Università degli Studi di Padova

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master Degree in Physics

Final Dissertation

Optical non-classicality as a Quantum Resource in Continuous-Variable Quantum Information

Thesis supervisor

Prof. Simone Montangero

Thesis co-supervisor

Prof. Martin Plenio

Candidate

Giovanni Ferrari

Academic Year 2019/2020

Abstract

In this thesis we address the problem of building a Quantum Resource Theory in infinite dimension. In particular, we study bosonic non-classicality as a Quantum Resource in continuous-variable Quantum Information. After reviewing the formalism of open quantum systems and Quantum Optics, we introduce the framework of Quantum Resource Theories and we discuss the case of non-classicality, and its applications in Quantum Optics and Quantum Technologies. Finally, we study a Resource Theory of non-classicality based on the standard and measured relative entropies of non-classicality as resource monotones and we prove, for the first time in an infinite-dimensional Resource Theory, a bound for asymptotic conversion rates.

Contents

Acknowledgments

Finalizing this thesis marks the end of the most intense and crucial part of my life so far, which began five years ago with my enrollment at the University of Padova and the Scuola Galileiana di Studi Superiori. It is superfluous to say that a number of people helped me getting through this journey, which would have been much worse, if not impossible, without them.
To begin with, I am grateful to all the professors I encountered throughout these years. One way or the other, they all taught me something, and this is certainly one of the best reasons to thank someone for. In particular, I want to mention Prof. Stefano Giusto, for being my bachelor supervisor, and giving me advice even after I abandoned the noble path of High Energy Physics; Prof. Luca Martucci, for being a valuable galileian tutor; Prof. Simone Montangero, for introducing me to Quantum Information Theory, but mostly for being so patient and supportive; Prof. Martin Plenio, for welcoming me into his group at the University of Ulm, for giving me the opportunity to write this thesis there, and for the rapidity of his responses, despite the dimension of the group; Prof. Pieralberto Marchetti, for being my referee for this thesis. I will not thank Paolo Ciatti for all the help he gave me, as he always tell me that he is just doing his job. At any rate, I am grateful to him for showing me what kind of scholar I would like to become.
I want to thank the Institute of Theoretical Physics at the University of Ulm, for being so welcoming during my very short (alas) stay there. In particular, Ludovico and Thomas have been wonderful mentors, and introduced me to the world of scientific research while giving me all the support I needed. Giovanni, Michele and Nicola deserve a special mention for making the coffee break one of the most pleasant moments of the day.
As for the Scuola Galileiana of Padova, I really need to say: "so long, and thanks for all the fish"[1]. Apart from the huge economical help, there is where I met some of the most extraordinary people of my whole life. A special thanks goes to some of its alumni, which gave me valuable advice during these years: Carlo Maria, Davide, Enrico, Gabriele, Omar, Roberto. A collective mention goes to all of its students whom I spent part of these five years with.
I want to thank all the people who struggled and suffered with me during these years. To begin with, a big thanks goes to Enrico and Umberto, members of the glorious laboratory group known as "I tre porcellini", whose deeds will be passed on to the next generations. Perhaps we were not the best at experiments, but we had a lot of fun. In particular, I want to thank Enrico for being the best neighbour I could dream of, and for being wiser than I will ever be (I am not quite sure whether this is an actual compliment); I want to thank Umberto for training my patience, and for being always by my side. I extend the

[1]Cited from "The Hitchhiker's Guide to the Galaxy".

gratitude to the other members of the group “i sei nani”, whom I shared my life at the Scuola Galileiana with: Andrea, Antonio and Matteo. We shared joys and sorrows for five long years, and I am sure this is not the end of it. Finally, I want to thank all of the other members of the group “Persone etc...”: Alessandro, Giacomo, Giovanni, Guglielmo, Luna, Nicola, Paolo. If leaving Padova hurts so much, it is because of all of these folks.

I also want to thank the canteens which nourished me while I was far from home. I am grateful to Piovego, for hardening my spirit; to Forcellini, for delighting my senses; to Pio X and Murialdo, for taking care of me after any tough day.

Even though I rarely stay in Ferrara, my hometown, there are a few people there I want to mention. I want to thank Alessandro, Edoardo, Manuel, Matteo, Mattia, Mirco, Mirko, Jacopo, Riccardo and Roberto for welcoming me any time I go back there. A very special thanks goes to Eleonora and her family, for being always by my side, even at great distances. Not to mention Luisa Maria, whom I will always be profoundly grateful to.

As obvious as it might sound, I have to say that I would not have been here without my family. I want to thank my brother Niccolò, my parents Andrea and Lucia, and all of my relatives who gave me support during these years: Antonio, Carla, Eliana, Lorenzo, Marco, Roberta, Stefano. In particular, I will always be grateful to my father for passing on to me his passion for Science, and for all the help he gave me for preparing for the entrance test of the Scuola Galileiana.

Finally, I reserved my biggest thank for Luna, who never abandoned me, and made everything much lighter and easier.

Introduction

More than a century after the seminal works which led to the complete formalisation of Quantum Mechanics, scientists are still far from making their peace with the disruptive innovations they introduced. Nonetheless, there is a rapidly increasing effort devoted to exploring how we can make use of the rules of Quantum Mechanics in order to obtain technological advantages. Interestingly enough, the most counterintuitive and debated features of Quantum Theory seem to be precisely the ones allowing for better performances with respect to classical technologies. It is likely that we will learn how to exploit these peculiar aspects of Quantum Theory in (every-day?) technological devices much before we will, if ever, get used to them.
The exploration of concepts such as quantum entanglement and quantum coherence paved the way for the birth of what is now called Quantum Information Theory, a flourishing subfield of Quantum Physics. Throughout this thesis, we will employ the language and technical tools of Quantum Information Theory and the theory of quantum open systems; at any rate, we will introduce in detail the large majority of the objects we will make use of, and point to references otherwise.
Technological implementations that make use of properties of Quantum Mechanics such as entanglement or coherence are collectively known as Quantum Technologies and, as we already anticipated, represent a very active area of research. While many challenges are faced at the experimental level, in the last few decades these developments has motivated a whole new line of theoretical research: Quantum Resource Theories. Actually, this rapidly emerging branch of Quantum Information Theory is deeply rooted in experimental research, as it sprang directly from the awareness that quantum objects are incredibly hard to manipulate. Indeed, if on one hand Quantum Mechanics describes how a quantum state can evolve in time in full generality, on the other hand Quantum Resource Theories considers the question: which kinds of quantum dynamics can we reproduce in an actual laboratory? In practice, any Quantum Resource Theory starts by identifying a set of "allowed" (i.e., experimentally accessible) quantum states $\mathcal{D}_f(\mathcal{H})$ and a set of allowed operations $\mathcal{O}_f(\mathcal{H} \to \mathcal{H}')$, based on some physically motivated assumptions. This peculiar approach to Quantum Theory also justifies the choice of the word "resource": just as mineral resources cannot be produced, but only consumed and exploited, quantum resources cannot be generated in a quantum system once experimental limitations are taken into account. Obviously, the notion of resources itself will vary depending on the particular experimental setting considered and on the choice of $\mathcal{D}_f(\mathcal{H})$ and $\mathcal{O}_f(\mathcal{H} \to \mathcal{H}')$: different initial hypothesis lead to completely different Quantum Resource Theories.
In general, a quantum resource is a certain feature that quantum objects such as states and operations may (or may not) display, that is not observed in our macroscopic, classical

world, and that can be of some practical use. Quantum states and operations containing some amount of resources are said to be resourceful, and can usually be exploited to enhance the performance of certain protocols, or even to achieve certain tasks that otherwise would be absolutely impossible.

A central problem of any Resource Theory is that of quantifying the amount of resource contained in a given quantum object, as our ability to make any quantitative statement depends upon it. To this end, a resource quantifier $\mathcal{F}$, assigning a "value" $\mathcal{F}(\rho) \in \mathbb{R}_+$ to any state ρ, can be defined; they are known also as resource monotones, as they cannot increase in value as long as we apply only operations in $\mathcal{O}_f(\mathcal{H} \to \mathcal{H}')$ to ρ. This is a direct consequence of the fact that resources cannot be produced with the readily available operations: otherwise, they would not be resources at all. We will see many examples of resource monotones in the course of the thesis.

Entanglement theory, being the first and most studied Resource Theory, is a canonical example. Here, all the ingredients of the Resource Theory are easily identifiable and physically intuitive: the experimental limitation consists in the spatial separation between two or more laboratories, and the fact that quantum states cannot be reliably sent over long distances; the quantum resource is, of course, entanglement itself (even though it is not trivial to determine what entanglement actually is); an information-theoretic task which cannot be performed without entanglement is the celebrated quantum teleportation of a quantum state; as for the resource monotones, many proposal have been studied throughout more than two decades, and some of them will be discussed in this thesis. For long, quantum entanglement was believed to be the most distinctive feature of Quantum Mechanics. With the advent of Quantum Resource Theories, it became clear that, depending on the context, different properties of quantum systems can emerge as "quantum signatures", i.e., as quantifiable departures from those behaviors Classical Mechanics got us accustomed to.

Now that we presented the questions which motivate the theoretical research in Quantum Resource Theories, it is time to ask ourselves what kind of answers they manage to provide. The most basic problem that can be addressed within a Quantum Resource Theory regards the manipulation of resources via state interconversion. Usually, when a certain amount of resource is at our disposal, it is not in the most useful form. Hence, it is fundamental to understand to what extent we can transform certain states into others once constraints on the set of operations are present. The monotonicity of resource quantifiers can be exploited to prove useful upper bounds on the efficiency of such processes: if resources cannot be generated during the evolution, the output state must be less resourceful than the input one.

In the spirit of Resource Theories, we want to obtain results which can be of some use in actual experimental situations. For instance, exact transformation between states are far too ideal, as no realistic state can actually be noiseless, and small deviations from the target are irrelevant for practical purposes. So, we want to relax the requirement from exact transformation between states to approximate (in a sense that we will make rigorous) ones. It is important to notice that this is not just a useless mannerism: typically, many state conversions become possible only if we allow for small errors in the result. Moreover, we want to allow for infinitely many identical copies of the input states, as in many experimental situations copies of a quantum state can be generated in sequence. In other words, we want to approximately convert the state $\rho^{\otimes n}$ into $\rho'^{\otimes m}$, with n being large by hypothesis and m being as large as possible. This particular type of state conversion is known as asymptotic state conversion, and its maximum efficiency is given in terms of the number of copies of the output state ρ' per copy of the input state ρ. Depending on whether we require the error of each single copy of the output state or their sum to be "small", we obtain two slighly different processes, with two different maximal efficiencies, denoted respectively as $\widetilde{R}(\rho \to \rho')$ and $R(\rho \to \rho')$. Since the latter case is more stringent, it clearly holds $R(\rho \to \rho') \leq \widetilde{R}(\rho \to \rho')$. Our aim is to arrive at a result in the form

$$\widetilde{R}(\rho \to \rho') \leq \frac{\mathcal{F}(\rho)}{\mathcal{F}(\rho')} \tag{1}$$

for some resource quantifier $\mathcal{F}$, by suitably exploiting its monotonicity.

A powerful characteristic of Quantum Resource Theories is their generality, as they do not depend directly from the details of the particular physical system we are considering. Nonetheless, when one constructs a Resource Theory, he/she usually has in mind a class of quantum systems motivating the choice of the constraints. For example, the physical framework which we will mostly work with throughout this thesis is that of Quantum Optics. Many Quantum Technologies, such as quantum computing, quantum communication and many more, rely on quantum optical platforms. These quantum systems represent some of the most promising realisations of quantum resources-driven technological devices. The easiest quantum states to be generated and manipulated in a quantum optical system are coherent states and their probabilistic mixtures. Moreover, states which do not belong to this class may display operational advantages in some relevant tasks such as entanglement generation, secure quantum key distribution, quantum computing, and quantum metrology. So, in this case, the choice of the set of available states is very natural. States which do not belong to the aforementioned set are said to be non-classical, and can can be regarded as resourceful. These premises lead to the Resource Theory of optical non-classicality.

Quantum optical systems are described by continuous variables; they display advantages with respect to discrete ones but, at the same time, their mathematical treatment is usually more cumbersome. For example, some mathematical results valid in finite-dimensional spaces may fail for systems living on infinite-dimensional Hilbert spaces. In particular, resource monotones usually becomes much "wilder", loosing some interesting analytical properties which can be crucial to prove certain results. For example, when dealing with approximate transformations, some kind of continuity is required, as we need to know whether similar states have also similar resource contents. Continuity is surprisingly hard to have in infinite-dimensional spaces, and hence the efficiency of asymptotic conversions typically cannot be bounded via resource monotones.

Despite the difficulties, we succeeded in proving a result in the form of (1) for the Resource Theory of optical non-classicality by introducing a resource monotone based on the measured relative entropy, and showing that it displays the properties we need. In particular, it is lower semi-continuous, meaning that small errors cannot cause the resource content of a quantum state to drop abruptly. This implies that it cannot be "much easier" to produce the approximate output state rather than the exact one. It is also super-additive, meaning that the resource content of many copies of a state cannot be less than the sum of the resource contents of each single copy. By combining these properties, we arrive at the desired result. To the best of our knowledge, this is the first example of such a bound in an infinite-dimensional Resource Theory. Furthermore, the whole machinery developed throughout the work can be readily applied to other infinite-dimensional Re-

source Theories. The bound can be easily approximated via a variational program as $\mathcal{F}$ can be written as a maximisation, and it can be upper bounded by a quantity written as a minimisation. By plugging suitable ansatzes in both the optimisations, we can easily upper bound the numerator and lower bound the denominator, hence approximating the bound over $R(\rho \to \rho')$ up to an arbitrary precision. This is a crucial point, as many bounds in Quantum Resource Theories are unmanageably hard to compute in practice, even in finite dimensional spaces. Furthermore, we prove that the resource content of any finite energy (and hence physical) state, is finite. This is a highly non-trivial yet crucial result in an infinite-dimensional Resource Theory: if the resource monotone diverges, the bound (1) may become meaningless.
Finally, we prove a number of additional results which help further the computation of the aforementioned bound. A method for efficiently approximating, up to an arbitrary precision, the resource content of general Fock diagonal states is given. It is then applied to Fock states and noisy Fock states, i.e., classical mixtures of Fock and thermal states. Noisy Fock states are often a good model for Fock states produced in a non-ideal laboratory. The resource content is bounded for cat states as well, for different values of the defining parameter. Protocols for the purification of Fock states and the concentration of cat states are considered. Both these classes of quantum states find applications, for instance, in linear optical quantum computation, where the considered experimental restrictions do apply.
In the following, we briefly summarize the content of the thesis.
In Chapter 1 we start by reviewing the Quantum Theory for open systems, which will be our natural habitat for the rest of the work. We introduce all the basic objects we need, such as density operators, quantum channels, POVMs, channels and measurements dilations, alongside with some more specific results which will come in handy in the following chapters. Then, we review the well-established theory of Quantum Optics, which will serve as the physical background for the last two chapters. After a brief review of the theoretical fundaments of the field, we will concentrate mostly on some specific concepts which will be extensively used afterwards: s-ordered characteristic functions, quasi-probability distributions, gaussian states and gaussian operations, linear optical operations. To conclude the chapter, we address a central problem of Quantum Information Theory, namely quantifying the distance (in a loose sense) between quantum states. To this end, we discuss some very popular quantifiers of such "distance": trace norm distance, relative entropy and measured relative entropy, and respective analytical properties. Arguably, the most important result presented in this chapter is Lemma 1.16, concerning a variational (i.e., expressed in terms of optimisations) representation of the meaured relative entropy. Ultimately, this result will stand at the core of Chapter 4.
In Chapter 2, the framework of Resource Theories, and in particular Quantum Resource Theories, is presented in detail. Here, the language and the majority of technical tools exploited in the rest of the work are introduced. In section 2.1, the general philosophy of Resource Theories is presented. Moreover, the difference between Quantum Resource Theories defined at the level of states and at the level of operations is studied. The discussion becomes more quantitative in Section 2.2: we present the notion of resource monotones, with relevant explicit examples such as relative entropies and robustnesses of resources, and many crucial analytical properties we might require for them. A basic resource-theoretic task, namely the interconversion between resourceful quantum states by means of free operations only, is addressed and discussed in detail for different scenar-

ios. A particular case of state interconversion, namely the asymptotic interconversion, will serve as a motivation for the definition of a new Resource Theory of non-classicality, which Chapter 4 is devoted to. Finally, some of the most known examples of Quantum Resource Theories are briefly considered: entanglement, coherence and quantum thermodynamics. They allow us to show some examples of resource monotones displaying appealing analytical properties and some typical resource-theoretic results.

In Chapter 3, the concept of (optical) non-classicality is finally introduced. We define the set of free states, i.e., classical states, and discuss possible choices for the set of free (classical) operations. Different mathematical characterisations for classical states and operations are discussed as well. Then, we start reviewing some past proposals for non-classicality Quantum Resource Theories. Our aim is twofold: we want to familiarize with the topic and, at the same time, to prepare the background for relating these Resource Theories with the one developed in the following chapter.

Finally, in Chapter 4, we present in full detail our proposal for a Resource Theory of non-classicality, based on the relative entropy and its measured version as resource monotones. The heart of the Chapter is represented by Theorem 4.9. It allows us to prove crucial properties for our monotones, and ultimately to prove a result which has no precedents in an infinite-dimensional Resource Theory: a bound on (maximal) asymptotic conversion rates. We also explain why other Resource Theories of non-classicality were not able to prove a similar result. At last, applications to specific states are considered.

1 Preliminaries and formalism

1.1 Foundations

1.1.1 Axioms

We start by briefly reviewing the formalism of open quantum systems, which represents the most sensible framework to deal with Quantum Resource Theories and Quantum Information in general. An open quantum system is a physical system which obeys the rule of Quantum Mechanics and interacts with an **environment** whose properties we are not interested in. On the contrary, a closed quantum system is a quantum system which does not interact with anything but itself. No realistic system can actually be considered closed, as any effort to perfectly isolate it from the environment is doomed to fail. Nonetheless, in many cases it is reasonable to approximate a physical system to be closed; at any rate, closed systems can be regarded as a special case of open ones, so we will focus our attention on the latter case from now on. For the theory of open quantum systems we will refer mostly to [1], [2] and [3].
Mathematically, any quantum system is associated to a **Hilbert space** $\mathcal{H}$ [4]; their formalism allows to define Quantum Mechanics by means of the following axioms.

A1) A physical system is completely described by its **state**, which is represented by a **density operator** $\rho \in \mathcal{D}(\mathcal{H})$, i.e., an operator acting on $\mathcal{H}$ and satisfying:

  S1) $\rho^\dagger = \rho$;

  S2) $\operatorname{Tr} \rho = 1$;

  S3) $\rho \geq 0$.

A2) The evolution of a quantum state is described by a **completely-positive trace-preserving** (**CPTP**) map $\mathcal{E} : \mathcal{D}(\mathcal{H}) \to \mathcal{D}(\mathcal{H}')$, i.e., a linear superoperator satisfying:

  E1) $\mathcal{E}(\rho)^\dagger = \mathcal{E}(\rho)$ for any $\rho \in \mathcal{D}(\mathcal{H})$;

  E2) $\operatorname{Tr} \mathcal{E}(\rho) = \operatorname{Tr} \rho$ for any $\rho \in \mathcal{D}(\mathcal{H})$;

  E3) $(\mathcal{E} \otimes I^E)(\rho) \geq 0$ for any $\rho \in \mathcal{D}(\mathcal{H} \otimes \mathcal{H}^E)$ (**complete positivity**).

A3) A **measurement** is described by a **positive operator-valued measure** (**POVM**), i.e., a set of operators $\{E_n\}_n$ satisfying:

M1) $E_n^\dagger = E_n$;

M2) $E_n \geq 0$;

M3) $\sum_n E_n = I$.

The probability of the n-th outcome is given by $\mathrm{Tr}(\rho E_n)$.

A4) Physical systems can be composed to form larger ones. The Hilbert space of a composite system is the tensor product of the Hilbert spaces of its subsystems. If each one of its subsystems is prepared in a state, the state of the total system is given by their tensor product.

It is rather easy to show that these axioms can be derived from those for closed systems in the case we have access to just a part of the total system, and the rest acts as an environment.
Throughout the work, we will often refer to topologies which can be defined on a Hilbert space, and to classes of operators acting on it. For a very brief compendium of these topics, see Appendix A.

1.1.2 Pure and mixed states

What is the link between density operators and states in the closed systems formulation of Quantum Mechanics? To begin with, an operator satisfying S1)-S3) has always a spectral decomposition in the form

$$\rho^A = \sum_{j=1}^{d} p_j \left|\phi_j\right\rangle\!\left\langle\phi_j\right| \,, \tag{1.1}$$

where the p_j are real non-negative numbers summing up to 1, $\{|\phi_j\rangle\}_{j=1,\dots,d}$ is an orthonormal basis for $\mathcal{H}$ and $d = \dim \mathcal{H}$. If we substitute the spectral decomposition of ρ_A in the formula for the outcome probabilities we get

$$\mathcal{P}_\rho(n) = \sum_{j=1}^{d} p_j \left\langle\phi_j\right|E_n\left|\phi_j\right\rangle \,. \tag{1.2}$$

By looking at this expression, we are enticed to interpret p_j as the probability associated to the state $|\phi_j\rangle$, and ρ^A as an **ensemble**. Since these probabilities arise from ignorance rather than quantum effects, we will sometimes refer to ρ^A as a **classical mixture** of the states $\{|\phi_j\rangle\}_j$.
Finally, we can see that density operators generalize the notion of closed system states. Indeed, a density operator

$$\psi = |\psi\rangle\!\langle\psi|$$

can be associated to any state $|\psi\rangle$ of a closed system. Such a density operator is an orthogonal projector, and represent what is called a **pure state** (what we would call simply "a state" in the theory of closed systems). Density operators whose spectral decomposition has more than just one term are not orthogonal projectors, and represent **mixed states** instead.

1.1.3 Quantum channels

Unitary operators emerge naturally in the context of closed quantum systems evolution once we require that the normalization of any state vector must remain constant; similarly, a superoperator fulfilling conditions E1)-E3) is clearly the most general object mapping a quantum state of an open system into another. Such an object is often referred to as a **quantum channel**, where the term "channel" is borrowed from classical information theory. The following result that gives a more explicit expression for these particular maps.

