. Oliveira, and G. Adesso, "Quantum discord determines the interferometric power of quantum states," *Physical Review Letters*, vol. 112, no. 21, p. 210401, 2014.

[51] S. Pirandola, "Quantum discord as a resource for quantum cryptography," *Scientific reports*, vol. 4, 2014.

[52] M.-D. Choi, "Completely positive linear maps on complex matrices," *Linear algebra and its applications*, vol. 10, no. 3, pp. 285–290, 1975.

[53] A. Jamiołkowski, "Linear transformations which preserve trace and positive semidefiniteness of operators," *Reports on Mathematical Physics*, vol. 3, no. 4, pp. 275–278, 1972.

[54] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Quantum circuit architecture," *Physical Review Letters*, vol. 101, no. 6, p. 060401, 2008.

[55] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, "Fundamental limits of repeaterless quantum communications," *Nature Communications*, vol. 8, pp. 15043 EP –, 04 2017.

[56] O. Oreshkov, F. Costa, and Č. Brukner, "Quantum correlations with no causal order," *Nature communications*, vol. 3, p. 1092, 2012.

[57] M. M. Wolf, "Quantum channels & operations: Guided tour," *Lecture notes available at http://www-m5.ma.tum.de/foswiki/pubM*, vol. 5, 2012.

[58] K. Kraus, "General state changes in quantum theory," *Annals of Physics*, vol. 64, no. 2, pp. 311–335, 1971.

[59] W. F. Stinespring, "Positive functions on C*-algebras," *Proceedings of the American Mathematical Society*, vol. 6, no. 2, pp. 211–216, 1955.

[60] R. A. Horn and C. R. Johnson, *Matrix analysis.* Cambridge University Press, 2013.

[61] J. Watrous, "Notes on super-operator norms induced by Schatten norms," *arXiv preprint quant-ph/0411077*, 2004.

[62] A. Y. Kitaev, "Quantum computations: algorithms and error correction," *Russian Mathematical Surveys*, vol. 52, no. 6, pp. 1191–1249, 1997.

[63] K. Chen and L.-A. Wu, "A matrix realignment method for recognizing entanglement," *Quantum Inf. Comput*, vol. 3, p. 193, 2003.

[64] O. Rudolph, "On the cross norm criterion for separability," *Journal of Physics A: Mathematical and General*, vol. 36, no. 21, p. 5825, 2003.

[65] C. Lupo, P. Aniello, and A. Scardicchio, “Bipartite quantum systems: on the realignment criterion and beyond,” *Journal of Physics A: Mathematical and Theoretical*, vol. 41, no. 41, p. 415301, 2008.

[66] P. Aniello and C. Lupo, “A class of inequalities inducing new separability criteria for bipartite quantum systems,” *Journal of Physics A: Mathematical and Theoretical*, vol. 41, no. 35, p. 355303, 2008.

[67] P. Aniello and C. Lupo, “On the relation between Schmidt coefficients and entanglement,” *Open Systems & Information Dynamics*, vol. 16, no. 02n03, pp. 127–143, 2009.

[68] S. Beigi, “A new quantum data processing inequality,” *Journal of Mathematical Physics*, vol. 54, no. 8, p. 082202, 2013.

[69] G. De las Cuevas, N. Schuch, D. Pérez-García, and J. I. Cirac, “Purifications of multipartite states: limitations and constructive methods,” *New Journal of Physics*, vol. 15, no. 12, p. 123021, 2013.

[70] O. Rudolph, “Further results on the cross norm criterion for separability,” *Quantum Information Processing*, vol. 4, no. 3, pp. 219–239, 2005.

[71] A. Gilchrist, D. R. Terno, and C. J. Wood, “Vectorization of quantum operations and its use,” *arXiv preprint arXiv:0911.2539*, 2009.

[72] R. A. Horn and C. R. Johnson, *Matrix analysis.* Cambridge University Press, 1990.

[73] J. A. Miszczak, “Singular value decomposition and matrix reorderings in quantum information theory,” *International Journal of Modern Physics C*, vol. 22, no. 09, pp. 897–918, 2011.