Theorem 1.1. *[1, Theorem 8.1] A linear superoperator $\mathcal{E}$ satisfies E1)-E3) if and only if*

$$\mathcal{E}(\rho) = \sum\nolimits_a K_a \rho K_a^\dagger \,, \tag{1.3}$$

with

$$\sum\nolimits_a K_a^\dagger K_a = I \,. \tag{1.4}$$

In order to derive the previous result it is crucial that in E3) we require the map $\mathcal{E}$ to be completely-positive and not only positive. From a physical point of view, this is well justified: we want our map $\mathcal{E}$ to map states into states even if we consider our system to be embedded in a larger one, with $\mathcal{E}$ acting trivially on the environment. It is not trivial to see why complete-positivity is indeed a stronger condition than poisitivity; an example of a positive-but-not-completely-positive superoperator is given by the transposition map $\rho \mapsto \rho^T$ [2].
As in the case of states, quantum channels are a generalization of unitary evolution. Indeed, if in equation (1.4) there is only one term, we are back to the definition of a unitary operator. Unitary operations are a very special case of quantum channels: it can be proven [2] that a quantum channel from a Hilbert space to itself is invertible if and only if it is a unitary map.
To conclude this section, we introduce the notion of the **dual** or **adjoint** $\mathcal{E}^*$ of a quantum channel $\mathcal{E}$, defined throught the following relation:

$$\mathrm{Tr}\left(O\mathcal{E}(\rho)\right) = \mathrm{Tr}\left(\mathcal{E}^*(O)\rho\right) \,, \tag{1.5}$$

valid for any observable O. Obviously

$$\mathcal{E}(\rho) = \sum\nolimits_a K_a \rho K_a^\dagger \implies \mathcal{E}^*(O) = \sum\nolimits_a K_a^\dagger O K_a \,. \tag{1.6}$$

One can always choose whether to evolve the state of a system with a quantum channel (**Schrödinger picture**) or the observables with its dual (**Heisemberg picture**): by definition of $\mathcal{E}^*$, all measurable quantities will result the same. A dual of a quantum channel has an important properties: from (1.4) and (1.6) we see that

$$\mathcal{E}^*(I) = I \,.$$

A superoperator with this property is called a **unital map**.
Remark 1.2. *Unital maps send POVMs into POVMs. Indeed:*

$$\sum\nolimits_a \sum\nolimits_b K_b^\dagger E_a K_b = \sum\nolimits_b K_b^\dagger \left(\sum\nolimits_a E_a\right) K_b = \sum\nolimits_b K_b^\dagger K_b = I \,.$$

Unital maps stand at the core of the following Theorem, which will be extensively used throughout the rest of the work.

Theorem 1.3. *(Jensen's operator inequality)[5, Theorem 2.1] Let $\mathcal{H}$ be a (possibly infinite-dimensional) Hilbert space, and $f : \mathcal{B}^I_{sa}(\mathcal{H}) \to \mathcal{B}_{sa}(\mathcal{H})$ an operatorial map, where $\mathcal{B}^I_{sa}(\mathcal{H})$ is the space of self-adjoint bounded operators with spectrum in the real interval I. Then, the following statements are equivalent:*

- *f is operator-convex, i.e.:*

$$f(\lambda h_1 + (1-\lambda)h_2) \leq \lambda f(h_1) + (1-\lambda)f(h_2) \quad \forall\, h_1,\, h_2 \in \mathcal{B}^I_{sa}(\mathcal{H}) \text{ and } \lambda \in [0,1]\,.$$

- *For any unital map*

$$\Lambda : (\cdot) \mapsto \sum_{a=1}^{n} K_a^\dagger (\cdot) K_a\,,$$

  it holds that

$$f\left(\sum_{a=1}^{n} K_a^\dagger h K_a\right) \leq \sum_{a=1}^{n} K_a^\dagger f(h) K_a\,.$$

- *For any projector P, $s \in I$ and $h \in \mathcal{B}^I_{sa}(\mathcal{H})$ it holds:*

$$Pf(PhP + s(1-P))P \leq Pf(h)P\,.$$

Remark 1.4. *A notable example of an operator function which is operator concave (and hence for which all the previous inequalities hold with opposite signs) is the operator logarithm.*

1.1.4 Generalized measurements

The first thing one should notice about axiom A3) is that, contrarily to what happens for projective measurements on closed systems, it does not state anything about the state of the system after the measurement. Indeed, not only POVMs can describe measurements which are not projective, but they also encompass those cases in which the state after the measurement is completely unknown, or it is destroyed during the process.
These **generalized measurements** are often obtained by entangling the system one wants to extract information from with an ancilla system (the measurement apparatus), and projectively measuring the latter. For example, starting from system S in the state $|\psi^S\rangle$ we can add the ancillary system A and apply the entangling unitary

$$U\left(|\psi^S\rangle \otimes |0^A\rangle\right) \mapsto \sum\nolimits_a M_a^S |\psi^S\rangle \otimes |a^A\rangle\;.$$

At this point, if we project the system A onto one of the (mutually orthogonal) states $|a^A\rangle$, system S will be brought in the (unnormalized) state $M_a^S|\psi^S\rangle$. States $M_a^S|\psi^S\rangle$ and $M_b^S|\psi^S\rangle$ need not be orthogonal for $a \neq b$ so, in turn, consecutive measures need not to agree:

$$\mathcal{P}_{\psi^S}(b|a) = \frac{\|M_b^S M_a^S |\psi^S\rangle\|^2}{\|M_a^S|\psi^S\rangle\|^2} \neq \delta_{ab}\,.$$

Note that unitary transformation preserves the trace, so

$$1 = \mathrm{Tr}\left(\sum\nolimits_a M_a^S \left|\psi^S\right\rangle \otimes \left|a^A\right\rangle\right) = \left\langle\psi^S\right| \sum\nolimits_a M_a^{S\dagger} M_a^S \left|\psi^S\right\rangle \; ;$$

being this relation true for any $|\psi^S\rangle$, it follows that

$$\sum\nolimits_a {M_a^S}^\dagger M_a^S = I^S .$$

So, any set of **measurement operators** $\{M_a\}_a$ identify a POVM $\{E_a\}_a = \{M_a^\dagger M_a\}_a$ satisfying M1)-M3), but left-multiplying any of the M_a by a unitary operator leads to the same POVM: this is why a POVM alone does not give information about the post-measurement state of the system.
Notice that even if they need not to be so, measurement operators might actually be mutually orthogonal projectors. In this case, we are back to projective measurements. This is also the only case in which the measurement operators and the POVM operators coincide, since $P_a^\dagger P_b = \delta_{ab} P_a$ for any couple of mutually orthogonal projectors P_a and P_b.
Finally, we point out that any measurement with a discrete set of possible outcomes and for which the measurement operators are specified, can be described with a CPTP map. Indeed, if ρ_a^S is the output state associated to the a-th outcome, we can encode the result of the measurement in an ancillary system A, and (post-)select the resulting state only at the end:

$$\rho \mapsto \sum\nolimits_a \rho_a^S \otimes \left|a^A\right\rangle\!\left\langle a^A\right| \; .$$

Such a state is said to be the **flagged outcome** of a measurement, and the $|a^A\rangle$ states are know as **flags**.

1.1.5 Back again to closed systems, or: the Church of the larger Hilbert space

Up to now, we introduced the framework of open quantum systems as a generalization of the closed quantum systems one. Now it is time to close the circle: we want to show that we can always consider an open system to be a part of a larger closed one, for which "usual" axioms hold.
We start from the following result.
Theorem 1.5. *For any mixed state ρ^S of the system S, it is possible to find a* ***purification****, i.e., a pure state $|\psi^{SE}\rangle$ of the system SE such that:*

$$\rho^S = \mathrm{Tr}_E \left|\psi^{SE}\right\rangle\!\left\langle\psi^{SE}\right| \; .$$

Moreover, any other state obtained via $|\psi^{SE}\rangle \mapsto (I^S \otimes U^E)\,|\psi^{SE}\rangle$ is a valid purification of ρ^S. Finally, the dimension of E can be taken to be not larger than that of S.

Proof. Starting from the spectral decomposition of ρ^S,

$$\rho^S = \sum_{j=1}^{d} p_j \left|\phi_j^S\right\rangle\!\left\langle\phi_j^S\right| \; ,$$

it is easy to see that the pure state

$$|\psi^{SE}\rangle = \sum_{j=1}^{d} \sqrt{p_j}\, |\phi_j^S\rangle \otimes |\phi_j^E\rangle$$

for the system SE does the job:

$$\mathrm{Tr}_E\, |\psi^{SE}\rangle\!\langle\psi^{SE}| = \rho^S\,.$$

Here, $\{|\phi_j^E\rangle\}_j$ is an arbitrary basis for the auxiliary system E: we could have chosen a different basis as well, so

$$|\psi^{SE}\rangle \mapsto (I^S \otimes U^E)\, |\psi^{SE}\rangle$$

gives another valid purification of the state ρ^S.

□

As for quantum channels and generalized measurements, we state without proof the following two fundamental results.

Theorem 1.6. *(Stinespring's dilation theorem) Given a CPTP map $\mathcal{E} : \mathcal{D}(\mathcal{H}^S) \to \mathcal{D}(\mathcal{H}^{S'})$, there exist a system E, a pure state $|0^E\rangle$ of E and a unitary transformation $U : \mathcal{H}^{SE} \to \mathcal{H}^{S'E'}$ such that*

$$\mathcal{E}(\rho^S) = \mathrm{Tr}_{E'}\left[U\left(\rho^A \otimes |O^E\rangle\!\langle O^E|\right) U^\dagger\right] \tag{1.7}$$

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

Theorem 1.7. *(Naimark's dilation theorem) Given a POVM $\{E_n^S\}_n$ on the system S, there exist a system E, a pure state $|0^E\rangle$ of E, a unitary transformation $U : \mathcal{H}^{SE} \to \mathcal{H}^{SE}$ and a set of orthogonal projectors $\{P_n^E\}_n$ on system E such that*

$$\mathrm{Tr}(E_n\rho) = \mathrm{Tr}\left[\left(I^S \otimes P_n^E\right) U \left(\rho^A \otimes |O^E\rangle\!\langle O^E|\right) U^\dagger\right] \tag{1.8}$$

for any $\rho^S \in \mathcal{D}(\mathcal{H}^S)$ and any E_n^S.

These three results ensure that any open system can always be seen as a subsystem of a bigger, closed one. Hence, the theory for closed quantum systems can be completely recovered from that for open ones. This paradigm is known as the *Church of the larger Hilbert space*, and suggests that there is not a preferred point of view, but it is just a matter of taste.

1.2 Elements of Quantum Optics

1.2.1 The physical background

Quantum optics [6, 7, 8, 9] is the study of a discrete set of electromagnetic modes obeying the rules of Quantum Mechanics, as opposed to Quantum Field Theory which deals with a continuum of modes, and their interaction with matter. In a typical quantum optical system, the light can be considered to be constantly travelling in vacuum and undergoing some discrete transformations or measurements every now and then: hence,

the starting point of this story is just the free electromagnetic field. It is well known that, after canonical quantisation, the dynamics of one mode of the electromagnetic field in vacuum is governed by the Hamiltonian of a harmonic oscillator, where the roles of position and momentum are played by the magnetic and electric fields respectively.
For m modes, the total Hilbert space of our system is then the tensor product of m infinite-dimensional Hilbert spaces, so we can consider (and we will) $\mathcal{H}_m = L^2(\mathbb{R}^m)$. The self-adjoint canonical operators acting on $\mathcal{H}$ are denoted with $\{\hat{x}_j\}_j$ and $\{\hat{p}_j\}_j$ and obey the **Canonical Commutation Relations** (**CCR**):

$$[\hat{x}_j, \hat{p}_k] = i\delta_{jk} I \qquad [\hat{x}_j, \hat{x}_k] = [\hat{p}_j, \hat{p}_k] = 0\,, \tag{1.9}$$

where we set $\hbar = 1$. Canonical operators are sometime grouped as:

$$\hat{\boldsymbol{r}} = (\hat{x}_1, \hat{p}_1, \ldots, \hat{x}_m, \hat{p}_m)^T \tag{1.10}$$

Starting from these operators, one usually define the **creation** and **annihilation operators** (collectively known as **ladder operators**) respectively as:

$$\hat{a}_j^\dagger = \frac{\hat{x}_j - i\hat{p}_j}{\sqrt{2}} \qquad \hat{a}_j = \frac{\hat{x}_j + i\hat{p}_j}{\sqrt{2}}\,; \tag{1.11}$$

they clearly fulfill the following relations:

$$[\hat{a}_j, \hat{a}_k^\dagger] = \delta_{jk} I \qquad [\hat{a}_j, \hat{a}_k] = [\hat{a}_j^\dagger, \hat{a}_k^\dagger] = 0\,. \tag{1.12}$$

We will use $\hat{}$ symbols for canonical and ladder operators in order to avoid ambiguities with real and complex variables.
The unique state which is mapped to the null vector by any of the annihilation operators is called the **vacuum** and denoted with $|\mathbf{0}\rangle = \bigotimes_{j=1}^m |0_j\rangle$; starting from it, we can construct the (normalized) **Fock states** as $|\boldsymbol{n}\rangle = |n_1 \ldots n_m\rangle := \bigotimes_{j=1}^m \left(\hat{a}_j^\dagger\right)^{n_j} |0_j\rangle / \sqrt{n_j!}$, which clearly obey the relations:

$$\begin{aligned} \hat{a}_j^\dagger |n_1 \ldots n_j \ldots n_m\rangle &= \sqrt{n_j + 1}\, |n_1 \ldots (n_j + 1) \ldots n_m\rangle\,, \\ \hat{a}_j |n_1 \ldots n_j \ldots n_m\rangle &= \sqrt{n_j}\, |n_1 \ldots (n_j - 1) \ldots n_m\rangle\,. \end{aligned} \tag{1.13}$$

Thanks to the creation and annihilation operators, the Hamiltonian of the system can now be written in the form

$$H = \omega \sum_{j=1}^m \hat{a}_j^\dagger \hat{a}_j\,, \tag{1.14}$$

where we already discarded the zero-point energy and set all the frequencies ω_j to the same value ω, because in this work we have no interest in considering modes with different frequencies. Starting from the CCR, it is easy to show that the eigenstates of this Hamiltonian are precisely the Fock states we just introduced [10]: as a consequence, they form a complete and orthonormal basis for $\mathcal{H}_m$.
The energy of a state is proportional to its **photon number**

$$N(\rho) = \mathrm{Tr}\left(\rho \sum_{j=1}^m \hat{a}_j^\dagger \hat{a}_j\right) = \sum_{j=1}^m \sum_{n_j=1}^\infty n_j \langle n_j | \rho_j | n_j\rangle\,, \tag{1.15}$$

where $\rho_j = \mathrm{Tr}_{\hat{j}}\, \rho$ is the partial trace of ρ with respect to all systems but the j-th.

1.2.2 Coherent states

For any $\boldsymbol{\alpha} \in \mathbb{C}^m$, we define the **displacement operator** as

$$\begin{aligned}\mathcal{D}(\boldsymbol{\alpha}) :=& e^{\sum_j (\alpha_j \hat{a}_j^\dagger - \alpha_j^* \hat{a}_j)} \\ \overset{1}{=}& e^{\frac{1}{2}\sum_j |\alpha_j|^2 [\hat{a}_j^\dagger, \hat{a}_j]} e^{\sum_j \alpha_j \hat{a}_j^\dagger} e^{-\sum_j \alpha_j^* \hat{a}_j} \\ \overset{2}{=}& e^{-\frac{1}{2}\sum_j |\alpha_j|^2} e^{\sum_j \alpha_j \hat{a}_j^\dagger} e^{-\sum_j \alpha_j^* \hat{a}_j} ,\end{aligned} \tag{1.16}$$

where in 1 we applied the Baker-Campbell-Hausdorff (BCH) formula [11], which gives a simple result in 2 because the commutator $[\hat{a}_j, \hat{a}_j^\dagger] \propto I$ commutes with every other operator in the algebra generated by $\hat{a}_j$ and $\hat{a}_j^\dagger$ (it is **central**). We can derive also the equivalent expression

$$\mathcal{D}(\boldsymbol{\alpha}) = e^{\frac{1}{2}\sum_j |\alpha_j|^2} e^{-\sum_j \alpha_j^* \hat{a}_j} e^{\sum_j \alpha_j \hat{a}_j^\dagger} \tag{1.17}$$

by applying the BCH formula with $\hat{a}_j$ and $\hat{a}_j^\dagger$ in the opposite order.
Two basic properties of these operators are that $\mathcal{D}(\alpha)^\dagger = \mathcal{D}(-\alpha)$ and $\mathcal{D}(\alpha)^\dagger \mathcal{D}(\alpha) = I$. A slighly less trivial property can again be proven with the BCH formula:

$$\begin{aligned}\mathcal{D}(\boldsymbol{\alpha})\mathcal{D}(\boldsymbol{\beta}) =& e^{-\frac{1}{2}\sum_{jk}(\alpha_j \beta_k^* [\hat{a}_j^\dagger, \hat{a}_k] + \alpha_j^* \beta_k [\hat{a}_j, \hat{a}_k^\dagger])} \mathcal{D}(\alpha + \beta) \\ =& e^{\frac{1}{2}\sum_j (\alpha_j \beta_j^* - \beta_j \alpha_j^*)} \mathcal{D}(\boldsymbol{\alpha} + \boldsymbol{\beta}) \\ =& e^{\frac{1}{2}(\langle \boldsymbol{\beta}, \boldsymbol{\alpha}\rangle - \langle \boldsymbol{\alpha}, \boldsymbol{\beta}\rangle)} \mathcal{D}(\boldsymbol{\alpha} + \boldsymbol{\beta}) ,\end{aligned} \tag{1.18}$$

where $\langle \boldsymbol{\alpha}, \boldsymbol{\beta}\rangle := \boldsymbol{\alpha}^* \cdot \boldsymbol{\beta} = \sum_{j=1}^m \alpha_j^* \beta_j$ is the usual hermitian product on $\mathbb{C}^n$. From (1.18) we derive also:

$$\mathcal{D}(\boldsymbol{\beta})^\dagger \mathcal{D}(\boldsymbol{\alpha}) \mathcal{D}(\boldsymbol{\beta}) = e^{\langle \boldsymbol{\beta}, \boldsymbol{\alpha}\rangle - \langle \boldsymbol{\alpha}, \boldsymbol{\beta}\rangle} \mathcal{D}(\boldsymbol{\alpha}) . \tag{1.19}$$

But the most important property, and the main reason why we introduced these operators, is the following:

$$\begin{aligned}\hat{a}_j \mathcal{D}(\boldsymbol{\alpha}) |0 \dots 0\rangle =& \mathcal{D}(\boldsymbol{\alpha}) \mathcal{D}(\boldsymbol{\alpha})^\dagger \hat{a}_j \mathcal{D}(\boldsymbol{\alpha}) |0 \dots 0\rangle \\ \overset{4}{=}& \mathcal{D}(\boldsymbol{\alpha}) \left(\hat{a}_j + \left[-\sum\nolimits_k (\alpha_k \hat{a}_k^\dagger - \alpha_k^* \hat{a}_k), \hat{a}_j \right] \right) |0 \dots 0\rangle \\ =& \alpha_j \mathcal{D}(\boldsymbol{\alpha}) |0\rangle \ ;\end{aligned} \tag{1.20}$$

in 4 we used the well-known formula:

$$e^A B e^{-A} = \sum_{n=0}^{\infty} \frac{1}{n!} [A, B]_n , \qquad [A, B]_n := \underbrace{A, [A, ..., [A, B]]...]}_{n \text{ nested commutators}} , \tag{1.21}$$

where all the terms with $n > 1$ vanish because of the CCR. Hence, the **coherent state** $|\alpha\rangle := \mathcal{D}(\alpha) |0\rangle$ is an eigenstate of $\hat{a}_j$ with eigenvalue α_j. The last relation we describe is the following partition of the identity:

$$I = \frac{1}{\pi^m} \int d^{2m}\alpha \ |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| \ , \tag{1.22}$$

which allows for writing the trace of an operator O as:

$$\mathrm{Tr}\,(O) = \frac{1}{\pi^m} \int d^{2m}\alpha \ \langle \boldsymbol{\alpha} | O | \boldsymbol{\alpha} \rangle \ . \tag{1.23}$$

Coherent states were introduced for the very first time by Schrödinger himself [12], and then re-discovered and studied by Klauder [13], Glauber [14] (whom their paternity is usually associated to) and Sudarshan [15]. Their expression in terms of Fock states can be determined directly from their definition and from the expression appearing in the last line of (1.16):

$$|\boldsymbol{\alpha}\rangle = e^{-\frac{|\boldsymbol{\alpha}|^2}{2}} e^{\boldsymbol{\alpha}\cdot\hat{\boldsymbol{a}}^\dagger} e^{-\boldsymbol{\alpha}^*\cdot\hat{\boldsymbol{a}}} |\mathbf{0}\rangle = \bigotimes_{j=1}^{m} e^{-\frac{|\alpha_j|^2}{2}} e^{\alpha_j \hat{a}_j^\dagger} |0_j\rangle = \bigotimes_{j=1}^{m} e^{-\frac{|\alpha_j|^2}{2}} \sum_{n_j=0}^{\infty} \frac{\alpha_j^n}{\sqrt{n_j!}} |n_j\rangle \ .$$

Hence, multi-mode coherent states are just tensor products of single mode coherent states. Coherent states display a number of interesting physical and mathematical properties.

- They form an over-complete basis for $\mathcal{H}$, in the sense that any state can be written as a superposition of coherent states, but they are not linearly independent. States of an over-complete basis cannot be mutually orthogonal: indeed, by applying equation (1.18) it is easy to see that

$$\langle\boldsymbol{\alpha}|\boldsymbol{\beta}\rangle = e^{\frac{1}{2}(\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle - \langle\boldsymbol{\beta},\boldsymbol{\alpha}\rangle)} e^{-\frac{|\boldsymbol{\alpha}-\boldsymbol{\beta}|^2}{2}} = e^{-\frac{|\boldsymbol{\alpha}|^2}{2}} e^{-\frac{|\boldsymbol{\beta}|^2}{2}} e^{\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle} \ .$$

- They have a "classical" evolution in vacuum, in the sense that they oscillate with constant velocity:

$$|\alpha_j(t)\rangle = e^{-i\omega\hat{a}_j^\dagger\hat{a}_j t} |\alpha_j\rangle = e^{-\frac{|\alpha_j|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha_j^n e^{-in\omega t}}{\sqrt{n!}} |n\rangle = |e^{-i\omega t}\alpha_j\rangle \ .$$

  In particular:

$$\langle\alpha_j(t)|\hat{a}_j|\alpha_j(t)\rangle = \alpha_j(t) = e^{-i\omega t}\alpha_j \ ;$$

  which implies that

$$\langle\hat{x}_j\rangle_{\alpha_j(t)} = \langle\alpha_j(t)|\hat{x}_j|\alpha_j(t)\rangle = \alpha_j \cos\omega t \ , \ \langle\hat{p}_j\rangle_{\alpha_j(t)} = \langle\alpha_j(t)|\hat{p}_j|\alpha_j(t)\rangle = -\alpha_j \sin\omega t \ ;$$

  thus, position and momentum follow (on average) the solutions of motion of a classical harmonic oscillator.
  Their temporal evolution is easily re-absorbed by moving to a "rotating reference frame", where any coherent state simply remain constant at any time if no other physical operation is applied. So, in the following we will forget about the evolution of the fields in vacuum.

- They saturate the uncertainty principle, because:

$$\langle\hat{x}_j^2\rangle_{\alpha_j} - \langle\hat{x}_j\rangle_{\alpha_j}^2 = \langle\left(\frac{\hat{a}_j + \hat{a}_j^\dagger}{\sqrt{2}}\right)^2\rangle_{\alpha_j} - \langle\left(\frac{\hat{a}_j + \hat{a}_j^\dagger}{\sqrt{2}}\right)\rangle_{\alpha_j}^2 = \frac{1}{2} \ ,$$

  and an analogue expression can be found for $\hat{p}_j$.

- They are relatively easy to synthetize in a real laboratory, being a good model for an ideal laser beam. For this reason, they are considered "operationally free"; we will make this statement more precise in the next chapters.

- The set of coherent states is mapped onto itself by a broad class of easily implementable unitaries (see [16, 17] and section 1.2.4). So, non-coherent states (more precisely, states which are not classical mixtures of coherent states) are often considered "hard" to obtain in quantum optical settings (this is not necessarily the case, for instance, in opto-mechanical settings). This idea is the backbone of the Resource Theory of optical non-classicality.

Especially for the classical aspect of their evolution and the minimization of uncertainty relations, coherent states are usually considered the "most classical" among quantum states of a harmonic oscillator.
For later convenience, we define also the **irreducible photon number**, which physically is the part of the energy that is not contained in the first moments:

$$N_0(\rho) := \min_{\boldsymbol{\alpha}\in\mathbb{C}^m} N\left(\mathcal{D}(\boldsymbol{\alpha})\,\rho\,\mathcal{D}(\boldsymbol{\alpha})^\dagger\right) = N(\rho) - \sum_j \left|\mathrm{Tr}[\rho\,\hat{a}_j]\right|^2 . \tag{1.24}$$

The last equality can be proven by simply setting the derivatives in the α_j of the definition of $N_0(\rho)$ to 0.

1.2.3 Gaussian states and gaussian operations

Coherent states are a special case of a more general family of quantum states: **gaussian states** [18, 19, 20, 21]. Gaussian states can be directly defined through the following parametrization [9]:

$$\rho_G = \frac{e^{-\beta H_G}}{\mathrm{Tr}\, e^{-\beta H_G}} , \tag{1.25}$$

where $\beta \in (0, \infty]$ and H_G is a self-adjoint operator fulfilling the two following conditions:

- it is at most quadratic in all the $\hat{a}_j$ and $\hat{a}_j^\dagger$;
- its spectrum is bounded from below, so that $\mathrm{Tr}\, e^{-\beta H_G}$ is finite and the state is normalizable.

Note that also the limiting case $\beta \to \infty$ is encompassed; if we denote with λ_j and $|\lambda_j\rangle$ the eigenvalues and eigenstates of $\hat{H}_G$ respectively, we can write:

$$\begin{aligned}\rho_G &= \lim_{\beta\to\infty} \frac{e^{-\beta\sum_{n=0}^{\infty}\lambda_n|\lambda_n\rangle\langle\lambda_n|}}{\sum_{n=0}^{\infty} e^{-\beta\lambda_n}} \\ &= \lim_{\beta\to\infty} \frac{e^{-\beta\sum_{n=0}^{\infty}(\lambda_n-\lambda_0)|\lambda_n\rangle\langle\lambda_n|}}{1+\sum_{n=1}^{\infty} e^{-\beta(\lambda_n-\lambda_0)}} \\ &= |\lambda_0\rangle\langle\lambda_0| .\end{aligned}$$

So, in other words, gaussian states are all the thermal (for finite β) and ground (otherwise) states of quadratic and bounded from below Hamiltonians. Here we presented a characterisation of gaussian states based on their explicit parametrisation, following the approach suggested in [9]. A more common, yet equivalent, definition of these states can be given in terms of their associated quasi-probability distributions, which we will introduce in the following sections. In particular, in section 1.3.4 a second definition of gaussian states will be provided.

We already encountered some gaussian states. For instance, the vacuum (which is at the same time a gaussian, a coherent and a Fock state) is the ground state of the Hamiltonian (1.14). Thermal states of the same Hamiltonian will be simply referred to as "thermal states", and parametrized as follows (for a single mode):

$$\tau_\nu = \frac{1}{1+\nu}\sum_{n=0}^{\infty}\left(\frac{\nu}{1+\nu}\right)^n |n\rangle\langle n| \,,$$

where $\nu = \frac{1}{e^{\beta\omega}-1}$ is the photon number of the state. Finally, the single-mode coherent state $|\alpha\rangle$ is the ground state of the following Hamiltonian (**driven harmonic oscillator**):

$$\begin{aligned}\mathcal{D}(\alpha)\hat{a}^\dagger\hat{a}\mathcal{D}(\alpha)^\dagger =&\hat{a}^\dagger\hat{a} + [\alpha\hat{a}^\dagger - \alpha^*\hat{a}, \hat{a}^\dagger\hat{a}] + \frac{1}{2}[\alpha\hat{a}^\dagger - \alpha^*\hat{a}, [\alpha\hat{a}^\dagger - \alpha^*\hat{a}, \hat{a}^\dagger\hat{a}]] \\ =&\hat{a}^\dagger\hat{a} - \alpha\hat{a}^\dagger - \alpha^*\hat{a} - \frac{1}{2}[\alpha\hat{a}^\dagger - \alpha^*\hat{a}, \alpha\hat{a}^\dagger + \alpha^*\hat{a}] \\ =&\hat{a}^\dagger\hat{a} - \alpha\hat{a}^\dagger - \alpha^*\hat{a} - |\alpha|^2 \,,\end{aligned}$$

since $\langle\alpha|\mathcal{D}(\alpha)\hat{a}^\dagger\hat{a}\mathcal{D}(\alpha)^\dagger|\alpha\rangle = \langle 0|\hat{a}^\dagger\hat{a}|0\rangle = 0$.
In addition to gaussian states, **gaussian operations** can be defined too. They are those operations which map any gaussian states into gaussian states. Let us consider for the moment gaussian unitaries. By virtue of the Stone's theorem [22], they can always be expressed in an exponential form: $\hat{U} = e^{i\hat{X}}$, with $\hat{X}$ being self-adjoint. By applying the BCH formula it is fairly easy to see that a unitary map is gaussian if and only if $\hat{X}$ is at most quadratic in all the $\hat{a}_j$ and $\hat{a}_j^\dagger$ (otherwise, higher order terms would appear), a condition which resembles the definition of gaussian states themselves. For an exhaustive treatment of gaussian operations, see for example [9]; we will focus mostly on a subset of gaussian operations, which play a privileged role in the theory of optical non-classicality.

1.2.4 Linear optical operations

A very special set of gaussian unitaries in a quantum optical setting are **linear optical** (**LO**) unitaries, which we are going to define throughout this section. An important subset of these maps are **passive linear** (**PL**) unitaries. As the name suggest, they are implemented without any external source of energy, and hence they preserve the total number of particles:

$$\begin{aligned}U^\dagger \sum_{j=1}^{m}\hat{a}_j^\dagger\hat{a}_j U &= \sum_{j=1}^{m}\hat{a}_j^\dagger\hat{a}_j \\ U^\dagger \sum_{j=1}^{m}\hat{a}_j^\dagger U U^\dagger\hat{a}_j U &= \sum_{j=1}^{m}\hat{a}_j^\dagger\hat{a}_j \\ \left(U^\dagger\hat{\boldsymbol{a}}U\right)^\dagger \cdot \left(U^\dagger\hat{\boldsymbol{a}}U\right) &= \hat{\boldsymbol{a}}^\dagger \cdot \hat{\boldsymbol{a}}\,,\end{aligned} \tag{1.26}$$

where we defined the vector of operators $\hat{\boldsymbol{a}} = (\hat{a}_1, \ldots, \hat{a}_m)^T$. Clearly, for the last equation to be satisfied, we must have:

$$U^\dagger\hat{\boldsymbol{a}}U = \mathbb{U}\hat{\boldsymbol{a}}\,, \tag{1.27}$$

where $\mathbb{U}$ is a unitary $m \times m$ matrix and must not be confused with U, which is an infinite dimensional operator. We cannot stress enough the fact that U and $\mathbb{U}$ are two completely different objects: the former acts on $\hat{\boldsymbol{a}}$ at the level of Hilbert space; the latter "sees" $\hat{\boldsymbol{a}}$ just as a vector. Furthermore, it is important to distinguish between the $|\boldsymbol{\alpha}_j\rangle$'s, which always denote coherent states as elements of $L^2(\mathbb{R}^m)$, and the $\boldsymbol{\alpha_j}$'s, which are instead m-uples of complex numbers and elements of $\mathbb{C}^m$.
At this point, one may wonder whether it is possible to find some elementary PL unitaries, from which any other can be derived by composition. Composition of PL unitaries results in the composition of unitary matrices acting on $\hat{\boldsymbol{a}}$; the question then becomes: which are the elementary building blocks in which we can decompose any unitary matrix? The answer is quite simple: any $m \times m$ unitary matrix can be written as the product of 2×2 unitary matrices (completed to the identity on the space of $m \times m$ matrices) [23], and any 2×2 unitary matrix can be written in terms of phase-multiplications and rotations:

$$\mathbb{U}_2 = e^{i\phi_1} \begin{pmatrix} e^{i\phi_2} & 0 \\ 0 & e^{-i\phi_2} \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} e^{i\phi_3} & 0 \\ 0 & e^{-i\phi_3} \end{pmatrix} .$$

Multiplications by a phase are achieved with **phase-shifters**:

$$\begin{aligned} e^{i\varphi \hat{a}^\dagger \hat{a}} \hat{a} e^{-i\varphi \hat{a}^\dagger \hat{a}} &= \sum_{n=0}^{\infty} \frac{(i\varphi)^n}{n!} [\hat{a}^\dagger \hat{a}, \hat{a}]_n \\ &= \sum_{n=0}^{\infty} \frac{(-i\varphi)^n}{n!} \hat{a} \\ &= e^{-i\varphi} \hat{a} , \end{aligned}$$

where we used the fact that $[\hat{a}^\dagger \hat{a}, \hat{a}] = -\hat{a}$. Instead, we obtain rotations by means of **beam-splitters**:

$$\begin{aligned} e^{-\theta(\hat{a}_1^\dagger \hat{a}_2 - \hat{a}_1 \hat{a}_2^\dagger)} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix} e^{\theta(\hat{a}_1^\dagger \hat{a}_2 - \hat{a}_1 \hat{a}_2^\dagger)} &= \sum_{n=0}^{\infty} \frac{(-\theta)^n}{n!} \begin{pmatrix} [\hat{a}_1^\dagger \hat{a}_2 - \hat{a}_1 \hat{a}_2^\dagger, \hat{a}_1]_n \\ [\hat{a}_1^\dagger \hat{a}_2 - \hat{a}_1 \hat{a}_2^\dagger, \hat{a}_2]_n \end{pmatrix} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n \theta^{2n}}{(2n)!} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix} + \sum_{n=0}^{\infty} \frac{(-1)^n \theta^{(2n+1)}}{(2n+1)!} \begin{pmatrix} \hat{a}_2 \\ -\hat{a}_1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta \hat{a}_1 + \sin\theta \hat{a}_2 \\ \cos\theta \hat{a}_2 - \sin\theta \hat{a}_1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix} , \end{aligned}$$

where we used the fact that $[\hat{a}_1^\dagger \hat{a}_2 - \hat{a}_1 \hat{a}_2^\dagger, \hat{a}_1] = -\hat{a}_2$ and $[\hat{a}_1^\dagger \hat{a}_2 - \hat{a}_1 \hat{a}_2^\dagger, \hat{a}_2] = \hat{a}_1$. We have not only proved that any PL unitary can be obtained by composition of phase-shifters and beam-splitters, but we also proved that any unitary matrix $\mathbb{U}$ acting on $\hat{\boldsymbol{a}}$ one can dream of can be obtained with a PL unitary. We summarize these results in the following proposition (see also [24]).
Proposition 1.8. *Any transformation of the type $\hat{\boldsymbol{a}} \mapsto \mathbb{U}\hat{\boldsymbol{a}}$ can be obtained with a passive linear unitary, and any such transformation can be decomposed as a product of only phase-shifters and beam-splitters.*

Now it is also straightforward to show how PL unitaries act on coherent states:

$$\hat{\boldsymbol{a}} U_{PL} |\boldsymbol{\alpha}\rangle = U_{PL} U_{PL}^{\dagger} \hat{\boldsymbol{a}} U_{PL} |\boldsymbol{\alpha}\rangle = U_{PL} \mathbb{U} \hat{\boldsymbol{a}} |\boldsymbol{\alpha}\rangle = \mathbb{U} \boldsymbol{\alpha} U_{PL} |\boldsymbol{\alpha}\rangle \ ; \tag{1.28}$$

thus, $U_{PL} |\boldsymbol{\alpha}\rangle = |\mathbb{U}\boldsymbol{\alpha}\rangle$. If we add also displacements, we can transform coherent states as follows:

$$|\boldsymbol{\alpha}\rangle \mapsto e^{i\phi(\boldsymbol{\alpha})} |\mathbb{U}\boldsymbol{\alpha} + \boldsymbol{\alpha}_0\rangle \ .$$

So, compositions of phase-shifters, beam-splitters and displacements map coherent states into coherent states. Together, they form the aforementioned class of linear optical unitaries.

As we will prove in Proposition 1.11, the expression for $\phi(\boldsymbol{\alpha})$ is fixed, up to an irrelevant constant, by unitarity of U_{PL}: $\phi(\boldsymbol{\alpha}) = \phi_0 - \frac{1}{2i}(\langle \boldsymbol{\alpha}_0, U\boldsymbol{\alpha}\rangle - \langle U\boldsymbol{\alpha}, \boldsymbol{\alpha}_0\rangle)$.

Remark 1.9. *Obviously such an expression for $\phi(\boldsymbol{\alpha})$ can always be achieved by means of PL unitaries and displacements, for example by building U with only PL unitaries and applying the displacement $\mathcal{D}(\boldsymbol{\alpha}_0)$ only at the end. So, the most general expression for a LO unitary is the following:*

$$|\boldsymbol{\alpha}\rangle \mapsto e^{i\phi_0} e^{\frac{1}{2}(\langle U\boldsymbol{\alpha}, \boldsymbol{\alpha}_0\rangle - \langle \boldsymbol{\alpha}_0, U\boldsymbol{\alpha}\rangle)} |U\boldsymbol{\alpha} + \boldsymbol{\alpha}_0\rangle \ . \tag{1.29}$$

Are there other unitaries, apart from LO ones, which preserve the set of coherent states? The following results answer to this question.

Lemma 1.10. *Let $\boldsymbol{f} : \mathbb{C}^n \to \mathbb{C}^n$ be such that $\langle \boldsymbol{f}(\boldsymbol{\alpha}) - \boldsymbol{f}(\boldsymbol{0}), \boldsymbol{f}(\boldsymbol{\beta}) - \boldsymbol{f}(\boldsymbol{0})\rangle = \langle \boldsymbol{\alpha}, \boldsymbol{\beta}\rangle$ for any $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. Then $\boldsymbol{f}$ acts as*

$$\boldsymbol{f}(\boldsymbol{\alpha}) = \mathbb{U}\boldsymbol{\alpha} + \boldsymbol{f}(\boldsymbol{0})$$

for any $\boldsymbol{\alpha} \in \mathbb{C}^n$, where $\mathbb{U}$ is unitary. In other words, $\boldsymbol{f}$ is a unitary affine map on $\mathbb{C}^n$.

Proof. We start by defining $\boldsymbol{g}(\boldsymbol{\alpha}) \coloneqq \boldsymbol{f}(\boldsymbol{\alpha}) - \boldsymbol{f}(\boldsymbol{0})$. Let $\{\boldsymbol{\alpha}_j\}_j$ be an arbitrary orthonormal basis for $\mathbb{C}^n$: then $\langle \boldsymbol{g}(\boldsymbol{\alpha}_j), \boldsymbol{g}(\boldsymbol{\alpha}_k)\rangle = \langle \boldsymbol{\alpha}_j, \boldsymbol{\alpha}_k\rangle$, which implies that $\{\boldsymbol{g}(\boldsymbol{\alpha}_j)\}_j$ is an orthonormal basis as well; we cannot conclude at this point because we do not know a priori whether $\boldsymbol{g}$ is linear or not.

Now we define the matrix $\mathbb{U}$ such that $\mathbb{U}^\dagger \boldsymbol{g}(\boldsymbol{\alpha}_j) = \boldsymbol{\alpha}_j$ for any j: this also ensures that $\mathbb{U}$ is unitary. Hence, for an arbitrary $\boldsymbol{\alpha}$ we have:

$$\langle \mathbb{U}^\dagger \boldsymbol{g}(\boldsymbol{\alpha}), \boldsymbol{\alpha}_j\rangle = \langle \mathbb{U}^\dagger \boldsymbol{g}(\boldsymbol{\alpha}), \mathbb{U}^\dagger \boldsymbol{g}(\boldsymbol{\alpha}_j)\rangle = \langle \boldsymbol{g}(\boldsymbol{\alpha}), \boldsymbol{g}(\boldsymbol{\alpha}_j)\rangle = \langle \boldsymbol{\alpha}, \boldsymbol{\alpha}_j\rangle \ .$$

Being true for all the elements of a basis, we have necessarily $\mathbb{U}^\dagger \boldsymbol{g}(\boldsymbol{\alpha}) = \boldsymbol{\alpha}$ for any $\boldsymbol{\alpha}$, so $\boldsymbol{g}$ is indeed linear (and unitary) after all. This conclude the proof. □

The previous result resembles the famous Wigner's theorem [25] applied to $\mathbb{C}^n$, but with the big difference that we require also the phase of scalar products to be preserved, and not only the modulus. This rules out antiunitary transformations.

Proposition 1.11. *Linear optical unitaries are the only unitaries which map any coherent state into another coherent state.*

Proof. Let U be a unitary mapping coherent states into coherent states. Then, for any $\boldsymbol{\alpha}$:

$$U |\boldsymbol{\alpha}\rangle = e^{i\phi(\boldsymbol{\alpha})} |\boldsymbol{f}(\boldsymbol{\alpha})\rangle \ ,$$

for some maps $\boldsymbol{f} : \mathbb{C}^n \to \mathbb{C}^n$ and $\phi : \mathbb{C}^n \to \mathbb{R}$. Now we have

$$\begin{aligned} e^{-\frac{|\alpha|^2}{2}} &= \langle \mathbf{0}|\boldsymbol{\alpha}\rangle \\ &= \langle \mathbf{0}|U^\dagger U|\boldsymbol{\alpha}\rangle \\ &= e^{i[\phi(\boldsymbol{\alpha})-\phi(\mathbf{0})]}\,\langle \boldsymbol{f}(\mathbf{0})|\boldsymbol{f}(\boldsymbol{\alpha})\rangle \\ &= e^{i[\phi(\boldsymbol{\alpha})-\phi(\mathbf{0})]} e^{\frac{1}{2}(\langle \boldsymbol{f}(\mathbf{0}),\boldsymbol{f}(\boldsymbol{\alpha})\rangle - \langle \boldsymbol{f}(\boldsymbol{\alpha}),\boldsymbol{f}(\mathbf{0})\rangle)} e^{-\frac{|\boldsymbol{f}(\boldsymbol{\alpha})-\boldsymbol{f}(\mathbf{0})|^2}{2}}\ , \end{aligned}$$

which implies $|\boldsymbol{f}(\boldsymbol{\alpha})-\boldsymbol{f}(\mathbf{0})|^2 = |\boldsymbol{\alpha}|^2$ and $\phi(\boldsymbol{\alpha})-\phi(\mathbf{0}) = -\frac{1}{2i}(\langle \boldsymbol{f}(\mathbf{0}), \boldsymbol{f}(\boldsymbol{\alpha})\rangle - \langle \boldsymbol{f}(\boldsymbol{\alpha}), \boldsymbol{f}(\mathbf{0})\rangle)$. We also have:

$$\begin{aligned} \langle \boldsymbol{\alpha}|\boldsymbol{\beta}\rangle &= \langle \boldsymbol{\alpha}|U^\dagger U|\boldsymbol{\beta}\rangle \\ &= e^{\frac{1}{2}(\langle \boldsymbol{f}(\mathbf{0}),\boldsymbol{f}(\boldsymbol{\alpha})-\boldsymbol{f}(\boldsymbol{\beta})\rangle - \langle \boldsymbol{f}(\boldsymbol{\alpha})-\boldsymbol{f}(\boldsymbol{\beta}),\boldsymbol{f}(\mathbf{0})\rangle)}\,\langle \boldsymbol{f}(\boldsymbol{\alpha})|\boldsymbol{f}(\boldsymbol{\beta})\rangle \\ &= \langle \boldsymbol{f}(\boldsymbol{\alpha}) - \boldsymbol{f}(\mathbf{0})|\boldsymbol{f}(\boldsymbol{\beta}) - \boldsymbol{f}(\mathbf{0})\rangle\ , \end{aligned}$$

which instead implies $\langle \boldsymbol{\alpha}, \boldsymbol{\beta}\rangle = \langle \boldsymbol{f}(\boldsymbol{\alpha}) - \boldsymbol{f}(\mathbf{0}), \boldsymbol{f}(\boldsymbol{\beta}) - \boldsymbol{f}(\mathbf{0})\rangle$. By virtue of Remark 1.9 and Lemma 1.10, we prove the claim. □

For the sake of completeness, and for later convenience, we mention that by adding just ont type of unitaries to LO ones, one can obtain all gaussian unitaries. The missing piece is represented by (single-mode) **squeezing unitaries**:

$$\hat{S}(\zeta) := e^{\zeta^* \hat{a}^2 - \zeta \hat{a}^{\dagger 2}}\ . \tag{1.30}$$

When applied to the vacuum, they give **squeezed states**:

$$|\psi_{r,\phi}\rangle = \frac{1}{\sqrt{\cosh r}} \sum_{n=0}^{\infty} \sqrt{\binom{2n}{n}} \left(-\frac{1}{2}\, e^{i\phi} \tanh r\right)^n |2n\rangle\ , \tag{1.31}$$

with $\zeta = re^{i\phi}$. Hence, squeezing unitaries does not map coherent states into coherent states.