[74] C. J. Oxenrider and R. D. Hill, “On the matrix reorderings Γ and Ψ,” *Linear Algebra and its Applications*, vol. 69, pp. 205–212, 1985.

[75] Y.-Z. Zhen, Y.-L. Zheng, W.-F. Cao, L. Li, Z.-B. Chen, N.-L. Liu, and K. Chen, "Certifying Einstein-Podolsky-Rosen steering via the local uncertainty principle," *Physical Review A*, vol. 93, no. 1, p. 012108, 2016.

[76] S. Lang, *Algebra, Graduate texts in mathematics*, vol. 211. Springer-Verlag, New York, 2002.

[77] C.-K. Li, Y.-T. Poon, and N.-S. Sze, "A note on the realignment criterion," *Journal of Physics A: Mathematical and Theoretical*, vol. 44, no. 31, p. 315304, 2011.

[78] C.-J. Zhang, Y.-S. Zhang, S. Zhang, and G.-C. Guo, "Entanglement detection beyond the cross-norm or realignment criterion," *Phys. Rev. A*, vol. 77, p. 060301.

[79] P. Horodecki, "Separability criterion and inseparable mixed states with positive partial transposition," *Phys. Lett. A*, vol. 232, p. 333, 1997.

[80] R. Orús, "A practical introduction to tensor networks: Matrix product states and projected entangled pair states," *Annals of Physics*, vol. 349, pp. 117–158, 2014.

[81] F. Verstraete, V. Murg, and J. I. Cirac, "Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems," *Advances in Physics*, vol. 57, no. 2, pp. 143–224, 2008.

[82] J. I. Cirac and F. Verstraete, "Renormalization and tensor product states in spin chains and lattices," *Journal of Physics A: Mathematical and Theoretical*, vol. 42, no. 50, p. 504004, 2009.

[83] G. Vidal, "Entanglement renormalization: an introduction," *arXiv preprint arXiv:0912.1651*, 2009.

[84] U. Schollwöck, "The density-matrix renormalization group in the age of matrix product states," *Annals of Physics*, vol. 326, no. 1, pp. 96–192, 2011.

[85] J. Eisert, "Entanglement and tensor network states," *arXiv preprint arXiv:1308.3318*, 2013.

[86] Y.-Y. Shi, L.-M. Duan, and G. Vidal, "Classical simulation of quantum many-body systems with a tree tensor network," *Physical Review A*, vol. 74, no. 2, p. 022320, 2006.

[87] S. Singh, R. N. Pfeifer, and G. Vidal, "Tensor network decompositions in the presence of a global symmetry," *Physical Review A*, vol. 82, no. 5, p. 050301, 2010.

[88] M. B. Hastings, "An area law for one-dimensional quantum systems," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2007, no. 08, p. P08024, 2007.

[89] F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac, "Matrix product density operators: Simulation of finite-temperature and dissipative systems," *Physical Review Letters*, vol. 93, no. 20, p. 207204, 2004.

[90] M. Zwolak and G. Vidal, "Mixed-state dynamics in one-dimensional quantum lattice systems: a time-dependent superoperator renormalization algorithm," *Physical Review Letters*, vol. 93, no. 20, p. 207205, 2004.

[91] P. W. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," in *Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on*, pp. 124–134, Ieee, 1994.

[92] A. Datta and G. Vidal, "Role of entanglement and correlations in mixed-state quantum computation," *Physical Review A*, vol. 75, no. 4, p. 042310, 2007.

[93] B. Lanyon, M. Barbieri, M. Almeida, and A. White, "Experimental quantum computing without entanglement," *Physical Review Letters*, vol. 101, no. 20, p. 200501, 2008.

[94] B. Dakić, V. Vedral, and Č. Brukner, "Necessary and sufficient condition for nonzero quantum discord," *Physical Review Letters*, vol. 105, no. 19, p. 190502, 2010.