We end this section by briefly discussing the actual physical implementations of these operations. As we already anticipated, they can be achieved with readily-accessible optical elements [26]. Arguably, the simplest ones are phase-shifters: they are obtained via dielectric plates, which change the dielectric constant of the medium in which the light is travelling, and hence the velocity of rotation of coherent states. The net effect, with respect to the "rotating frame" which already accounts for the evolution in vacuum, is a discrete rotation of any coherent state, independent from α. Beam-splitter are obtained via semi-reflective mirrors, with reflectivity r and transmittivity t, where different modes are mixed. Finally, displacements can be implemented with a low-transmittivity semi-reflective mirror and a high-intensity laser source [27]: indeed

$$\hat{U}_{BS}\,|\alpha\alpha'\rangle = |(\cos\theta\alpha + \sin\theta\alpha')(\cos\theta\alpha' - \sin\theta\alpha)\rangle \approx |(\alpha + \alpha_0)(\alpha' - \epsilon\alpha)\rangle\ ,$$

where we set $\theta = \epsilon$ and $\alpha_0 = \epsilon\alpha'$.

1.3 Phase-space quantisation

1.3.1 The Fourier-Weyl transform

From now on, we will focus on a single electromagnetic mode, unless stated otherwise. Displacement operators fulfill the following relation:

$$\begin{aligned}\mathrm{Tr}\left(\mathcal{D}(\alpha)^\dagger\mathcal{D}(\beta)\right) =& \frac{1}{\pi}\int d^2\gamma\ \langle\gamma|\mathcal{D}(\alpha)^\dagger\mathcal{D}(\beta)|\gamma\rangle \\ =& \frac{1}{\pi}\int d^2\gamma\ \langle 0|\mathcal{D}(\gamma)^\dagger\mathcal{D}(\alpha)^\dagger\mathcal{D}(\gamma)\mathcal{D}(\gamma)^\dagger\mathcal{D}(\beta)\mathcal{D}(\gamma)|0\rangle \\ \stackrel{1}{=}& \frac{1}{\pi}\int d^2\gamma\, e^{\frac{1}{2}[(\beta-\alpha)\gamma^*-\gamma(\beta-\alpha)^*]}\ \langle\alpha|\beta\rangle \\ =& \pi\delta^2(\beta-\alpha)\,,\end{aligned} \tag{1.32}$$

where in 1 we used (1.19). The expression we derived is formally analogue to the orthonormality condition of Fourier oscillatory factors. So, we might be enticed to define an "operatorial Fourier transform": this can indeed be done (at least for bounded operators), and it takes the name of **Fourier-Weyl transform** [9, 28]:

$$O = \frac{1}{\pi}\int d^2\alpha\,\mathrm{Tr}\left(O\mathcal{D}(\alpha)\right)\mathcal{D}(\alpha)^\dagger\,. \tag{1.33}$$

The Fourier-Weyl transform of a density operator ρ is called the **characteristic function** of ρ, and denoted with $\chi_0^\rho(\alpha)$ or simply $\chi^\rho(\alpha)$. It obeys the following basic properties:

- $\chi^\rho(0) = \mathrm{Tr}\,\rho = 1$;
- $\chi^\rho(\alpha)^* = \mathrm{Tr}\left(\rho\mathcal{D}(\alpha)\right) = \mathrm{Tr}\left(\rho\mathcal{D}(\alpha)^\dagger\right) = \chi^\rho(-\alpha)$;
- $|\chi^\rho(\alpha)| \leq 1$, since $\mathcal{D}(\alpha)$ is unitary.

Since it depends on both position and momentum at the same time, the procedure of describing a quantum state via its characteristic function is known as **phase-space quantisation**.

Characteristic functions can also be used in order to compute the **Hilbert-Schmidt** scalar product between two bounded operators:

$$\mathrm{Tr}\left(O_1^\dagger O_2\right) = \frac{1}{\pi^2}\int d^2\alpha d^2\beta\,\chi^1(\alpha)^*\chi^2(\beta)\,\mathrm{Tr}\left(\mathcal{D}(\alpha)^\dagger\mathcal{D}(\beta)\right) = \frac{1}{\pi}\int d^2\alpha\,\chi^1(\alpha)^*\chi^2(\alpha)\,. \tag{1.34}$$

In particular:

$$1 \geq \mathrm{Tr}(\rho^\dagger\rho) = \frac{1}{\pi}\int d^2\alpha\,|\chi^\rho(\alpha)|^2\,, \tag{1.35}$$

so that $\chi^\rho \in L^2(\mathbb{C})$.

1.3.2 Characteristic functions

Starting from the characteristic function of a density operator we can define a whole family of related objects: the **s-ordered characteristic functions**, defined as

$$\chi_s^\rho(\alpha) = e^{s\frac{|\alpha|^2}{2}}\chi^\rho(\alpha) = e^{s\frac{|\alpha|^2}{2}}\,\mathrm{Tr}\left(\rho\mathcal{D}(\alpha)\right)\,, \qquad -1 \leq s \leq 1\,. \tag{1.36}$$

It is immediate to see that for any s it still holds that $\chi_s^\rho(0) = 1$ and $\chi_s^\rho(\alpha)^* = \chi_s^\rho(-\alpha)$, while $\chi_s^\rho(\alpha)$ might not be bounded for a strictly positive s. The name "characteristic function" is borrowed from probability theory, where characteristic functions can be used in order to efficiently compute all the momenta of a probability distribution. Even though s-ordered characteristic functions have indeed a similar property, as we are going to show in a moment, they are not proper characteristic functions, as their symplectic Fourier transforms are not proper probability distributions.
From (1.16) it is easy to see that:

$$\left(\frac{\partial^n}{\partial\alpha^n}\right)_{|\alpha=0}\left(-\frac{\partial^m}{\partial\alpha^{*m}}\right)_{|\alpha=0}\chi_1^\rho(\alpha) = \mathrm{Tr}\left(\rho\hat{a}^{\dagger n}\hat{a}^m\right) = \langle\hat{a}^{\dagger n}\hat{a}^m\rangle_1\,,$$

where $\langle\cdot\rangle_1$ denotes the **normal-ordered expectation value** of an operator. Similarly, it is easy to show that:

$$\left(\frac{\partial^n}{\partial\alpha^n}\right)_{|\alpha=0}\left(-\frac{\partial^m}{\partial\alpha^{*m}}\right)_{|\alpha=0}\chi_{-1}^\rho(\alpha) = \langle\hat{a}^{\dagger n}\hat{a}^m\rangle_{-1}\,,$$

where $\langle\cdot\rangle_{-1}$ denotes the **anti-normal-ordered expectation value** of an operator. Finally, it holds:

$$\begin{aligned}\left(\frac{\partial^n}{\partial\alpha^n}\right)_{|\alpha=0}\left(-\frac{\partial^m}{\partial\alpha^{*m}}\right)_{|\alpha=0}\chi_0^\rho(\alpha) =& \mathrm{Tr}\left[\rho\left(\frac{\partial^n}{\partial\alpha^n}\right)_{|\alpha=0}\left(-\frac{\partial^m}{\partial\alpha^{*m}}\right)_{|\alpha=0}\sum_{j=0}^{\infty}\frac{(\alpha\hat{a}^\dagger-\alpha^*\hat{a})^j}{j!}\right]\\ =&\langle\hat{a}^{\dagger n}\hat{a}^m\rangle_0\,,\end{aligned}$$

where $\langle\cdot\rangle_0$ denotes the normalized **symmetrically-ordered expectation value**, obtained by summing over all possible re-orderings of the operators and dividing by $\binom{n+m}{n}$. To obtain the last expression it is sufficient to note that the only terms which survive after the derivation are those in which $\hat{a}^\dagger$ appears exactly n times and $\hat{a}$ m times. The normalization and symmetric ordering in the result come from the explicit expression of $(\alpha\hat{a}^\dagger - \alpha^*\hat{a})^{n+m}$. More in general, we can just define the s-**ordered expectation value** as

$$\langle\hat{a}^{\dagger n}\hat{a}^m\rangle_s := \left(\frac{\partial^n}{\partial\alpha^n}\right)_{|\alpha=0}\left(-\frac{\partial}{\partial\alpha^{*m}}\right)_{|\alpha=0}\chi_s^\rho(\alpha)\,.$$

1.3.3 Quasi-probability distributions

We will often consider also the symplectic Fourier transform of the s-ordered characteristic functions: the **s-ordered quasi-probability distributions**:

$$W_s^\rho(\beta) = \frac{1}{\pi^2}\int d^2\alpha\, e^{\beta\alpha^*-\alpha\beta^*}\chi_s^\rho(\alpha)\,. \tag{1.37}$$

Also $W_s^\rho(\alpha)$ can be used to compute the s-ordered expectation values of $\hat{a}$ and $\hat{a}^\dagger$. Thanks to the properties of the Fourier transform we have indeed:

$$\langle\hat{a}^{\dagger n}\hat{a}^m\rangle_s = \int d^2\beta\,\beta^{*n}\beta^m W_s^\rho(\beta)\,.$$

Quasi-probability distribution, just like standard probability distributions, are normalized to 1:

$$\int d^2\beta W_s^\rho(\beta) = \chi_s^\rho(0) = 1\,;$$

however, they can be negative (as functions or as distributions, as we will see) at some points in phase space. Negativity of quasi-probability distribtions is often considered as a signature (or even as a "measure", in a sense that we will make more precise in the next chapters) of "quantumness". At any rate, the fact that they cannot be interpreted as actual probability distributions (even when they are positive at every point!) can be seen as a reflection of the uncertainty principle: we cannot associate to a quantum state a probability distribution of position and momentum at the same time.
However, we can obtain actual probability distributions from the W_s^ρ by marginalization. Let us consider W_0^ρ, which we will denote simply with W^ρ from now on. It was introduced by Wigner himself [29], and hence it is also known as the **Wigner's function** of the state ρ. Parametrizing α as $\alpha = \frac{x+ip}{\sqrt{2}}$ and β as $\beta = \frac{x'+ip'}{\sqrt{2}}$, we can prove the following relation (see [9] for a slightly different derivation):

$$\begin{aligned}\int dp\, W^\rho(x,p) &= \int dp \int \frac{dx'dp'}{2\pi^2}\, e^{i(px'-xp')}\chi^\rho(x',p')\\ &= \int \frac{dx'dp'}{2\pi^2} \int dp\, e^{ipx'}e^{-ixp'}\, \mathrm{Tr}\left(\rho e^{i(p'\hat{x}-x'\hat{p})}\right)\\ &\stackrel{1}{=} \int \frac{dp'}{\pi}\, e^{-ixp'}\, \mathrm{Tr}\left(\rho e^{ip'\hat{x}}\right)\\ &\stackrel{2}{=} \int \frac{dp'}{\pi} e^{-ixp'} \int dq\, \langle q|\rho e^{ip'\hat{x}}|q\rangle\\ &= \int dq \int \frac{dp'}{\pi}\, e^{i(x-q)p'}\, \langle q|\rho|q\rangle\\ &\stackrel{3}{=} 2\,\langle x|\rho|x\rangle\ ,\end{aligned}$$

where in 1 we used the integral representation of $\delta(x')$, in 2 we expressed the trace with the eigenstates $|q\rangle$ of the position operator, and in 3 we used the integral representation of $\delta(x-q)$. If we define the self-adjoint operators $\hat{x}_\phi = \cos\phi\hat{x}+\sin\phi\hat{p}$ and the corresponding eigenstates $|x_\phi\rangle$, we can repeat the above procedure and prove the more general formula below:

$$\frac{1}{2}\int x_{\phi+\frac{\pi}{2}}\, W^\rho(x,p) = \langle x_\phi|\rho|x_\phi\rangle\ . \tag{1.38}$$

In this sense, marginalizations of W^ρ give proper probability distributions of observables. By means of equation (1.17), we can rewrite $\chi^\rho(\alpha)$ as:

$$\begin{aligned}\chi^\rho(\alpha) &= \frac{e^{\frac{|\alpha|^2}{2}}}{\pi} \int d^2\beta\ \langle\beta|\rho e^{-\alpha^* a}e^{\alpha a^\dagger}|\beta\rangle\\ &= \frac{e^{\frac{|\alpha|^2}{2}}}{\pi} \int d^2\alpha\ \langle\beta|e^{\alpha a^\dagger}\rho e^{-\alpha^* a}|\beta\rangle\\ &= \frac{e^{\frac{|\alpha|^2}{2}}}{\pi} \int d^2\alpha\, e^{\alpha\beta^*-\alpha^*\beta}\,\langle\beta|\rho|\beta\rangle\ ,\end{aligned} \tag{1.39}$$

which implies that $W^{\rho}_{-1}(\beta) = \frac{1}{\pi}\langle\beta|\rho|\beta\rangle$. This quasi-probability distribution is knows also as the **Husimi Q-function**, and often denoted with Q^{ρ}. The boundedness of χ^{ρ} implies that $\chi^{\rho}_{-1} \in L^1(\mathbb{C})$, which in turns implies that Q^{ρ} is a continuous function. Being continuous and positive everywhere, it is the "most regular" of all quasi-probability distributions: for this reason, it is widely used in Quantum Optics and Quantum Information, as we will see in the next section.
On the other hand, the "wildest" of all quasi-probability distributions is W^{ρ}_1, which is also known as **Glauber-Sudarshan P-function** and denoted with P^{ρ}. It is easy to see why this is the case: loosely speaking, the regularity of a function is correlated to the asymptotic behaviour of its Fourier transform, and by definition χ^{ρ}_1 is the "most divergent" of all s-characteristic functions. To be precise, W^{ρ}_1 might not even exist not only as a function, but also as a tempered distribution, as χ^{ρ}_1 might diverge exponentially. For this reason, we will try to avoid using W^{ρ}_1 directly, and we will do it only when it is a well-defined object. However, when P can actually be constructed, it satisfies a very important property; in order to prove it, let us compute the characteristic function for a coherent state $|\alpha_0\rangle$:

$$\chi^{\alpha_0}(\alpha) = \langle\alpha_0|\mathcal{D}(\alpha)|\alpha_0\rangle = e^{\alpha\alpha_0^* - \alpha_0\alpha^*} e^{-\frac{|\alpha|^2}{2}}\,, \tag{1.40}$$

from which we derive the corresponding P-function:

$$P^{\alpha_0}(\beta) = \frac{1}{\pi^2}\int d^2\alpha\, e^{\alpha\alpha_0^* - \alpha_0\alpha^*} = \delta^2(\beta - \alpha_0)\,. \tag{1.41}$$

Being the Fourier transform linear, we conclude that, when P^{ρ} exists, it is possible to write

$$\rho = \int d^2\alpha\, P^{\rho}(\alpha)\,|\alpha\rangle\!\langle\alpha|\;. \tag{1.42}$$

1.3.4 Some examples

For later convenience, we perform the explicit computation of the s-ordered characteristic functions and quasi-probability distributions for some relevant quantum states. We already computed the characteristic function for a coherent state in (1.40), from which we also see that all of its characteristic functions and quasi-probability distributions are (possibly degenerate) gaussians. More in general, it is easy to see from (1.25) that a state is gaussian if and only if it has gaussian characteristic functions and quasi-probability distributions. Actually, this is often taken to be the definition of a gaussian state, and it will be helpful in the future. More precisely, if we define the **quantum covariance matrix** of a state ρ as

$$\mathbb{V}_{jk} = \frac{1}{2}\,\mathrm{Tr}[\rho\{\hat{r}_j - \mathrm{Tr}(\rho\hat{r}_j), \hat{r}_k - \mathrm{Tr}(\rho\hat{r}_k)\}]\,, \tag{1.43}$$

the Q-function of a gaussian state is a gaussian with classical covariance matrix equal to $\mathbb{V} + \frac{1}{2}\mathbb{I}$.

Let us consider for example a thermal state τ_ν. We have:

$$\begin{aligned} Q^\nu(\alpha) =& \frac{1}{\pi} \langle\alpha|\tau_\nu|\alpha\rangle \\ =& \frac{1}{\pi(1+\nu)} \sum_{n=0}^{\infty} \left(\frac{\nu}{1+\nu}\right)^n |\langle n|\alpha\rangle|^2 \\ =& \frac{e^{-|\alpha|^2}}{\pi(1+\nu)} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{\nu|\alpha|^2}{1+\nu}\right)^n \\ =& \frac{1}{\pi(1+\nu)} e^{-\frac{|\alpha|^2}{1+\nu}} . \end{aligned} \tag{1.44}$$

Its characteristic function is then:

$$\chi^\nu(\alpha) = e^{-\left(\frac{1}{2}+\nu\right)|\alpha|^2} , \tag{1.45}$$

while its P-function is:

$$P^\nu(\alpha) = \frac{1}{\pi\nu} e^{-\frac{|\alpha|^2}{\nu}} . \tag{1.46}$$

Another example of a gaussian state is the aforementioned squeezed state, defined in (1.31). Again, we start from its Q-function:

$$\begin{aligned} Q^{r,\phi} =& \frac{1}{\pi} |\langle\alpha|\psi_{r,\phi}\rangle|^2 \\ =& \frac{1}{\pi} \left| \frac{1}{\sqrt{\cosh r}} \sum_{n=0}^{\infty} \frac{\cancel{\sqrt{(2n)!}}}{n!} \left(-\frac{1}{2} e^{i\phi} \tanh r\right)^n \frac{e^{-\frac{|\alpha|^2}{2}} \alpha^{*2n}}{\cancel{\sqrt{(2n)!}}} \right|^2 \\ =& \frac{1}{\pi \cosh r} e^{-|\alpha|^2} \left| e^{-\frac{1}{2} e^{i\phi} \alpha^{*2} \tanh r} \right|^2 \\ =& \frac{1}{\pi \cosh r} e^{-|\alpha|^2} e^{-\Re[e^{i\phi}\alpha^{*2}] \tanh r} \\ =& \frac{1}{\pi \cosh r} e^{-(1+\cos\phi \tanh r)\Re[\alpha]^2 - (1-\cos\phi\tanh r)\Im[\alpha]^2 - 2\sin\phi\cos\phi\Re[\alpha]\Im[\alpha]} . \end{aligned} \tag{1.47}$$

From this result it is easy to obtain the characteristic functions (we just have to Fourier transform a gaussian function) and see that $\chi_1^{r,\phi}(\alpha)$ is exponentially divergent for any $r > 0$: as a consequence, $P^{r,\phi}$ is not even a tempered distribution.
Let us consider now a single-mode Fock state $|n\rangle$. We have:

$$\begin{aligned} \chi^n(\alpha) =& e^{-\frac{|\alpha|^2}{2}} \langle n|e^{\alpha\hat{a}^\dagger} e^{-\alpha^*\hat{a}}|n\rangle \\ =& e^{-\frac{|\alpha|^2}{2}} \langle n| \left(\sum_{m'=0}^{\infty} \frac{\alpha^{m'} \hat{a}^{\dagger m'}}{m'!}\right) \left(\sum_{m=0}^{\infty} \frac{(-\alpha^*)^m \hat{a}^m}{m!}\right) |n\rangle \\ =& e^{-\frac{|\alpha|^2}{2}} \left(\sum_{m'=0}^{n} \frac{\alpha^{m'} \langle n-m'|}{\sqrt{m'!}}\right) \left(\sum_{m=0}^{n} \frac{(-\alpha^*)^m |n-m\rangle}{\sqrt{m!}}\right) \\ & e^{-\frac{|\alpha|^2}{2}} \sum_{m=0}^{n} \frac{(-1)^m |\alpha|^{2m}}{m!} , \end{aligned} \tag{1.48}$$

from which we derive

$$\chi_1^n(\alpha) = \sum_{m=0}^{n} \frac{(-1)^m |\alpha|^{2m}}{m!} \implies P^n(\alpha) = \pi^2 \sum_{m=0}^{n} \frac{\partial_\alpha^m \partial_{\alpha^*}^m \delta^2}{m!} . \tag{1.49}$$

This highly-singular P-function can only be written in terms of derivatives of a δ function. The Q-function is:

$$Q^n(\alpha) = \frac{1}{\pi} | \langle \alpha | n \rangle |^2 = \frac{e^{-|\alpha|^2} |\alpha|^{2n}}{\pi n!} . \tag{1.50}$$

The last quantum state we consider is the **cat state**, whose name is a tribute to the famous dead-and-alive Schrödinger's cat. Indeed, it consists of a superposition of two coherent states with opposite amplitude:

$$|\psi_\alpha\rangle := \frac{|\alpha\rangle + |-\alpha\rangle}{\sqrt{2 + 2e^{-2|\alpha|^2}}} \tag{1.51}$$

The weird normalization factor is due to the non-orthonomality of the coherent states. We start from the characteristic function:

$$\begin{aligned} \chi^{\psi_{\alpha_0}}(\alpha) &= \langle \psi_{\alpha_0} | \mathcal{D}(\alpha) | \psi_{\alpha_0} \rangle \\ &= e^{-\frac{|\alpha|^2}{2}} \langle \psi_{\alpha_0} | e^{\alpha \hat{a}^\dagger} e^{-\alpha^* \hat{a}} | \psi_{\alpha_0} \rangle \\ &= \frac{e^{-\frac{|\alpha|^2}{2}}}{2 + 2e^{-2|\alpha_0|^2}} \left(e^{\alpha \alpha_0^*} \langle \alpha_0 | + e^{-\alpha \alpha_0^*} \langle -\alpha_0 | \right) \left(e^{-\alpha^* \alpha_0} |\alpha_0\rangle + e^{\alpha^* \alpha_0} |-\alpha_0\rangle \right) \\ &= \frac{e^{-\frac{|\alpha|^2}{2}}}{1 + e^{-2|\alpha_0|^2}} \left(\cos(2\Im[\alpha \alpha_0^*]) + e^{-2|\alpha_0|^2} \cosh(2\Re[\alpha \alpha_0^*]) \right) . \end{aligned} \tag{1.52}$$

Hence, also $\chi_1^{\psi_{\alpha_0}}$ is exponentially divergent. The Q-function is:

$$\begin{aligned} Q^{\psi_{\alpha_0}}(\alpha) &= \frac{1}{\pi} | \langle \alpha | \psi_{\alpha_0} \rangle |^2 \\ &= \frac{e^{-|\alpha|^2} e^{-|\alpha_0|^2}}{\pi(2 + 2e^{-|\alpha_0|^2})} \left| e^{\alpha^* \alpha_0} + e^{-\alpha^* \alpha_0} \right|^2 \\ &= \frac{2e^{-|\alpha|^2} e^{-|\alpha_0|^2}}{\pi(1 + e^{-|\alpha_0|^2})} |\cosh(\alpha^* \alpha_0)|^2 . \end{aligned} \tag{1.53}$$

1.4 Quantum distances and quantum entropies

1.4.1 Trace distance

A fundamental question in Quantum Information Theory is: how far apart are two given quantum states? As we will see, there are many ways to answer this question, but they all start by defining a quantity which acts as a "distance" between states. Quotations marks are needed as some of the most popular choices are not distances in the mathematical

sense at all. Some of them, however, are so: an example is given by the **trace distance**. The trace distance between A and B is defined as $\|A - B\|_1$, where $\|\cdot\|_1$ is the L^1 norm of an operator defined in Appendix A. For a self-adjoint operator, the trace norm is just the sum of the absolute values of its eigenvalues.
Trace distance has also a nice operational interpretation, as it is shown in the proposition below.
Proposition 1.12. *The trace distance between two density operators ρ and σ is equal to the maximum L^1 distance between the probability distributions associated to them by a measurement $\{E_a\}_a$.*

Proof. Let us consider a generic POVM $\{E_a\}_a$ and the two probability distributions

$$p_a = \mathrm{Tr}(\rho E_a)\,, \qquad q_a = \mathrm{Tr}(\sigma E_a)\,.$$

Since $\rho - \sigma$ is self-adjoint, we can write it as:

$$\rho - \sigma = \sum\nolimits_a \lambda_a \, |\lambda_a\rangle\!\langle\lambda_a|$$

Now we have that:

$$\begin{aligned}\sum\nolimits_a |p_a - q_a| &= \sum\nolimits_a |\,\mathrm{Tr}[E_a(\rho-\sigma)]| \\ &= \sum\nolimits_a \left|\sum\nolimits_b \lambda_b \, \langle\lambda_b|E_a|\lambda_b\rangle\right| \\ &\leq \sum\nolimits_{a,b} |\lambda_b| \, \langle\lambda_b|E_a|\lambda_b\rangle \\ &\overset{1}{=} \sum_b |\lambda_b| \\ &= \|\rho - \sigma\|_1\,,\end{aligned}$$

where in 1 we used the completeness relation $\sum_a E_a = I$. Moreover, the inequality can always be saturated by choosing $E_a = |\lambda_a\rangle\!\langle\lambda_a|$. This completes the proof. □

Being a true distance, the trace distance generates an actual topology as well (again, see Appendix A).

1.4.2 Kullback-Leibler divergence

The **Kullback-Leibler divergence** [30, 31] is a very popular quantifier of the similarity of two classical probability distributions. For p and q probability distributions on the set D, it is defined as follows:

$$D_{KL}(p\|q) := \begin{cases} \sum_{x\in D} p(x)\log\left(\frac{p(x)}{q(x)}\right) & \text{if } \mathrm{supp}(p) \subset \mathrm{supp}(q) \\ +\infty & \text{otherwise}\end{cases}, \tag{1.54}$$

where the convention $0 \cdot \log 0 = 0$ has been adopted, and the the basis of the logarithm is not specified. Note that the sum is replaced by an integral whenever D ceases to be a discrete set.
D_{KL} has some useful properties:

- $D_{KL}(p\|q) \geq 0$ for any p and q;
- $D_{KL}(p\|q) = 0$ if and only if $p = q$;
- $D_{KL}(\lambda p_1 + (1-\lambda)p_2 \| \lambda q_1 + (1-\lambda)q_2) \leq \lambda D_{KL}(p_1\|q_1) + (1-\lambda)D_{KL}(p_2\|q_2)$ for any $p_{1,2}$ and $q_{1,2}$ and $\lambda \in [0,1]$ (**joint convexity**).

Despite the fact that in some sense it measures the "distance" between p and q, D_{KL} is not a distance: it fulfills neither the symmetry requirement nor the triangular inequality. Another fundamental property of D_{KL} is that it can be lower bounded by the L^1 norm, as a consequence of the famous **Pinkser's inequality** [32, 33, 34].

Proposition 1.13. *(Pinkser's inequality) The following inequality holds for any probability distributions p and q:*

$$D_{KL}(p\|q) \geq \frac{\log e}{2} \|p - q\|_1^2 \,. \tag{1.55}$$

1.4.3 Measured relative entropy

The trace distance we introduced is the quantum version of the L^1 distance between classical probability distributions. At this point, it is natural to ask whether the Kullback-Leibler divergence can be "quantised" as well. A first approach is that of going back from quantum states to classical probabilities via generalized measurements; in other words, given two quantum states ρ and σ, we want to compute D_{KL} for the outcome probability distributions $\mathcal{P}_\rho^{\mathcal{M}}(x) = \mathrm{Tr}(\rho E_x)$ and $\mathcal{P}_\sigma^{\mathcal{M}}(x) = \mathrm{Tr}(\sigma E_x)$, where $\mathcal{M} = \{E_x\}_x$ is a POVM. Since we want to measure the distinguishability of two quantum states, it makes sense to consider the POVM which makes the resulting probability distributions as diverse as possible; we define the **measured relative entropy** as [35, 36]:

$$D^M(\rho\|\sigma) := \sup_{\mathcal{M}} D_{KL}\left(P_\rho^{\mathcal{M}} \| P_\rho^{\mathcal{M}}\right) \,. \tag{1.56}$$

D^M inherits directly from D_{KL} its properties:

- $D^M(\rho\|\sigma) \geq 0$ for any ρ and σ;
- $D^M(\rho\|\sigma) = 0$ if and only if $\rho = \sigma$;
- $D^M(\lambda\rho_1 + (1-\lambda)\rho_2 \| \lambda\sigma_1 + (1-\lambda)\sigma_2) \leq \lambda D_{KL}(\rho_1\|\sigma_1) + (1-\lambda)D^M(\rho_2\|\sigma_2)$ for any $\rho_{1,2}$ and $\sigma_{1,2}$ and $\lambda \in [0,1]$ (**joint convexity**).

Other useful properties of D^M are summarized in the results below.

Lemma 1.14. *D^M is monotonically decreasing under the joint action of any CPTP map Λ:*

$$D^M(\Lambda(\rho)\|\Lambda(\sigma)) \geq D^M(\rho\|\sigma) \,.$$

In particular, it is invariant under the joint action of any unitary map.

Proof. The proof is straightforward:

$$
\begin{aligned}
D^M(\Lambda(\rho)\|\Lambda(\sigma)) &= \sup_{\mathcal{M}=\{E_a\}_a} D_{KL}(\mathcal{P}^{\mathcal{M}}_{\Lambda(\rho)}\|\mathcal{P}^{\mathcal{M}}_{\Lambda(\sigma)}) \\
&= \sup_{\mathcal{M}=\{\Lambda^*(E_a)\}_a} D_{KL}(\mathcal{P}^{\mathcal{M}}_{\rho}\|\mathcal{P}^{\mathcal{M}}_{\sigma}) \\
&\overset{1}{\geq} \sup_{\mathcal{M}=\{E_a\}_a} D_{KL}(\mathcal{P}^{\mathcal{M}}_{\rho}\|\mathcal{P}^{\mathcal{M}}_{\sigma}) \,,
\end{aligned}
$$

where in 1 we used Remark 1.2. The fact that unitary maps are invertible is enough to conclude. □

Lemma 1.15. *For any quantum states ρ and σ it holds:*

$$
D^M(\rho\|\sigma) \geq \frac{\log e}{2}\|\rho-\sigma\|_1 \,. \tag{1.57}
$$

Proof. Let us consider the spectral decomposition of $\rho-\sigma$:

$$
\rho-\sigma = \sum\nolimits_a \lambda_a \,|\lambda_a\rangle\!\langle\lambda_a| \;.
$$

As we already saw in Proposition 1.12, we have:

$$
\|\rho-\sigma\|_1 = \sum\nolimits_a \left|\mathrm{Tr}\left[\hat{E}_a(\rho-\sigma)\right]\right| \,, \tag{1.58}
$$

where we defined the POVM $\mathcal{M}=\{\hat{E}_a\}_a = \{|\lambda_a\rangle\!\langle\lambda_a|\}_a$. Now by definition we have also

$$
D^M(\rho\|\sigma) \geq D_{KL}(\mathcal{P}^{\mathcal{M}}_{\rho}\|\mathcal{P}^{\mathcal{M}}_{\sigma}) \overset{1}{\geq} \frac{\log e}{2}\left(\sum\nolimits_a \left|\mathrm{Tr}\left[\hat{E}_a(\rho-\sigma)\right]\right|\right)^2 = \frac{\log e}{2}\|\rho-\sigma\|_1^2 \,,
$$

where in 1 we used Proposition 1.13. □

An interesting fact about the measured relative entropy is that it admits a nice variational expression, which will be used extensively in the final chapter of this work. The following proof is an adaptation of [37, Lemma 1] to the infinite-dimensional case.

Lemma 1.16. *Let ρ and σ be two density operators. Then*

$$
\begin{aligned}
D^M(\rho\|\sigma) &= \sup_{h\in\mathcal{B}_{\mathrm{sa}}(\mathcal{H})} \left\{\mathrm{Tr}\,\rho h - \log \mathrm{Tr}\,\sigma e^h\right\} && (1.59)\\
&= \sup_{h\in\mathcal{B}_{\mathrm{sa}}(\mathcal{H})} \left\{\mathrm{Tr}\,\rho h + 1 - \mathrm{Tr}\,\sigma e^h\right\} . && (1.60)
\end{aligned}
$$

Proof. Let us parametrize h as $h=\log L$. Since h is bounded, there exists a finite M such that L has spectrum in $[1/M, M]$. Let us also fix $\epsilon>0$. We want to prove that (1.60) is still valid if we restrict L to be in the form $L = I+R$, with $\mathrm{rk}\,R<\infty$. Construct a finite-dimensional projector P such that $\|\rho-P\rho P\|_1\,,\, \|\sigma-P\sigma P\|_1 \leq \epsilon$. Then,

$$
\begin{aligned}
\mathrm{Tr}\,\rho\log L + 1 - \mathrm{Tr}\,\sigma L &\overset{1}{\leq} \mathrm{Tr}\,P\rho P\log L + 1 - \mathrm{Tr}\,P\sigma P L + \epsilon(\log M + M) \\
&\overset{2}{\leq} \mathrm{Tr}\,\rho\log(PLP + \mathbb{1} - P) + 1 - \mathrm{Tr}\,\sigma PLP + \epsilon(\log M + M) \\
&\overset{3}{\leq} \mathrm{Tr}\,\rho\log(PLP + \mathbb{1} - P) + 1 - \mathrm{Tr}\,\sigma(PLP + \mathbb{1} - P) \\
&\quad + \epsilon(\log M + M + 1) \,.
\end{aligned}
$$

Here, 1 follows because $\|\log L\|_\infty \leq \log M$ and $\|L\|_\infty \leq M$ ($\|\cdot\|_\infty$ is the operator norm defined in A), in 2 we applied the Jensen's operator inequality (Theorem 1.3, and 3 is an application of the estimate $\mathrm{Tr}[\sigma(\mathbb{1} - P)] = \mathrm{Tr}[\sigma - P\sigma P] \leq \|\sigma - P\sigma P\|_1 \leq \epsilon$. We see that up to introducing an arbitrarily small error we can substitute $L \mapsto PLP + \mathbb{1} - P = \mathbb{1} + R$, where $\mathrm{rk}\, R = \mathrm{rk}\, P < \infty$.
Now, let R be of finite rank, and denote with $R = \sum_{n=1}^N \lambda_n P_n$ its spectral decomposition. Then $L = \mathbb{1} + R = \sum_{n=0}^N (1+\lambda_n) P_n$, where $P_0 := \mathbb{1} - \sum_{n=1}^N P_n$ and $\lambda_0 = 0$, and consequently

$$\begin{aligned}
\mathrm{Tr}[\rho \log L] + 1 - \mathrm{Tr}[\sigma L] &= \sum_{n=0}^N \big(\log(1+\lambda_n)\,\mathrm{Tr}[\rho P_n] + \cancel{1} - \cancel{\mathrm{Tr}[\sigma]} - \lambda_n \mathrm{Tr}[\sigma P_n]\big) \\
&= \sum_{n=1}^N (\log(1+\lambda_n)\,\mathrm{Tr}[\rho P_n] - \lambda_n \mathrm{Tr}[\sigma P_n]) \\
&\overset{4}{\leq} \sum_{n=1}^N \left(\mathrm{Tr}[\rho P_n] \log \frac{\mathrm{Tr}[\rho P_n]}{\mathrm{Tr}[\sigma P_n]} - \mathrm{Tr}[\rho P_n] + \mathrm{Tr}[\sigma P_n]\right) \\
&\overset{5}{\leq} \sum_{n=0}^N \left(\mathrm{Tr}[\rho P_n] \log \frac{\mathrm{Tr}[\rho P_n]}{\mathrm{Tr}[\sigma P_n]} - \mathrm{Tr}[\rho P_n] + \mathrm{Tr}[\sigma P_n]\right) \\
&\overset{6}{=} D_{KL}\left(P_\rho^{\mathcal{M}} \| P_\sigma^{\mathcal{M}}\right) \\
&\leq D^M(\rho\|\sigma).
\end{aligned}$$

Here, the inequality in 4 comes from the estimate $a\log(1+x) - bx \leq a \log\frac{a}{b} - a + b$, (which can be proven simply by maximisation in x), 5 is a consequence of the fact that $a \log\frac{a}{b} - a + b \geq 0$ for all $a, b \geq 0$, and in 6 we introduced the measurement $\mathcal{M} := \{P_x\}_{x \in \{0,\ldots,N\}}$.

The converse is proved with exactly the same argument as in the proof of [37, Lemma 1]. Namely, for a measurement $\mathcal{M} = \{E_n\}_{n\in\mathcal{X}}$, introduce the set:

$$\widetilde{\mathcal{X}} := \{n \in \mathcal{X} : \mathrm{Tr}[\rho E_n]\,\mathrm{Tr}[\sigma E_n] > 0\},$$

and write:

$$\begin{aligned}
D_{KL}\left(P_\rho^{\mathcal{M}} \| P_\sigma^{\mathcal{M}}\right) &= \sum_{n\in\widetilde{\mathcal{X}}} \mathrm{Tr}[\rho E_n]\,(\log \mathrm{Tr}[\rho E_n] - \log \mathrm{Tr}[\sigma E_n]) \\
&= \mathrm{Tr}\left[\rho \sum\nolimits_{n\in\widetilde{\mathcal{X}}} \sqrt{E_n}\, \log\left(\frac{\mathrm{Tr}[\rho E_n]}{\mathrm{Tr}[\sigma E_n]} \cdot \mathbb{1}\right) \sqrt{E_n}\right] \\
&\overset{7}{\leq} \mathrm{Tr}\left[\rho \log\left(\sum\nolimits_{n\in\widetilde{\mathcal{X}}} \frac{\mathrm{Tr}[\rho E_n]}{\mathrm{Tr}[\sigma E_n]}\, E_n\right)\right] \\
&\overset{8}{=} \mathrm{Tr}\,[\rho \log L] + 1 - \mathrm{Tr}[\sigma L],
\end{aligned}$$

where 7 is again an application of Jensen's operator inequality, and in 8 we defined $L := \sum_n \frac{\mathrm{Tr}[\rho E_n]}{\mathrm{Tr}[\sigma E_n]} E_n$.
Starting from (1.60) it is easy to prove also (1.59): Indeed, from the inequality $\log x \leq x - 1$ we see that

$$\mathrm{Tr}\,\rho h - \log \mathrm{Tr}\,\sigma e^h \geq \mathrm{Tr}\,\rho h + 1 - \mathrm{Tr}\,\sigma e^h$$

for any h. At the same time, the expression (1.59) is manifestly invariant under transformations of the type $h \mapsto h + \lambda I$ for any $\lambda \in \mathbb{R}$. So, we can always choose a λ in both expressions such that $\operatorname{Tr} \sigma e^h = 1$, which saturates hence the aforementioned inequality. □

Lemma 1.17. *It is possible to rewrite*

$$D^M(\rho\|\sigma) = \sup_{0<L\in\mathcal{B}_{\mathrm{sa}}(\mathcal{H})} \{\operatorname{Tr} \rho \log L - \log \operatorname{Tr} \sigma L\} \tag{1.61}$$

$$= \sup_{0<L\in\mathcal{B}_{\mathrm{sa}}(\mathcal{H})} \{\operatorname{Tr} \rho \log L + 1 - \operatorname{Tr} \sigma L\}\,. \tag{1.62}$$

Proof. Expressions (1.61) and (1.62) might seem equivalent to (1.59) and (1.60), but here we require only L to be bounded, and not $\log L$. So, we are encompassing also the case in which the spectrum of L is in the form $(0, M]$ for some M. Since we are including more L, we just have to prove that

$$\sup_{h\in\mathcal{B}_{\mathrm{sa}}(\mathcal{H})} \left\{\operatorname{Tr} \rho h - \log \operatorname{Tr} \sigma e^h\right\} \geq \sup_{0<L\in\mathcal{B}_{\mathrm{sa}}(\mathcal{H}_m)} \{\operatorname{Tr} \rho \log L - \log \operatorname{Tr} \sigma L\}\ .$$

For any positive and bounded $L = \sum_j \ell_j \,|\ell_j\rangle\!\langle\ell_j|$ we can define $L_\delta = \sum_j \ell_{j,\delta}\,|\ell_j\rangle\!\langle\ell_j| = \sum_j \max\{\ell_j, \delta\}\,|\ell_j\rangle\!\langle\ell_j|$. By construction $L_\delta > \delta\mathbb{1}$: now we just have to prove that

$$\lim_{\delta\to 0} \operatorname{Tr} \rho \log L_\delta - \log \operatorname{Tr} \sigma L_\delta \geq \operatorname{Tr} \rho \log L - \log \operatorname{Tr} \sigma L \tag{1.63}$$

for any $\rho = \sum_j p_j\,|a_j\rangle\!\langle a_j|$ and σ. Thanks to the scale invariance in L of the expression (1.61), we can assume $L \leq \mathbb{1}$. Then the series

$$\operatorname{Tr} \rho \log L_\delta = \sum_{j,k} p_j |\langle a_j|\ell_k\rangle|^2 \log \ell_{k,\delta} \tag{1.64}$$

is well defined since all its terms are negative for $\delta < 1$. Moreover, each one of its terms is monotonically decreasing in δ, which ensures that the series is continue in δ:

$$\lim_{\delta\to 0} \operatorname{Tr} \rho \log L_\delta = \operatorname{Tr} \rho \log L\,. \tag{1.65}$$

We can apply a similar procedure for $\operatorname{Tr} \sigma A$. Hence, (1.63) is actually an equality, and the claim is proved. □

Remark 1.18. *The definition of the measured relative entropy can easily be extended to $\sigma \in \mathcal{T}_+(\mathcal{H}) \setminus \{0\}$. Indeed, if $\sigma = \lambda\tilde{\sigma}$, with $\tilde{\sigma} \in \mathcal{D}(\mathcal{H})$ and $\lambda \in (0, 1]$, we simply have: $D_{KL}(\rho\|\sigma) = D_{KL}(\rho\|\tilde{\sigma}) - \log\lambda$, and all the variational expressions we proved so far remain valid. If $\sigma = 0$, expression* (1.60) *gives $+\infty$, as it should.*

1.4.4 Other entropies

We want to introduce another quantised version of the Kullback-Leibler divergence. To do so, we start introducing a functional of a single operator: the **von Neumann entropy**. The von Neumann entropy (or simply "entropy") of some quantum state $\rho \in \mathcal{D}(\mathcal{H})$ can be defined as

$$S(\rho) \coloneqq -\operatorname{Tr}[\rho \log \rho]\,. \tag{1.66}$$

This is clearly a quantised version of the **Shannon entropy** of a classical probability distribution $p(x)$ with domain D:

$$S_S(p) = -\sum_{x\in D} p(x)\log p(x)\,.$$

Note that (1.66) is a well-defined although possibly infinite quantity. One way to define it is via the infinite sum $S(\rho)=\sum_i(-p_i\log p_i)$, where $\rho=\sum_n p_n\,|a_n\rangle\!\langle a_n|$ is the spectral decomposition of ρ. Since all the terms of this sum are non-negative, to the sum itself can be assigned a well-defined value, possibly $+\infty$.
An alternative quantised version of the Shannon entropy is the **Wehrl entropy**, which for a quantum state ρ is defined through its Q-function:

$$S_W(\rho) = -\int d^2\alpha\, Q^\rho(\alpha)\log Q^\rho(\alpha)\,. \tag{1.67}$$

It holds that $S_W(\rho)\geq S(\rho)$, and the minimum Wehrl entropy for a given von Neumann entropy is attained for thermal states [38].
Following the same philosophy of the von Neumann entropy, we define the **relative entropy** between two quantum states $\rho,\sigma\in\mathcal{D}(\mathcal{H})$ as [39]:

$$D(A\|B) \coloneqq \mathrm{Tr}\left[\rho(\log\rho-\log\sigma)\right]. \tag{1.68}$$

Again, the above expression is well defined and possibly infinite [40]. To see why, we represent it as the infinite sum $D(A\|B)\coloneqq\sum_{i,j}|\langle a_i|b_j\rangle|^2\,(a_i\log a_i - a_i\log b_j + b_j - a_i) + \mathrm{Tr}[A-B]$, where $A=\sum_i a_i\,|a_i\rangle\!\langle a_i|$ and $B=\sum_j b_j\,|b_j\rangle\!\langle b_j|$ are the spectral decompositions of A and B, respectively. As detailed in [40], the convexity of $a\mapsto a\log a$ implies that all terms of the sum are non-negative, making the expression well defined. Clearly, a necessary condition for $D(A\|B)$ to be finite is that $\mathrm{supp}\,A\subseteq\mathrm{supp}\,B$. D has three interesting properties [41].

- It is additive on tensor product states:

$$D(\rho^A\otimes\rho^B\|\sigma^A\otimes\sigma^B) = D(\rho^A\|\sigma^A)+D(\rho^B\|\sigma^B)\,.$$

- It is jointly convex:

$$D(\lambda\rho_1+(1-\lambda)\rho_2\|\lambda\sigma_1+(1-\lambda)\sigma_2)\leq\lambda D(\rho_1\|\sigma_1)+(1-\lambda)D(\rho_2\|\sigma_2)$$

  for any $\lambda\in[0,1]$.

- It is monotonically decreasing under the joint action of CPTP maps:

$$D(\Lambda(\rho)\|\Lambda(\sigma))\leq D(\rho\|\sigma)\,,$$

  and hence in particular it is invariant under unitary maps.

Finally, it is well known that one has $D^M(\rho\|\sigma)\leq D(\rho\|\sigma)$ for all pairs of states ρ,σ [35], with equality if and only if $[\rho,\sigma]=0$ [37, 42]. This also implies that Proposition 1.15 holds for D as well.