[95] M. Gessner, E.-M. Laine, H.-P. Breuer, and J. Piilo, "Correlations in quantum states and the local creation of quantum discord," *Physical Review A*, vol. 85, no. 5, p. 052122, 2012.

[96] B. Lanyon, P. Jurcevic, C. Hempel, M. Gessner, V. Vedral, R. Blatt, and C. Roos, "Experimental generation of quantum discord via noisy processes," *Physical Review Letters*, vol. 111, no. 10, p. 100504, 2013.

[97] M. A. Schlosshauer, *Decoherence: and the quantum-to-classical transition.* Springer Science & Business Media, 2007.

[98] G. Lindblad, "Completely positive maps and entropy inequalities," *Communications in Mathematical Physics*, vol. 40, no. 2, pp. 147–151, 1975.

[99] M. Horodecki and P. Horodecki, "Reduction criterion of separability and limits for a class of distillation protocols," *Physical Review A*, vol. 59, no. 6, p. 4206, 1999.

[100] G. Gour, "Family of concurrence monotones and its applications," *Physical Review A*, vol. 71, no. 1, p. 012318, 2005.

[101] R. Doran, *Characterizations of C* Algebras: the Gelfand Naimark Theorems*, vol. 101. CRC press, 1986.

[102] A. McIntosh, "Heinz inequalities and perturbation of spectral families," *Macqaurie Mathematical Reports*, 1979.

[103] R. Bhatia, "Interpolating the arithmetic–geometric mean inequality and its operator version," *Linear Algebra and its Applications*, vol. 413, no. 2-3, pp. 355–363, 2006.

[104] T. Ando, "Majorizations and inequalities in matrix theory," *Linear Algebra and its Applications*, vol. 199, pp. 17–67, 1994.

[105] H. Brunn, *Ueber Ovale und Eiflächen.* Akademische Buchdruckerei von R. Straub, 1887.

[106] H. Minkowski, *Geometrie der zahlen*, vol. 40. , 1910.

[107] A. W. Roberts, "Convex functions," in *Handbook of Convex Geometry, Part B*, pp. 1081–1104, Elsevier, 1993.

[108] M. A. Nielsen, "Conditions for a class of entanglement transformations," *Physical Review Letters*, vol. 83, no. 2, p. 436, 1999.

[109] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: theory of majorization and its applications.* Springer, 1979.

[110] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities.* Cambridge University Press, 1988.

[111] P. Alberti and A. Uhlmann, "Stochasticity and partial order: Doubly stochastic maps and unitary mixing," 1982.

[112] R. Bhatia, *Matrix analysis*, vol. 169. Springer Science & Business Media, 2013.

[113] M. A. Nielsen, "An introduction to majorization and its applications to quantum mechanics," *Lecture Notes, Department of Physics, University of Queensland, Australia*, 2002.

[114] A. S. Holevo, "On quasiequivalence of locally normal states," *Teoreticheskaya i Matematicheskaya Fizika*, vol. 13, no. 2, pp. 184–199, 1972.

[115] M. M. Wilde, "Recoverability for Holevo's just-as-good fidelity," *arXiv preprint arXiv:1801.02800*, 2018.

[116] A. Uhlmann, "The transition probability in the state space of a $*$-algebra," *Reports on Mathematical Physics*, vol. 9, no. 2, pp. 273–279, 1976.

[117] M. Tomamichel, M. Berta, and M. Hayashi, "Relating different quantum generalizations of the conditional Rényi entropy," *Journal of Mathematical Physics*, vol. 55, no. 8, p. 082206, 2014.

[118] V. Giovannetti, S. Lloyd, and L. Maccone, "Quantum-enhanced measurements: beating the standard quantum limit," *Science*, vol. 306, no. 5700, pp. 1330–1336, 2004.

[119] G. Tóth and I. Apellaniz, "Quantum metrology from a quantum information science perspective," *Journal of Physics A: Mathematical and Theoretical*, vol. 47, no. 42, p. 424006, 2014.