2 | Quantum Resource Theories

2.1 An introduction to the framework

2.1.1 Physical motivations

Up to now, we were concerned with defining physical entities such as quantum states and quantum operations, without questioning much whether they could be reproduced and observed in a realistic experimental setting. Now we want to adopt a new point of view, and start from the question: what are we actually able to do? As long as anything can be achieved, nothing has value; on the contrary, whenever our capabilities turns out to be limited, anything is beyond them suddenly becomes a precious resource.
This practical idea can be readily applied to Quantum Mechanics as well. As quantum technologies grow in popularity, the following questions arise naturally.

- Which peculiar features of Quantum Theory are responsible for the supposed operational advantages of quantum technologies with respect to classical ones?
- If, on the contrary, our devices cannot produce, manipulate and exploit "quantumness" (whatever this means), which tasks become absolutely out of reach?
- How to establish which quantum states or processes are more precious than others within a given scenario?
- Which physical properties are responsible for a certain exquisitely quantum phenomenon?

In recent years, **Quantum Resource Theories** [43] emerged as the natural framework where these questions can be addressed. In the context of Quantum Physics, resources (or better, quantum resources) arise for instance whenever we have not access to some quantum properties of a system due to experimental limitations. Usually, these properties can be exploited in some practical task: by showing what we cannot do without them, we prove why and how they are indeed precious. Quantum Resource Theories allow also for rigorously quantifying the resourcefulness of quantum objects, and the price to pay, in terms of resources, for completing certain tasks.
For the sake of concreteness, let us make some examples. Consider two experimentalists, Alice and Bob, working in distant laboratories with only a classical phone for communicating: this means that they can neither make their systems interact directly, nor freely share quantum objects with each other. This is clearly an example of an experimental limitation. To keep things simple (for the moment) let us stick to pure states and unitary

evolution. If they start from a tensor product state $|\psi^A\rangle \otimes |\psi^B\rangle$ and apply only operations in the form $U^A \otimes U^B$, they will inevitably end up in another tensor product state $|\psi'^A\rangle \otimes |\psi'^B\rangle$. So, everything is not in such a form (**entangled states**) cannot be freely generated. But would it be also useful for something? As it is well-known, the answer is a resounding yes, and prominent examples of applications are **quantum teleportation** [44], **dense coding** [45, 46] and **quantum criptography** [47, 48, 49, 50]. But **entanglement** initially drew attention [51, 52] as the culprit of **locality violations** [53] via **Bell's inequalities**[54, 55, 56]. Hence, in addition to being an example of a useful resource which cannot be generated freely in many realistic scenarios, it is also a nice example of how a property of a quantum system can be identified to be responsible (in a loose sense, as entanglement do not implies violations of Bell's inequalities in general [57]) of a fundamental phenomenon. **Entanglement** theory [58, 59, 60, 61] has also the honour to be the first quantum resource to be studied as such.
Another feature of Quantum Mechanics which has been elevated from being a counter-intuitive quirkiness to being a desirable property to be exploited is **coherence** [62, 63, 64, 65, 66, 67]. This rather recent Resource Theory is motivated by the fact that, as we pointed out at the really beginning of this work, interaction with the environment are unavoidable. These interactions usually select a preferred basis (the **incoherent basis**) for the Hilbert space, whose states are stable under the open dynamics of the systems. Any state which is not a convex combination of them is unstable, and rapidly decay in a stable one. This process is called **decoherence**, and is ubiquitous in open quantum systems. Again, being quantum coherence hard to preserve, any state containing some of it (for example, an undecohered Schrödinger's cat) becomes a resource.
Finally, it was recently realized that Quantum Thermodynamics [68] can be rephrased as a Quantum Resource Theory, with out-of-thermal-equilibrium states being the resourceful ones [69, 70].

2.1.2 Free states or free operations?

We now start to introduce in a more systematic way the ingredients of a Resource Theory. To begin with, we need to establish which are the freely accessible objects within a certain experimental setting. The reason why we keep talking generically about "objects" will be clear in a moment.
We could start by defining the set of **free states** (as opposed to resourceful, precious, "costly" ones) $\mathcal{D}_f(\mathcal{H}) \subset \mathcal{D}(\mathcal{H})$ as the state that can be prepared without effort in our laboratory. In some cases, starting from $\mathcal{D}_f(\mathcal{H})$ is the more natural approach, as free states are particularly easy to characterize. An example is given by the resource theory of coherence: as we said, interactions with the environment select a stable basis of states $\{|\phi_n\rangle\}_n$, and all classical mixtures of these states, i.e. **incoherent states**, are considered free. In this case, experimental conditions dictate in which states the system can be prepared without immediately decohere in something else (in this sense they are easy to prepare), but do not select unambiguously a set of easily implementable operations, which can instead be chosen depending on additional conditions which can be specified from case to case.
We have just seen an example of a much more general fact: whenever we build our Resource Theory starting from the set of free states, a natural choice for the set $\mathcal{O}_f(\mathcal{H} \to \mathcal{H}')$ of **free operations**, i.e., operations which can be implemented at no cost, does not

emerge (please note the slight abuse of notation: elements of $\mathcal{O}_f(\mathcal{H} \to \mathcal{H}')$ are not maps from $\mathcal{H}$ to $\mathcal{H}'$, but from $\mathcal{D}(\mathcal{H})$ to $\mathcal{D}(\mathcal{H}')$). Indeed, there might be multiple classes of quantum operations which are compatible with $\mathcal{D}_f(\mathcal{H})$, i.e., which map free states into free states. For certain, any operation which do not fulfill this basic requirement cannot be considered free: if we could easily generate non-free states from free ones, the whole concept of resources would not make much sense. Moreover, a class of free operation must be closed under composition: if two or more operations are free, then also all the operations obtained by combining them in sequence must be so. Identity must be a free operation, as "doing nothing" is always possible. Wrapping up:

R1) $\rho \in \mathcal{D}_f(\mathcal{H})$, $\Lambda \in \mathcal{O}_f(\mathcal{H} \to \mathcal{H}') \implies \Lambda(\rho) \in \mathcal{D}_f(\mathcal{H}')$;

R2) $\Lambda_1 \in \mathcal{O}_f(\mathcal{H} \to \mathcal{H}')$, $\Lambda_2(\mathcal{H}' \to \mathcal{H}'') \implies \Lambda_2 \circ \Lambda_1 \in \mathcal{O}_f(\mathcal{H} \to \mathcal{H}'')$;

R3) $I \in \mathcal{O}_f(\mathcal{H} \to \mathcal{H}')$.

Now we re ready to give a definition of a **Quantum Resource Theory**: a Quantum Resource Theory defined on a class of Hilbert spaces $\mathfrak{H}$ is a map which associates to any couple of Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$ in $\mathfrak{H}$ the corresponding sets of free states $\mathcal{D}_f(\mathcal{H})$ and $\mathcal{D}_f(\mathcal{H}')$, and a set of free operations $\mathcal{O}_f(\mathcal{H} \to \mathcal{H}')$ satisfying R1)-R3). If $\mathcal{D}_f(\mathcal{H})$ is convex for any $\mathcal{H}$ in $\mathfrak{H}$, then the Quantum Resource Theory is said to be **convex** as well. A Resource Theory is instead **affine** if for any collection of free states $\{\sigma_a\}_a$ and (possibly negative) real numbers $\{c_a\}_a$ such that $\sum_a c_a = 1$, also $\sigma = \sum_a c_a \sigma_a$ is a free state (as long as it is a physical state, i.e., if $\sigma \geq 0$, which is not guaranteed a priori).
The majority of interesting Quantum Resource Theories have some additional constraints. For examples, they are usually compatible with the tensor product. At the level of states, this means that $\mathcal{D}_f(\mathcal{H})$ is such that if $\rho \in \mathcal{D}_f(\mathcal{H}^A)$ and $\sigma \in \mathcal{D}_f(\mathcal{H}^B)$ then also $\rho \otimes \sigma \in \mathcal{D}_f(\mathcal{H}^A \otimes \mathcal{H}^B)$. Interestingly enough, some physically motivated Resource Theories do not have this property [71], which means that $\rho \otimes \sigma$ can contain some amount of a resource even if ρ and σ alone do not. At any rate, we will not be concerned with them.
At the level of operations, compatibility with the tensor product implies that free operations are "completely free", i.e., they remain free even when acting on just a part of a larger system. The partial trace must be a free operation as discarding a system is always possible. Finally, we require appending free states to be a free operation. Summarizing:

- $\rho^A \in \mathcal{D}_f(\mathcal{H}^A)$, $\rho^B \in \mathcal{D}_f(\mathcal{H}^B) \implies \rho^A \otimes \rho^B \in \mathcal{D}_f(\mathcal{H}^A \otimes \mathcal{H}^B)$;
- $\Lambda \in \mathcal{O}_f(\mathcal{H} \to \mathcal{H}') \implies \Lambda \otimes I \in \mathcal{O}_f(\mathcal{H} \otimes \mathcal{H}^E \to \mathcal{H}' \otimes \mathcal{H}^E)$;
- $\mathrm{Tr}_E(\cdot) \in \mathcal{O}_f(\mathcal{H} \otimes \mathcal{H}^E \to \mathcal{H})$;
- $\Lambda_\sigma : (\cdot) \mapsto (\cdot) \otimes \sigma \in \mathcal{O}_f(\mathcal{H} \to \mathcal{H} \otimes \mathcal{H}') \, \forall \, \sigma \in \mathcal{D}_f(\mathcal{H}')$.

By combining channels which append free states with the partial trace, we can always throw away any state and substitute it with a free one.
As we said, there is not a unique choice for the class of free operations. In any case, there is a class which is "special": the one which contains as much operations as possible, which we will denote with $\mathcal{O}_f^{max}(\mathcal{H} \to \mathcal{H}')$. Equivalently, it is the unique class of operations which is defined by all the condition we stated until now, and nothing else. Usually, such a set has no strong operational motivations, as it encompasses much more operations than those which are actually available in a laboratory. Despite this fact, it is often a

smart choice because of its simplicity; since many of the results in a Resource Theory take the form of no-go theorems or upper bound over the efficiency of some process, they will still hold under stronger assumptions.
The other approach is to start from the set of free operations $\mathcal{O}_f(\mathcal{H} \to \mathcal{H}')$, which are usually motivated by limitations on our capabilities of manipulating quantum states. It is easy to see that in this case, contrarily to what happened before, there is no freedom in the choice of the set of free states $\mathcal{D}_f(\mathcal{H})$. On one hand, any state which can be freely obtained starting from any other state is necessarily free: if this was not the case, either $\mathcal{D}_f(\mathcal{H})$ is empty (and the Resource Theory is trivial), or we could freely generate resource. On the other hand, as we already pointed out, we can always throw away the state we have and substitute it with a free one. Hence:

$$\mathcal{D}_f(\mathcal{H}) = \{\rho \in \mathcal{D}(\mathcal{H}) : \forall \omega \in \mathcal{D}(\mathcal{H}) \, \exists \Lambda \in \mathcal{O}_f(\mathcal{H} \to \mathcal{H}) : \Lambda(\omega) = \rho\} \, .$$

2.2 General features of Quantum Resource Theories

2.2.1 Resource monotones

By introducing the concepts of free states and free operations, we addressed the problem of formally distinguishing physical objects that can be considered free from those which can instead be regarded as valuable. Now we need to take a step further and answer the question: how much a given object is resourceful?
If we want a quantitative response, we need some sort of functional assigning to each quantum object a value. We will stick to quantifying the resource content of quantum states. The problem of quantifying the resourcefulness of quantum operations is a very active area of research [72, 73, 74] at the moment but goes beyond the scope of this work. Since we want to remain as general as possible, we ask ourselves which are the most basic assumptions we need to make. For sure, any functional which aim for quantifying the resource content of states, must attain the minimum value when computed on a free state. Assuming that this minimum is different from $-\infty$, we can always re-scale the functional in such a way that its minimum becomes 0. The intuition that the resource cannot be increased by free operations, otherwise it would be easily produced and hence not a resource anymore, leads to the requirement that the functional must monotonically decrease under free operations. For this reason, such a functional is usually called a **resource monotone**, or simply a **monotone**.
Definition 2.1. *Given a class of Hilbert spaces $\mathscr{H}$ and a Quantum Resource Theory defined on it, a functional $\mathcal{F} : \mathcal{H} \to \mathbb{R}$, where $\mathcal{H}$ is any Hilbert space in $\mathscr{H}$, is called a* ***resource monotone*** *if:*

- $\mathcal{F}(\rho) = 0 \, \forall \, \rho \in \mathcal{D}_f(\mathcal{H})$;
- $\mathcal{F}(\rho) \geq 0 \, \forall \, \rho \in \mathcal{D}(\mathcal{H})$;
- $\mathcal{F}(\Lambda(\rho)) \leq \mathcal{F}(\rho) \, \forall \, \rho \in \mathcal{D}(\mathcal{H}), \, \forall \, \Lambda \in \mathcal{O}_f(\mathcal{H} \to \mathcal{H}')$.

A resource monotone can be defined in many ways. One of the most popular is based on the following result.

Proposition 2.2. *Let consider a Quantum resource Theory on $\mathcal{H}$. For any $\rho \in \mathcal{D}(\mathcal{H})$ with $\mathcal{H}$ in $\mathcal{H}$, we define the following functional:*

$$\mathcal{F}_r(\rho) := \inf_{\sigma \in \mathcal{D}_f(\mathcal{H})} \delta(\rho, \sigma)\,, \tag{2.1}$$

where $\delta : \mathcal{D}(\mathcal{H}) \times \mathcal{D}(\mathcal{H}) \to \mathbb{R}_+$ is a generic functional satisfying:

- $\delta(\rho, \sigma) \geq 0 \,\forall\, \rho,\, \sigma$*;*
- $\delta(\rho, \rho) = 0$*;*
- $\delta(\Lambda(\rho), \Lambda(\sigma)) \leq \delta(\rho, \sigma)$ *for any CPTP map Λ.*

Then, $\mathcal{F}_r$ is a resource monotone.

Proof. Being δ always non-negative, also $\mathcal{F}_r$ is so. Moreover, if $\rho \in \mathcal{D}_f(\mathcal{H})$, it sufficies to take $\sigma = \rho$ and we get $\mathcal{F}_r(\rho) = 0$. Now, if $\Lambda \in \mathcal{O}_f(\mathcal{H} \to \mathcal{H}')$ we have

$$\begin{aligned}
\mathcal{F}_r(\Lambda(\rho)) &= \inf_{\sigma \in \mathcal{D}_f(\mathcal{H}')} \delta(\Lambda(\rho), \sigma) \\
&\overset{1}{\leq} \inf_{\sigma \in \Lambda(\mathcal{D}_f(\mathcal{H}))} \delta(\Lambda(\rho), \sigma) \\
&= \inf_{\sigma \in \mathcal{D}(\mathcal{H})} \delta(\Lambda(\rho), \Lambda(\sigma)) \\
&\leq \inf_{\sigma \in \mathcal{D}(\mathcal{H})} \delta(\rho, \sigma)\,.
\end{aligned}$$

where in 1 we used the fact that $\Lambda(\mathcal{D}_f(\mathcal{H})) \subseteq \mathcal{D}_f(\mathcal{H}')$. Hence, $\mathcal{F}_r$ is also monotonically decreasing under free operations, and it is a proper resource monotone. □

Definition 2.3. *If we choose $\delta(\cdot\,,\cdot) = D^{(M)}(\cdot\|\cdot)$, we obtain the* ***(measured) relative entropy of resource****.*

Another important family of resource monotones is represented by robustness measures. Below we report the two most important examples.

Definition 2.4. *We define the* ***absolute robustness of resource*** *for a state $\rho \in \mathcal{D}(\mathcal{H})$ as:*

$$\mathcal{R}_A(\rho) := \inf_{\sigma \in \mathcal{D}_f(\mathcal{H})} \left\{ r > 0 : \frac{\rho + r\sigma}{1 + r} \in \mathcal{D}_f(\mathcal{H}) \right\}\,. \tag{2.2}$$

Definition 2.5. *We define the* ***global robustness of resource*** *for a state $\rho \in \mathcal{D}(\mathcal{H})$ as:*

$$\mathcal{R}_G(\rho) := \inf_{\sigma \in \mathcal{D}(\mathcal{H})} \left\{ r > 0 : \frac{\rho + r\sigma}{1 + r} \in \mathcal{D}_f(\mathcal{H}) \right\}\,. \tag{2.3}$$

Remark 2.6. *It trivially holds that $\mathcal{R}_G(\rho) \leq \mathcal{R}_A(\rho)$ for any quantum state $\rho \in \mathcal{D}(\mathcal{H})$.*

It is very easy to see that the robustness measures we just defined are resource monotones as well.

Proposition 2.7. *$\mathcal{R}_A$ and $\mathcal{R}_G$ are resource monotones.*

Proof. Non-negativity and vanishing of free states are trivial properties. Also monotonicity is readily proven: indeed,

$$\frac{\rho + r\sigma}{1+r} \in \mathcal{D}_f(\mathcal{H})\,,\; \sigma \in \mathcal{D}_{(f)}(\mathcal{H}) \implies \frac{\Lambda(\rho) + r\Lambda(\sigma)}{1+r} \in \mathcal{D}_f(\mathcal{H}')\,,\; \Lambda(\sigma) \in \mathcal{D}_{(f)}(\mathcal{H}')\,,$$

for any $\Lambda \in \mathcal{O}_f(\mathcal{H} \to \mathcal{H}')$.

□

In many cases, it is useful to define ad-hoc monotones, based on some peculiar feature of a given Quantum Resource Theory rather than general structures of Quantum Resource Theories, as we have done until now. Usually, these functional have a more direct operational meaning, but it is much harder to prove that they are indeed monotones. We will see some example of such monotones in the following.

2.2.2 Properties of the monotones

Resource monotones are not just a way to merely assign to any state a number symbolizing its corresponding resource content, but they represent the core of any Resource Theory and the main tool in order to obtain results about what can and cannot be done with a certain amount of resource. In order to prove something useful, we need our monotone to have some additional properties; different properties are need for different types of result.
The simplest property a resource monotone can have is convexity. Physically, it implies that classically mixing states cannot increase their resource content.
Another basic property one can require for a functional is some sort of continuity. From a physical point of view, continuity of resource monotones corresponds to the intuition that similar states should have similar resource contents, and slightly perturbing a state should not change dramatically its usefulness; from a practical point of view, dealing with continuous function is much easier, as it allows for considering approximations of the states we are actually interested in. Moreover, some tasks may require to produce states within a certain error: without any kind of continuity we could state nothing. We will see examples of this situation in the future. The best case scenario is to have a continuous resource monotones; a very strong type of continuity, crucial for the following, is presented below.
Definition 2.8. *A real-valued functional $f : \mathcal{D}(\mathcal{H}) \to \mathbb{R}$ is said to be* ***asymptotically continuous*** *[75] if*

$$|f(\rho) - f(\sigma)| \leq c\|\rho - \sigma\|_1 \log d + \eta(\|\rho - \sigma\|_1) \tag{2.4}$$

for any ρ and σ, where $d = \dim \mathcal{H}$, c is a constant and $\eta(x)$ is a dimension-independent function which vanishes for $x \to 0$.
Remark 2.9. (2.4) *makes sense only in finite-dimensional spaces.*

In particular, it f is asymptotically continuous, it holds:

$$\liminf_{\rho \xrightarrow{\text{TNT}} \rho_0} f(\rho) = f(\rho_0)\,. \tag{2.5}$$

Sometimes, it is too much to require asymptotic continuity. An example is given by infinite-dimensional Quantum Resource Theories, as we already pointed out, but this discussion is not restricted to them. Luckily, it is often enough to require something much weaker than continuity, i.e., semi-continuity.

Definition 2.10. *A real-valued functional $f : \mathcal{D}(\mathcal{H}) \to \mathbb{R}$ is said to be **lower semi-continuous** if*

$$\liminf_{\rho \xrightarrow{\text{TNT}} \rho_0} f(\rho) \geq f(\rho_0) \tag{2.6}$$

for any ρ.

The definition of **upper semi-continuity** can be easily derived from the previous one. The physical meaning of lower semi-continuity is straightforward: it means that the resource content of a state cannot decrease abruptly if it gets slightly perturbed; equivalently, it implies that if we aim for approximating a state within a very small error, we will need at least as much resource as to produce the exact state.

The next property we present is the following.

Definition 2.11. *A resource monotone satisfying the condition $\mathcal{F}(\rho) = 0 \iff \rho \in \mathcal{D}_f(\mathcal{H})$ is said to be **faithful**.*

The concept of faithfulness is linked to the idea that is desirable to have a monotone that never fails to detect the presence of resource. In any case, many important and useful resource monotones are not faithful, and vanish on non-free states [76, 77]. Usually, this is linked to the fact that some states contain resource, but it cannot be extracted from them. At any rate, the results below ensure that a large class of resource monotones are faithful.

Proposition 2.12. *Let us consider again $\mathcal{F}_r$ as defined in Proposition 2.2. If we have, in addition, the following conditions:*

- *any $\mathcal{H}$ in $\mathfrak{H}$ is finite-dimensional;*
- *$\mathcal{D}_f(\mathcal{H})$ is closed for any $\mathcal{H}$ in $\mathfrak{H}$;*
- *$\delta(\cdot\,,\cdot)$ is lower semi-continuous in the second argument;*
- *$\delta(\rho, \sigma) = 0 \iff \rho = \sigma$.*

then $\mathcal{F}_r$ is also faithful.

Proof. Being density operators normalized, $\mathcal{D}(\mathcal{H})$ is always limited. By hypothesis $\mathcal{D}_f(\mathcal{H})$ is closed and, being $\mathcal{H}$ finite dimensional, it is also compact. Note that we do not have to specify in which topology $\mathcal{D}_f(\mathcal{H})$ is closed since we are working in finite dimension by hypothesis. Now let us assume that $\mathcal{F}_r(\rho) = 0$: this means that it exists a sequence $\{\sigma_n\}_n \subset \mathcal{D}_f(\mathcal{H})$ such that $\lim_{n\to\infty} \delta(\rho, \sigma_n) = 0$. But being $\mathcal{D}_f(\mathcal{H})$ compact we can extract a subsequence $\{\sigma_{k_n}\}_n$ converging at $\sigma_0 \in \mathcal{D}_f(\mathcal{H})$. By lower semi-continuity we have:

$$0 \leq \delta(\rho, \sigma_0) \leq \liminf_n \delta(\rho, \sigma_n) = 0 \implies \delta(\rho, \sigma_0) = 0 \implies \rho = \sigma_0\,.$$

Thus, $\rho \in \mathcal{D}_f(\mathcal{H})$ as well. □

An example of a $\delta(\cdot\,,\cdot)$ functional which is lower semi-continuous in the second argument is given by the relative entropy [41]. But in this proof it was crucial that $\dim\mathcal{H} < \infty$; nonetheless, in many cases resource monotones based on relative entropies are faithful even in infinite-dimensional Quantum resource Theories, as we prove in the following results.

Proposition 2.13. *If $\mathcal{D}_f(\mathcal{H})$ is closed with respect to the norm topology for any $\mathcal{H}$ in $\mathscr{H}$ and $\delta(\cdot\,,\cdot) = D^M(\cdot\|\cdot)$ or $\delta(\cdot\,,\cdot) = D(\cdot\|\cdot)$, $\mathcal{F}_r$ is faithful.*

Proof. Let us assume that $\rho \notin \mathcal{D}_f(\mathcal{H})$. Being $\mathcal{D}_f(\mathcal{H})$ closed in trace norm, it exists a ball (again, in trace norm) of radius $\delta > 0$ centered in ρ and with no intersection with $\mathcal{D}_f(\mathcal{H})$. So, $\|\rho - \sigma\|_1 > \delta$ for any $\sigma \in \mathcal{D}_f(\mathcal{H})$. Now it is sufficient to invoke Pinsker's inequality to conclude the proof. □

Proposition 2.14. *$\mathcal{R}_A$ and $\mathcal{R}_G$ are faithful whenever $\mathcal{D}_f(\mathcal{H})$ is closed in trace norm.*

Proof. Let us consider a non-free state $\rho \notin \mathcal{D}_f(\mathcal{H})$. Being $\mathcal{D}_f(\mathcal{H})$ closed, it exists a ball in trace norm with radius $\delta > 0$ centered in ρ and with no intersection with $\mathcal{D}_f(\mathcal{H})$. If $\mathcal{R}_G(\rho) = 0$, it exists a sequence r_n such that $\lim_{n\to\infty} r_n = 0$ and a sequence $\{\sigma_n\}_n \subset \mathcal{D}(\mathcal{H})$ such that

$$\frac{\rho + r_n\sigma_n}{1 + r_n} =: \omega_n \in \mathcal{D}_f(\mathcal{H})\,,\ \forall\, n\,.$$

But this means that

$$\lim_{n\to\infty} \|\rho - (1 + r_n)\omega_n\|_1 = \lim_{n\to\infty} \|\rho - \omega_n\|_1 = 0$$

which is in contrast with the fact that ρ is at a non-zero distance in trace norm from $\mathcal{D}_f(\mathcal{H})$. We conclude by recalling that $\mathcal{R}_A(\rho) \geq \mathcal{R}_G(\rho)$. □

Another class of properties we could want to require for our resource monotones is related to how a resource monotone behave on bipartite states. The simplest condition is the following.

Definition 2.15. *A real-valued functional $f : \mathcal{D}(\mathcal{H}^A \otimes \mathcal{H}^B) \to \mathbb{R}$ is said to be **weakly additive** if*

$$f(\rho^A \otimes \rho^B) = f(\rho^A) + f(\rho^B)$$

*for any $\rho^A \in \mathcal{D}(\mathcal{H}^A)$ and $\rho^B \in \mathcal{D}(\mathcal{H}^B)$; it is said to be **strongly additive** if*

$$f(\rho^{AB}) = f(\mathrm{Tr}_B\,\rho^{AB}) + f(\mathrm{Tr}_A\,\rho^{AB})$$

for any $\rho^{AB} \in \mathcal{D}(\mathcal{H}^{AB})$.

If instead of an equality we have inequalities, we obtain the definition of **weak/strong sub-additivity** and **weak/strong super-additivity**. Clearly, they extend automatically to generic multipartite states. These properties are particularly useful when it comes to convert many copies of a state into as may copies as possible of another; we will come back on this later. We also mention that given a resource monotone $\mathcal{F}$, it is always possible to define the corresponding **regularized monotone** as follows:

$$\mathcal{F}^\infty(\rho) := \lim_{n\to\infty} \frac{1}{n}\mathcal{F}(\rho^{\otimes n})\,. \tag{2.7}$$

This new monotone is additive by construction; the drawback is that it is usually impossible to compute exactly. An example of weakly sub-additive monotone is given by the relative entropy of resource.

Proposition 2.16. *The relative entropy of resource is a weakly sub-additive monotone.*

Proof. The proof is straightforward:

$$\inf_{\sigma^{AB}\in\mathcal{D}_f(\mathcal{H}^A\otimes\mathcal{H}^B)} D(\rho^A\otimes\rho^B\|\sigma^{AB}) \leq \inf_{\substack{\sigma^A\in\mathcal{D}_f(\mathcal{H}^A)\\ \sigma^B\in\mathcal{D}_f(\mathcal{H}^B)}} D(\rho^A\otimes\rho^B\|\sigma^A\otimes\sigma^B)$$

$$\stackrel{1}{=} \inf_{\sigma^A\in\mathcal{D}_f(\mathcal{H}^A)} D(\rho^A\|\sigma^A) + \inf_{\sigma^B\in\mathcal{D}_f(\mathcal{H}^B)} \mathcal{D}(\mathcal{H}_()\rho^B\|\sigma^B)\,,$$

where in 1 we used the additivity of relative entropy on tensor product states. □

Some monotones display a stronger form of monotonicity. In particular, for a functional to be a resource monotone is required to monotonically decrease under any quantum channel, but do not state anything about probabilistic transformations. If we require our monotone not to increase even on average when probabilistic protocols are applied, we obtain the following, stronger, property.

Definition 2.17. *A resource monotone $\mathcal{F}$ is said to display **strong monotonicity** if, for any free operation giving as a result a flagged outcome in the form:*

$$\rho \mapsto \sum\nolimits_n p_n\rho_n' \otimes |n^A\rangle\langle n^A|\ ,$$

it holds:

$$\mathcal{F}(\rho) \geq \sum\nolimits_n p_n\mathcal{F}(\rho_n)\,. \tag{2.8}$$

When a resource monotone displays strong monotonicity, resources cannot be increased even on average, performing post-selection on the outcomes. It is obviously a stronger condition than standard monotonicity.

To conclude this section, we give the following, last definition.

Definition 2.18. *The resource monotone $\mathcal{F}$ is said to satisfy **tensorisation** if*

$$\mathcal{F}(\rho\otimes\rho') = \max\left\{\mathcal{F}(\rho), \mathcal{F}(\rho')\right\}\ ., \tag{2.9}$$

for any $\rho, \rho' \in \mathcal{D}(\mathcal{H})$.

2.2.3 States convertibility

When a limited amount of resource is available, the simplest task we can aim for is to transform resourceful states from one another using only free operations. Indeed, in a typical experimental setting, some amount of resource is given, but usually it is not in the form it is needed. So, it is fundamental to understand to what extent we can manipulate resources in order to pose stricter and more realistic constraints on the minimum amount of resource needed for a certain task.

States convertibility is also the first and most striking example in which resource monotones play a major role in proving rigorous results. Indeed, starting from the paradigm that resources cannot be increased by free operations, we can rule out all the transitions

which do not respect this simple yet unavoidable requirement. As we will see, there are many types of state transformations, and different properties of the monotones are required in different scenarios.
Of course, in the context of a Quantum Resource Theory, the simplest question we can ask ourselves is: can we go from state ρ to ρ' with only free operations? A necessary, but not sufficient in general, condition, is that

$$\mathcal{F}(\rho) \geq \mathcal{F}(\rho') \tag{2.10}$$

for any resource monotone $\mathcal{F}$ we can define for our Resource Theory. To prove that we cannot freely go from ρ to ρ', then, it sufficies to find a monotone (with no particular additional properties) such that condition (2.10) is not satisfied. However, this is not the clever question we may ask, for several reasons:

- the requirement is too stringent, as in most cases the transformation cannot be freely achieved even if $\mathcal{F}(\rho) \geq \mathcal{F}(\rho')$;
- it might be easier to produce ρ' probabilistically with some finite probability p, instead of converting ρ with a deterministic CPTP map;
- it would sufficies to repreduce ρ' within a certain, reasonably small, error;
- with the realistic experimental situation in mind, it does not make much sense to talk about exact transformation;
- it might be more convenient to transform many copies of the input in many copies of the output, instead of focusing on single-copy tasks.

Based on this observations, we can define many other types of states conversions; the downside is that, being more general tasks, they require more specific monotones to be constrained. In particular, some of the properties we presented in the previous section might come in handy. The first kind of conversions we want to consider are **probabilistic conversions**. More precisely, we ask whether it exists a free CPTP map acting as follows:

$$\rho \mapsto p_0 \rho' \otimes |0^A\rangle\langle 0^A| + \sum_{j=1}^{N} p_j \rho_j \otimes |j^A\rangle\langle j^A| \ ,$$

for some $N \in \mathbb{N}$. Again, we used a flagged outcome to describe the result of a probabilistic transformation. The question then becomes: what is $P^{max}(\rho \to \rho')$, the maximum value p_0 can assume? The answer can be given via a monotone satisfying strong monotonicity:

$$\mathcal{F}(\rho) \geq p_0 \mathcal{F}(\rho') + \sum_{j=1}^{N} p_j \mathcal{F}(\rho_j) \geq p_0 \mathcal{F}(\rho') \implies p_0 \leq P^{max}(\rho \to \rho') \leq \frac{\mathcal{F}(\rho)}{\mathcal{F}(\rho')} \ .$$

Obviously, the question we started from at the beginning of the section corresponds to the particular case in which we ask whether p_0 can be equal to 1.
Another kind of conversions we will deal with in the following are **asymptotic conversions**. In this case, the setting is much different: we start from n identical input states, $\rho^{\otimes n}$, and we aim for obtaining as much copies of the output state as possible, within a certain error. Then, we tke the limit $n \to \infty$, and we require the error in the preparation of the output to vanish in the limit. For the sake of precision, let us present the two following definitions.

Definition 2.19. *Within a certain Quantum Resource Theory, the* ***asymptotic conversion rate*** *from state $\rho \in \mathcal{D}(\mathcal{H})$ to $\rho' \in \mathcal{D}(\mathcal{H}')$ is defined as*

$$R(\rho \to \rho') \coloneqq \sup \left\{ r : \lim_{n\to\infty} \inf_{\Lambda \in \mathcal{O}_f(\mathcal{H}^{\otimes n} \to \mathcal{H}'^{\otimes \lfloor rn \rfloor})} \left\| \Lambda(\rho^{\otimes n}) - \rho'^{\otimes \lfloor rn \rfloor} \right\|_1 = 0 \right\}.$$

Definition 2.20. *Within a certain Quantum Resource Theory, the* ***maximal asymptotic conversion rate*** *from state $\rho\mathcal{D}(\mathcal{H})$ to $\rho'\mathcal{D}(\mathcal{H}')$ is defined as*

$$\widetilde{R}(\rho \to \rho') \coloneqq \sup \left\{ r : \lim_{n\to\infty} \inf_{\Lambda \in \mathcal{O}_f(\mathcal{H}^{\otimes n} \to \mathcal{H}'^{\otimes \lfloor rn \rfloor})} \sup_{j=1,\ldots,\lfloor rn \rfloor} \left\| \mathrm{Tr}_{\hat{j}} \left(\Lambda(\rho^{\otimes n}) \right) - \rho' \right\|_1 = 0 \right\}.$$

Note that the integer part $\lfloor \cdot \rfloor$ is needed in order to obtain an integer number, but it becomes irrelevant in the limit. In (2.20), as opposed to (2.19), we require only the error associated to each copy to go to 0, while the global error might even diverge. It is sensible reasonable choice whenever we want to use the output copies independently from one another, for example if they have to be distributed to different non-interacting parties. It trivially holds that $\widetilde{R}(\rho \to \rho') \geq R(\rho \to \rho')$. Moreover, we also have $P^{max}(\rho \to \rho') \leq R(\rho \to \rho')$, since we convert each input state independently with some probabilisitic protocol in the asymptotic setting as well. These quantities can be bounded using resource monotones, as the results below show.

Theorem 2.21. *Let us assume that the resource monotone $\mathcal{F}$ is:*

- *asymptotically continuous;*
- *weakly additive;*

Then:

$$R(\rho \to \rho') \leq \frac{\mathcal{F}(\rho)}{\mathcal{F}(\rho')}. \tag{2.11}$$

Proof. Let us consider two quantum states $\rho \in \mathcal{D}(\mathcal{H})$ and $\rho' \in \mathcal{D}(\mathcal{H}')$, with $\dim \mathcal{H}' = d'$. Let us assume that it exists a sequence of free operations $\{\Lambda_n\}_n$ such that $\Lambda_n \in \mathcal{O}_f(\mathcal{H}^{\otimes n} \to \mathcal{H}'^{\otimes \lfloor rn \rfloor})$ and:

$$\lim_{n\to\infty} \|\Lambda_n(\rho^{\otimes n}) - \rho'^{\otimes \lfloor rn \rfloor}\|_1 = 0.$$

We define $\epsilon_n \coloneqq \|\Lambda_n(\rho^{\otimes n}) - \rho'^{\otimes \lfloor rn \rfloor}\|_1$; then we have:

$$\begin{aligned}
\mathcal{F}(\rho) &= \lim_{n\to\infty} \frac{1}{n} \mathcal{F}(\rho^{\otimes n}) \\
&\geq \lim_{n\to\infty} \frac{1}{n} \mathcal{F}\left(\Lambda_n(\rho^{\otimes n})\right) \\
&= \lim_{n\to\infty} \frac{1}{n} \left[\mathcal{F}(\rho'^{\otimes \lfloor rn \rfloor}) + c\epsilon_n \log \left(d'^{\lfloor rn \rfloor} \right) + \eta\left(\epsilon_n\right) \right] \\
&= \lim_{n\to\infty} \left[\frac{\lfloor rn \rfloor}{n} \left[\mathcal{F}(\rho') + c\epsilon_n \log d' \right] + \frac{\eta(\epsilon_n)}{n} \right] \\
&= r\mathcal{F}(\rho').
\end{aligned}$$

□

Remark 2.22. *In the previous proof it was crucial the logarithmic scaling of the continuity bound. This is why simple continuity in trace norm is not enough, and asymptotic continuity is required instead.*

Sometimes, asymptotic continuity simply does not hold, or it does not even make sense, for example when we consider infinite dimensional Resource Theories. Luckily, the following result holds.

Theorem 2.23. *Let us assume that the resource monotone $\mathcal{F}$ is:*

- *lower semi-continuous;*
- *weakly additive;*
- *strongly super-additive.*

Then:

$$R(\rho \to \rho') \leq \widetilde{R}(\rho \to \rho') \leq \frac{\mathcal{F}(\rho)}{\mathcal{F}(\rho')} \,. \tag{2.12}$$

Proof. Again, let us consider two quantum states $\rho \in \mathcal{D}(\mathcal{H})$ and $\rho' \in \mathcal{D}(\mathcal{H}')$. Let us assume that it exists a sequence of free operations $\{\Lambda_n\}_n$ such that $\Lambda_n \in \mathcal{O}_f(\mathcal{H}^{\otimes n} \to \mathcal{H}'^{\otimes \lfloor rn \rfloor})$ and:

$$\lim_{n\to\infty} \|\Lambda_n(\rho^{\otimes n}) - \rho'^{\otimes \lfloor rn \rfloor}\|_1 = 0\,.$$

We define $\epsilon_n \coloneqq \sup_{j=1,\dots,\lfloor rn \rfloor} \| \operatorname{Tr}_{\hat{j}} \Lambda_n(\rho^{\otimes n}) - \rho'\|_1$; then we have:

$$\begin{aligned}
\mathcal{F}(\rho) &= \lim_{n\to\infty} \frac{1}{n} \mathcal{F}(\rho^{\otimes n}) \\
&\geq \lim_{n\to\infty} \frac{1}{n} \mathcal{F}\left(\Lambda(\rho^{\otimes n})\right) \\
&\geq \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{\lfloor rn \rfloor} \mathcal{F}\left(\operatorname{Tr}_{\hat{j}} \Lambda(\rho^{\otimes n})\right) \\
&\geq \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{\lfloor rn \rfloor} \left[\mathcal{F}(\rho') + \mathcal{O}(\epsilon_n)\right] \\
&\geq \lim_{n\to\infty} \frac{\lfloor rn \rfloor}{n} \left[\mathcal{F}(\rho') + \mathcal{O}(\epsilon_n)\right] \\
&= r\mathcal{F}(\rho')\,.
\end{aligned}$$

□

Finally, we mention the fact that resource monotones fulfilling the tensorisation condition (see Definition 2.18) can be used to prove that **distillation** of a resurce, i.e., the conversion of many states into a single, more resourceful, one, is impossible with only free operations. Indeed, if $\mathcal{F}$ satisfy tensorisation, we have:

$$\mathcal{F}(\rho^{\otimes n}) = \mathcal{F}(\rho) \geq \mathcal{F}(\Lambda(\rho))\,,$$

for any $n \in \mathbb{N}$, no matter how big it is. Loosely speaking, Resource Theories displaying tensorisation can be seen as physical realisations of the common saying "it's quality, not quantity".

2.3 Some celebrated examples

2.3.1 Entanglement

As we already anticipated, the operational limitations which select entanglement as a quantum resource are those of a "distant laboratories and classical phones" scenario. Namely, two or more experimentalists cannot make their quantum systems interact directly, but they can share classical probability distributions over a classical channel (using standard telecom communication, for instance), and perform operations according to them.

The quantum channels one can construct with these restrictions are collectively called **local operations and classical communication** (**LOCC**) channels. In the case of N laboratories $A_1, \dots, A_N$, and by composing local (i.e., in a tensor product form) operations according to some shared probability distribution $\{p_a\}_a$, we can immediately see that LOCC channels admit a Kraus decomposition in the following form:

$$\Lambda(\cdot) = \sum\nolimits_a \left(\bigotimes_{n=1}^{N} K_{a,n}^{A_n} \right) (\cdot) \left(\bigotimes_{n=1}^{N} \left(K_{a,n}^{A_n}\right)^\dagger \right) . \tag{2.13}$$

Hence, if we starting from a quantum state which can be written as follows:

$$\sum\nolimits_a p_a \bigotimes_{n=1}^{N} \rho_{a,n}^{A_n} , \quad \sum\nolimits_a p_a = 1 , \quad p_a \geq 0 \, \forall \, a , \tag{2.14}$$

and we apply operations in the form (2.13) we will always end up again in a state admitting a decomposition as in (2.14). Such states are known as **separable states** [57], and we will denote their set with $\mathcal{S}(\mathcal{H}) \subset \mathcal{D}(\mathcal{H})$. Conversely, a state in the form (2.14), can always be generated with only LOCC operations starting from a generic multipartite state: since any local operation is allowed, we can just discard the state we have at the beginning and command, accordingly to some a priori fixed probability distribution $\{p_a\}_a$, the experimentalist A_n to prepare his system in the state $\rho_{a,n}^{A_n}$. This simple argument shows that, if LOCC channels are taken to be our free operations, we necessarily have:

$$\mathcal{D}_f^E(\mathcal{H}) = \mathcal{S}(\mathcal{H}) . \tag{2.15}$$

Note that the condition $\sum_a p_a = 1$ in (2.14) must be valid for any pysical state, while is the $p_a \geq 0$ condition which distinguishes separable ones from the others. Entanglement theory is an example of how choosing the set of free operations leaves no freedom in the definition of the set of free states. It is also a convex Resource Theory.

Now we could ask ourselves whether LOCC channels are the most general quantum operations mapping separable states into separable states. In fact, they are not, but it is not completely trivial to see why. The set of LOCC operations is strictly smaller than the set of operations which can be written as in (2.13), i.e., **separable operations**. Perhaps surprisingly, neither separable operations are the most general operations which cannot generate entanglement. The set of all operations mapping $\mathcal{S}(\mathcal{H})$ into itself is unsurprisingly known as the set of **non-entangling operations**.

To understand why this set is bigger than those of both LOCC and separable channels, let us notice that if Λ^S is a separable operation (and then a LOCC one), also $\Lambda^S \otimes I^E$ is so. So, it cannot generate entanglement starting from separable states even when acting on a part of a larger system. Now, ley us consider two isomorphic Hilbert spaces $\mathcal{H}^{A,B}$, and the **swap map**:

$$|j^A k^B\rangle \mapsto |k^A j^B\rangle \, .$$

Clearly, this map will preserve the form (2.14) of a quantum state of the system AB, and hence it is a non-entangling operation. Nonetheless, when acting on subsystems A and B of the larger system $AA'BB'$ prepared in the separable (with respect to the bipartition AA'/BB') state:

$$\rho \propto \left(|0^A 0^{A'}\rangle + |1^A 1^{A'}\rangle\right) \otimes \left(|0^B 0^{B'}\rangle + |1^B 1^{B'}\rangle\right) \, ,$$

it will produce a non-separable (with respect to the same bipartition) state. Hence, it cannot be a separable (or LOCC) operation [43].

Here we have a good example of how sometimes it is preferable to enlarge the set of free operations in order to characterize it easily. Determining whether a generic operation can be constructed with only LOCC protocols is indeed a very hard problem, while separable operations, for instance, have a straightforward parametrisation. Doing so, however, we encompass also physical maps which do not have a clear operational justification, such as generic non-entangling operations. As we already pointed out, this allow us to prove only weaker results than with LOCC, but with much less effort. So, there is a trade-off between the manageability of the maps we define to be free in our Resource Theory, and their relation with reality: where the optimum is strongly depends on the result we want to prove.

For later convenience, let us introduce and discuss very briefly two very important proposals for entanglement measures. We will focus on the bipartite setting, i.e., when only two parties are considered.

- **Relative entropy of entanglement**. [61]. This measure of entanglement is based on Definition 2.3:

  $$E_r(\rho^{AB}) := \inf_{\sigma \in \mathcal{S}(\mathcal{H}^A \otimes \mathcal{H}^B)} D(\rho \| \sigma) \, .$$

  Thus, we have already proven that it is indeed a proper resource monotone. It is also convex (as a consequence of the joint convexity of $D(\cdot\|\cdot)$), weakly sub-additive (see Proposition 2.16) and, in the finite-dimensional case, asymptotically continuous [78]. So, by virtue of Theorem 2.21, its regularized version can be used to upper bounds asymptotic conversion rates:

  $$R(\rho \to \rho') \le \frac{E_r^\infty(\rho)}{E_r^\infty(\rho')} \, .$$

- **Logarithmic negativity**. Logarithmic negativity is defined starting from the partial transpose of a state ρ:

  $$\rho = \sum\nolimits_{j,k,l,h} \rho_{\substack{jk \\ lh}} \, |jl\rangle\langle kh| \mapsto \rho^\Gamma := \sum\nolimits_{j,k,l,h} \rho_{\substack{jk \\ lh}} \, |jh\rangle\langle kl| \, . \tag{2.16}$$

  As we already pointed out at the very beginning of this work, transposition is a positive-but-not-completely-positive map: it maps physical states into physical states only whe acting on the whole system. It can also be proven that if ρ is separable, ρ^Γ is a physical state, but the converse is not true [76, 77]. This means that the logarithmic negativity, defined as follows:

$$E_N(\rho) \coloneqq \log \|\rho^\Gamma\|_1 \,, \tag{2.17}$$

is not a faithful resource monotone. However, entanglement cannot be distilled from non-separable states with vanishing logarithmic negativity: they display **bound entanglement**[79].
Despite the fact that it is not even convex, it displays strong monotonicity [80]. Moreover, it is one of the few entanglement monotone that can be computed efficiently. Note that neither the choice of the basis nor that of the subsystem influence E_N.