[120] G. M. D'Ariano, P. LoPresti, and M. G. A. Paris, "Using entanglement improves the precision of quantum measurements," *Phys. Rev. Lett.*, vol. 87, no. 27, p. 270404, 2001.

[121] A. M. Childs, J. Preskill, and J. Renes, "Quantum information and precision measurement," *Journal of Modern Optics*, vol. 47, no. 2-3, pp. 155–176, 2000.

[122] A. Acín, "Statistical distinguishability between unitary operations," *Physical Review Letters*, vol. 87, no. 17, p. 177901, 2001.

[123] B. Rosgen and J. Watrous, "On the hardness of distinguishing mixed-state quantum computations," in *Computational Complexity, 2005. Proceedings. Twentieth Annual IEEE Conference on*, pp. 344–354, IEEE, 2005.

[124] M. F. Sacchi, "Optimal discrimination of quantum operations," *Physical Review A*, vol. 71, no. 6, p. 062340, 2005.

[125] M. F. Sacchi, "Entanglement can enhance the distinguishability of entanglement-breaking channels," *Physical Review A*, vol. 72, no. 1, p. 014305, 2005.

[126] S. Lloyd, "Enhanced sensitivity of photodetection via quantum illumination," *Science*, vol. 321, no. 5895, pp. 1463–1465, 2008.

[127] B. Rosgen, "Additivity and distinguishability of random unitary channels," *Journal of Mathematical Physics*, vol. 49, no. 10, p. 102107, 2008.

[128] J. Watrous, "Notes on super-operator norms induced by Schatten norms," *Quantum Info. Comput.*, vol. 5, pp. 58–68, Jan. 2005.

[129] J. Watrous, "Distinguishing quantum operations having few kraus operators," *Quantum Information & Computation*, vol. 8, no. 8, pp. 819–833, 2008.

[130] G. M. D'Ariano and P. L. Presti, "Imprinting complete information about a quantum channel on its output state," *Physical Review Letters*, vol. 91, no. 4, p. 047902, 2003.

[131] A. Jenčová and M. Plávala, "Conditions for optimal input states for discrimination of quantum channels," *arXiv preprint arXiv:1603.01437*, 2016.

[132] F. G. Brandão, M. Piani, and P. Horodecki, "Generic emergence of classical features in quantum Darwinism," *Nature communications*, vol. 6, p. 7908, 2015.

[133] G. Vidal, "Entanglement of pure states for a single copy," *Physical Review Letters*, vol. 83, no. 5, p. 1046, 1999.

[134] O. Rudolph, "Computable cross-norm criterion for separability," *Letters in Mathematical Physics*, vol. 70, no. 1, pp. 57–64, 2004.

[135] S. Luo, "Using measurement-induced disturbance to characterize correlations as classical or quantum," *Phys. Rev. A*, vol. 77, no. 2, p. 022301, 2008.

[136] G. M. D'Ariano, M. G. Paris, and M. F. Sacchi, "Quantum tomography," *Advances in Imaging and Electron Physics*, vol. 128, pp. 206–309, 2003.

[137] A. Bisio, G. Chiribella, G. DAriano, S. Facchini, and P. Perinotti, "Optimal quantum tomography of states, measurements, and transformations," *Physical Review Letters*, vol. 102, no. 1, p. 010404, 2009.

[138] G. D'Ariano and P. L. Presti, "Quantum tomography for measuring experimentally the matrix elements of an arbitrary quantum operation," *Physical Review Letters*, vol. 86, no. 19, p. 4195, 2001.

[139] R. J. Duffin and A. C. Schaeffer, "A class of nonharmonic fourier series," *Transactions of the American Mathematical Society*, vol. 72, no. 2, pp. 341–366, 1952.

[140] N. R. Howes, *Modern analysis and topology.* Springer Science & Business Media, 2012.

[141] I. Nechita, Z. Puchała, Ł. Pawela, and K. Życzkowski, "Almost all quantum channels are equidistant," *Journal of Mathematical Physics*, vol. 59, no. 5, p. 052201, 2018.