2.3.2 Coherence

As for the Quantum Resource Theory of coherence, there is an obvious choice for the set of free states. The decoherence process is modeled by the **totally dephasing channel**:

$$\Delta(\cdot) = \sum_{n=1}^{d} |n\rangle\!\langle n| \, (\cdot) \, |n\rangle\!\langle n| \ , \tag{2.18}$$

with $d = \dim \mathcal{H}$ and with $\{|n\rangle\}$ being the incoherent basis. Given a generic quantum state ρ expressed in the incoherent basis, the totally dephasing channel kills all of its off-diagonal elements. Free states are those which are left unchanged by Δ, i.e., those which are stable under the open dynamic of the system, and hence are easy to preserve:

$$\mathcal{D}_f^C(\mathcal{H}) = \left\{ \rho \in \mathcal{D}(\mathcal{H}) : \rho = \sum_{n=1}^{d} p_n \, |n\rangle\!\langle n| \ , \, p_n \geq 0 \, \forall \, n \right\} . \tag{2.19}$$

Δ fulfills the following conditions:

- $\Delta(\sigma) = \sigma \ \forall \ \sigma \in \mathcal{D}_f(\mathcal{H})$;
- $\Delta(\rho) \in \mathcal{D}_f(\mathcal{H}) \ \forall \ \rho \in \mathcal{D}(\mathcal{H})$.

In a general Resource Theory, such a map is called a **resource-destroying map**. A necessary (but not sufficient) condition for a Resource Theory for admitting a linear resource-destroying map is to be affine. Indeed, if $\{\sigma_a\}_a$ are free states, then $\Delta\left(\sum_a c_a \sigma_a\right) = \sum_a c_s \Delta(\sigma_a) = \sum_a c_a \sigma_a$. So, affine combinations of free states must be free, i.e., the Resource Theory must be affine. For necessary and sufficient conditions for admitting a linear resource-destroying map see [81].
Coherence theory is also a striking example of how starting from $\mathcal{D}_f(\mathcal{H})$ leaves the door open for a plethora of possible sets of free operations. We report some examples below.

- **Maximally incoherent operations** (**MIO**) [62]. This is the maximal set of free operations, i.e., it encompasses any channel Λ such that

$$\Lambda(\mathcal{D}_f^C(\mathcal{H})) \subseteq \mathcal{D}_f^C(\mathcal{H}') \,.$$

- **Strictly incoherent operations** (**SIO**) [67]. An operation Λ is in SIO if it can be written by means of Kraus operators K_a such that:

$$\langle n|K_a\rho K_a^\dagger|n\rangle = \langle n|K_a\Delta(\rho)K_a^\dagger|n\rangle \ ,$$

  for any incoherent state $\{|n\rangle\}_n$. In other words, SIO are all those operations which cannot exploit the coherence of the input state in order to obtain an effect which is detectable by an incoherent measurement. In [82], an operational construction for any element in SIO is given.
  We cited this set of free operations because it provides an example of a Resource Theory admitting tensorisation. In [83], it was introduced the following monotone:

$$\eta(\rho) := \max_{j\neq k} \frac{\langle j|\rho|k\rangle}{\sqrt{\langle j|\rho|j\rangle \langle k|\rho|k\rangle}} \ ,$$

  where again $|j\rangle$ and $|k\rangle$ are incoherent states. It can be shown that η is indeed a monotone under SIO, and that:

$$\eta(\rho_1 \otimes \rho_2) = \max\{\eta(\rho_1), \eta(\rho_2)\} \ ,$$

  and hence, in particular, $\eta(\rho^{\otimes n}) = \eta(\rho)$. Thanks to this property, it can br proved that there are states from which we cannot distill coherence: they display **bound coherence** (under SIO).

- **Physical incoherent operations** (**PIO**). They are all the incoherent operations admitting a Stinespring dilation which is incoherent as well. This means that if Λ is in PIO, it can be written as follows:

$$\Lambda(\cdot) = \mathrm{Tr}_E \left[U \left((\cdot) \otimes \sigma \right) U^\dagger \right] \ ,$$

  with σ being an incoherent state and U being an incoherent (i.e., diagonal in the incoherent basis) unitary. We cited this set of free operations because the notion of "free dilation" will come back in the next chapter.

It can be proven [84] that PIO⊂SIO⊂MIO.

2.3.3 Thermodynamics

Given a quantum system with Hilbert space $\mathcal{H}$, Hamiltonian H and inverse temperature β, the set of free states for the Resource Theory of Thermodynamics is just the thermal state $\tau_H := e^{-\beta H}$:

$$\mathcal{D}_f^T(\mathcal{H}) = \{\tau_H\} \ . \tag{2.20}$$

Also in this case there is more than just one possible set of free operations. For example, we call **thermal operations** (acting on the system S) the set of physical maps Λ which act on a state ρ^S as follows:

$$\Lambda(\rho^S) = \mathrm{Tr}_E \left[U^{SE} \left(\rho^S \otimes \tau^E_{H^E} \right) U^{SE\dagger} \right] \ , \tag{2.21}$$

where $\tau^E_{H^E}$ is the thermal state of the system E for the Hamiltonian H^E, and U^{SE} is a unitary such that:

$$[U^{SE}, H^S \otimes I^E + I^S \otimes H^E] = 0\,.$$

By construction, thermal operations have a free dilation. The maximal set of free operations is represented by **Gibbs-preserving operations**, i.e., all those operations which leave the thermal state invariant.
Once again, a good resource monotone can be obtained via the relative entropy:

$$T_r(\rho) \coloneqq D(\rho\|\gamma)\,. \tag{2.22}$$

Then, if the dimension of $\mathcal{H}$ is finite, the following result can be proven.
Theorem 2.24. *[69, Theorem 1] Using only thermal operations at background temperature T, asymptotic conversion at non-zero rate is possible between all non-thermal states ρ and σ of a system with Hamiltonian H. Being τ_H the thermal state, the optimal rate is given by:*

$$R(\rho \to \sigma) = \frac{D(\rho\|\tau_H)}{D(\sigma\|\tau_H)}\,. \tag{2.23}$$

It is easy to see that $D(\rho\|\tau_H) = \beta F_\beta(\rho) - \beta F_\beta(\tau_H)$, with $F_\beta(\rho) \coloneqq \langle H\rangle_\rho - \frac{1}{\beta}S(\rho)$ being the **free energy** of the state ρ.

3 Non-classicality as a resource

3.1 What is non-classicality?

3.1.1 Classical and non-classical states

As we pointed out in section 1.2.2, there are several reasons to consider coherent states as the most classical among the states of a quantum harmonic oscillator. By linearity, this can be extended to probabilistic mixtures of coherent states. We then define the set of **classical states** to be the closure (in trace norm) of the convex hull of the set of all coherent states [85]:

$$\mathcal{C}_m = \mathcal{C}(\mathcal{H}_m) := \overline{\mathrm{conv}}\,\{|\alpha\rangle\!\langle\alpha| : \alpha \in \mathbb{C}^m\}\ , \tag{3.1}$$

with $\mathcal{H}_m = L^2(\mathbb{R}^m)$ (we will denote $\mathcal{H}_1 = \mathcal{H}$ for simplicity). Any Quantum Resource Theory of **optical non-classicality** (it will be usually referred simply as non-classicality in the following for simplicity) [86] start from the following identification:

$$\mathcal{D}_f^{NC}(\mathcal{H}_m) = \mathcal{C}_m\,. \tag{3.2}$$

The Resource Theory of non-classicality is, by construction, a convex Resource Theory, but obviously it is not affine. Different choices for $\mathcal{O}_f$ lead to different Resource Theories. In the following, we will make use of some results and terminology of distribution Theory: a brief review of the topic is presented in Appendix B.
Every classical state $\sigma \in \mathcal{C}_m$ can be represented as

$$\sigma = \int d^{2m}\alpha\, P(\boldsymbol{\alpha})\, |\alpha\rangle\!\langle\alpha|\ , \tag{3.3}$$

with $P(\boldsymbol{\alpha})$, which is the P-function of the classical state σ (see section 1.3.3), being a proper probability distribution, i.e., a normalised and non-negative measure. Loosely speaking, non-classical states are usually said to have a "negative" or "more singular than a δ function" P-function. Actually, we can give a rigorous meaning to these colloquial expressions: "negative" distributions are those to which we cannot associate a positive sign at any point in phase space, while any distribution of order at least 1 is said to be "more singular than a δ function". From Proposition B.3 we see that the latter condition implies the former one, but the converse is not true: in the following, we will se an example of a state with a "negative" P-function, which is not "more singular than a δ". Since the P-function is defined as the Fourier transform of a well-defined object, χ_1, and the Fourier transform is always unique as a distribution (once it is well-defined), this definition of singularity of the P-function is unambiguous (in earlier works some confusion has been made, see for example [87]).

At any rate, dealing with the P-function of a state is a bit awkward because it can be a highly singular object. Thanks to the famous Bochner's theorem [88], we have an alternative characterisation of classical states via characteristic functions instead of quasi-probability distributions: σ is a classical state if and only if its χ_1 function is **positive-definite**, i.e., if and only if the matrix

$$M_{jk} = \chi_1^\sigma(\boldsymbol{\alpha}_j - \boldsymbol{\alpha}_k) \tag{3.4}$$

is positive definite for any choice of $\{\alpha_j\}_{j=1,\dots,n}$. This characterisation has the advantage that χ_1, contrarily to P, is well-defined and regular for any quantum state. Note that if we choose $\boldsymbol{\alpha}_1 = \mathbf{0}$ and $\boldsymbol{\alpha}_2 = \boldsymbol{\alpha}$ the positive-definiteness condition immediately imply that:

$$0 \le \det \begin{pmatrix} \chi_1^\sigma(\mathbf{0}) & \chi_1^\sigma(\boldsymbol{\alpha}) \\ \chi_1^\sigma(-\boldsymbol{\alpha}) & \chi_1^\sigma(\mathbf{0}) \end{pmatrix} = 1 - |\chi_1^\sigma(\boldsymbol{\alpha})|^2 \,,$$

where we used $\chi_1(\mathbf{0}) = 1$ and $\chi_1(-\boldsymbol{\alpha}) = \chi_1^*(\boldsymbol{\alpha})$. So, we can state the following result.

Lemma 3.1. *If $\sigma \in \mathcal{C}_m$ then $|\chi_1^\sigma(\boldsymbol{\alpha})| \le 1$ for any $\boldsymbol{\alpha} \in \mathbb{C}^m$.*

Non-classicality can manifest itself in the P-function in different ways. For example, let us consider the thermal state with the vacuum removed:

$$\widetilde{\tau}_\nu = \frac{1}{\nu} \sum_{n=1}^{\infty} \left(\frac{\nu}{1+\nu} \right)^n |n\rangle\langle n| \,. \tag{3.5}$$

The state is non-classical: its P-function can be inferred from equations (1.41) and (1.46):

$$P^{\widetilde{\tau}_\nu}(\alpha) = \frac{1+\nu}{\pi \nu^2} e^{-\frac{|\alpha|^2}{\nu}} - \frac{1}{\nu} \delta^2(\alpha) \,, \tag{3.6}$$

and clearly it is not positive at $\alpha = 0$ (actually, it does not have a well defined weak sign). Moreover, no classical state can have vanishing overlap with the vacuum [89], therefore the state is definitely non-classical. Nonetheless, its P-function is a well-defined tempered distribution, and it is not even more singular than a δ function.

Another example is given by Fock states. From equation (1.49) we see that the P-function of $|n\rangle$ is a tempered distribution which must be expressed in terms of derivatives of δ functions: it is not properly negative at some point, since it vanishes everywhere but in 0, where its weak sign is not even defined.

At last, we cite the case of squeezed and cat states, whose χ_1 functions diverge exponentially. Hence, their P-functions cannot be defined in the space of tempered distributions either.

Note that the Resource Theory of non-classicality is somehow similar to that of coherence, as in both cases the starting point is the identification of a preferred basis and the definition of $\mathcal{D}_f$ as its closed convex hull. However, there are two important differences:

- it is necessarily formulated in infinite-dimensional Hilbert spaces;
- the preferred basis is not an orthonormal one, and it is even over-complete.

As we will see, these two points will lead to many additional difficulties. We will come back to this analogy later in this work.

3.1.2 Classical and non-classical operations

The maximal set of free operations $\mathcal{O}_f^{max}$ is that of **classicality-preserving operations** (**CPO**). Unfortunately, there is not a clear characterisation or a strong operational motivation for these operations. On the contrary, a set of operations which is well justified from an operational point of view is that of linear optical unitaries, discussed in section 1.2.4. Indeed, as we already pointed out, they are easily implementable in linear optical settings. Moreover, as proved in Proposition 1.11, they are the only unitaries which preserve the set of coherent states: this also implies that they are the only unitary CPOs. Therefore, by Stinespring dilation, we see that any quantum channel which can be physically implemented by means CPOs only must be obtained from a linear optical unitary applied to a larger system. More precisely, the only trace-preserving CPOs whose dilation is still a CPO are those which can be written as follows:

$$\Lambda_L(\cdot) = \mathrm{Tr}_E \left[U_L \left((\cdot) \otimes \sigma^E \right) U_L^\dagger \right] ,$$

where σ^E is a classical state of $\mathcal{H}^E$ and U_L is a linear optical unitary. We call these operations **linear optical operations**. We can encompass maps giving a flagged outcome simply by taking as ancillary state $\sigma^{EE'}$, where $\mathcal{H}^{E'}$ is the classical system containing the classical flags.
In order to describe a more experimentally significant scenario, we might want to allow for multiple linear optical operations with feed forward between them, i.e., the conditioning of an operation to be performed on a state based on the result of a previous measurement. In [17], the following sets of free operations are proposed:

- P_0: any operations obtained by appending a classical ancilla, performing a linear optical unitary and tracing out a set of modes;
- P_1: the set of modes to be traced out can be measured first, and the information about the outcome can be totally or partially retained (in other words, destructive measurements are allowed);
- $P_{\mathbb{N}}$: multiple operations of P_1 can be performed subsequently, and feed-forward is allowed.

The following result ensures that $P_{\mathbb{N}}$ is a subset of the set of CPOs.
Theorem 3.2. *[17, Theorem 1] Every quantum operation from $\mathcal{B}(\mathcal{H}_m)$ to $\mathcal{B}(\mathcal{H}_{m'})$ in $P_{\mathbb{N}}$ admits a set of Kraus operators in the following form:*

$$K_n \left|\boldsymbol{\alpha}\right\rangle = c_n(\boldsymbol{\alpha}) \left|M_n \boldsymbol{\alpha} + \boldsymbol{\delta}_n\right\rangle ,$$

where $c_n(\boldsymbol{\alpha}) \in \mathbb{C}$, $\boldsymbol{\delta}_n \in \mathbb{C}^{m'}$ and M_n is a $m \times m'$ matrix with singular value not exceeding 1.

Physically, this result implies that any operation in $P_{\mathbb{N}}$ results in a combination of contractions and displacements in phase space. From this theorem we immediately see that any coherent state is mapped into a classical state by any element of $P_{\mathbb{N}}$. By writing any classical state in terms of its P-function, it is easy to deduce than any element in $P_{\mathbb{N}}$ is also a CPO. However, it is not clear whether the inclusion of $P_{\mathbb{N}}$ in the set of CPOs is strict or not. In any case, the previous result gives a strong motivation to the

Resource Theory of non-classicality, by proving that this quantum resource cannot be generated when only operations in $P_{\mathbb{N}}$ are available, which is a plausible assumption in many situations. Actually, in standard quantum optical settings we cannot implement all the operations in $P_{\mathbb{N}}$ either, since the condition of "arbitrary destructive measurements" is way too loose. At any rate, any result obtained with this set of operations holds true when only a subset of them are actually available.
An example of a CPO is the totally dephasing map with respect to the Fock basis. Indeed, from the following identity:

$$\frac{1}{2\pi}\int_0^{2\pi} d\varphi\, e^{i\varphi \hat{a}^\dagger \hat{a}} |n\rangle\langle m| \, e^{-i\varphi \hat{a}^\dagger \hat{a}} = \frac{1}{2\pi}\int_0^{2\pi} d\varphi\, e^{i\varphi(n-m)} |n\rangle\langle m| = \delta_{n,m} |n\rangle\langle m| \ ,$$

we see that the action of Δ (for a single mode state) can be written as follows:

$$\Delta(\rho) = \frac{1}{2\pi}\int_0^{2\pi} d\varphi\, e^{i\varphi \hat{a}^\dagger \hat{a}} \rho e^{-i\varphi \hat{a}^\dagger \hat{a}} \ .$$

We can use this expression to prove that Δ is a classical operation, and in particular $\Delta(\mathcal{C}_m) \subset \mathcal{C}_m^{FD}$, where $\mathcal{C}_m^{FD}$ is the set of states in $\mathcal{C}_m$ which are also diagonal in the Fock basis. Indeed:

$$\begin{aligned}\Delta(\rho) &= \int d^2\alpha\, P(\alpha) \int_0^{2\pi} d\varphi\, e^{i\varphi \hat{a}^\dagger \hat{a}} |\alpha\rangle\langle\alpha| \, e^{-i\varphi \hat{a}^\dagger \hat{a}} \\ &= \int d^2\alpha\, P(\alpha) \int_0^{2\pi} d\varphi\ |e^{i\varphi}\alpha\rangle\langle e^{i\varphi}\alpha| \\ &= \int d^2\alpha \int_0^{2\pi} d\varphi\, P(e^{-i\varphi}\alpha)\ |\alpha\rangle\langle\alpha| \ .\end{aligned}$$

Clearly, if $P(\alpha)$ is a positive and normalized measure, also $\int_0^{2\pi} d\varphi P(e^{-i\varphi}\alpha)$ is so.
Note that Δ is a resource-destroying map for the Resource Theory of coherence, but not for the Resource Theory of non-classicality: neither it leaves any free state invariant, nor it maps any non-free state into a free one. Actually, no physical resource-destroying map is allowed in the resource Theory of non-classicality, since it is not affine. At any rate, we can define the following linear map:

$$\Phi(\rho) = \frac{1}{\pi}\int d^2\alpha\ |\alpha\rangle\langle\alpha| \, \rho \, |\alpha\rangle\langle\alpha| \ .$$

Its net effect is to substitute the P-function of a state with its Q-function: since the latter is always non-negative and well-behaved, the resulting state is always classical. So, Φ is a CPO. However, $\Phi(\rho) \neq \rho$ even if ρ is classical, so it is not a resource-destroying map.

3.2 Quantum Resource Theories of non-classicality

3.2.1 P-negativity and the absolute robustness of non-classicality

The first Resource Theory of non-classicality we want to consider [90] is based on a very intuitive idea: if the P-function of a non-classical state is not non-negative at any point

in phase space, it is reasonable to quantify the degree of non-classicality by measuring the negative volume of P. However, as we already thoroughly discussed, P is usually a highly singular object whose "negative part" might not be easily identifiable.
The solution is to consider the **filtered P-functions** introduced in [91], which consists of a regularization of the actual P-functions, in such a way that the negative part is well-defined. Since singularities of the P-function arise from divergences of the χ_1 function (with "singularities" we always refer to "beyond δ function singularities"), it seems a good idea to multiply the latter for a **filter** $\Omega_\omega(\alpha)$ to make it bounded. For a single mode quantum state, the filter has to satisfy the following conditions (note that [90] uses a slightly different convention for the symplectic Fourier transform):

N1) $\Omega_\omega(\alpha)$ can be factorized as $\Omega^1_\omega(\alpha)\Omega^2_\omega(\alpha)$, with $\Omega^1_\omega(\alpha)$, $\Omega^2_\omega(\alpha) \in L^2(\mathbb{C})$;

N2) $\Omega^1_\omega(\alpha)e^{\frac{|\alpha|^2}{2}} \in L^2(\mathbb{C})$;

N3) $\Omega_\omega(0) = 1$ and $\lim_{\omega\to\infty} \Omega_\omega(\alpha) = 1\,\forall\,\alpha \in \mathbb{C}$.

Points N1) and N2) are justified by the fact that since $\mathcal{D}(\alpha)$ is a unitary operator for any α, it holds $|\chi^\rho_0(\alpha)| \leq 1$ for any ρ. So, $\chi^\rho_1(\alpha)$ does not diverge faster than $e^{\frac{|\alpha|^2}{2}}$. Point N3) is needed in order to obtain a sensible regularization, i.e., to go back to the non-filtered P-function in some limit.
Starting from the filter Ω_ω, we can finally construct the filtered P-function P_ω by taking the symplectic Fourier transform. To begin with, we report the following result about P_ω.
Theorem 3.3. *[90, Theorem 1] If Ω_ω satisfies N1) and N2), then P_ω has no singularities and is finite everywhere.*

Hence, P_ω can now be used to construct the following, well-defined, quantity.
Definition 3.4. *The **negativity** of a state ρ is defined as:*

$$\mathcal{N}(\rho) := \lim_{\omega\to\infty} \int d^2\alpha P^-_\omega(\alpha) \tag{3.7}$$

for some filter Ω_ω, and with P^-_ω being the negative part of the filtered P-function P_ω.

A fundamental question that we may ask at this level is whether $\mathcal{N}$ is independent or not from the filter Ω_ω. We will see in the following that this is indeed the case. Moreover, $\mathcal{N}$ can be used as a resource monotone in a Resource Theory of non-classicality, as the following result shows.
Theorem 3.5. *[90, Theorem 2] If we identify the set of free operations $\mathcal{O}_f$ with the set of linear optical operations, the negativity $\mathcal{N}$ becomes a resource monotone. Moreover, it satisfies:*

- *faithfulness;*
- *strong monotonicity;*
- *convexity.*

The negativity can be related to another non-classicality monotone, i.e., the celebrated absolute robustness. Indeed, the following result holds true.

Theorem 3.6. *[90, Theorem 3]* $\mathcal{N}(\rho) = \mathcal{R}_A^{NC}(\rho)\,\forall\,\rho \in \mathcal{D}(\mathcal{H}_m)$, *where* $\mathcal{R}_A^{NC}(\rho)$ *is the absolute robustness of non-classicality:*

$$\mathcal{R}_A^{NC}(\rho) := \inf_{\sigma\in\mathcal{C}_m} \left\{ r > 0 : \frac{\rho + r\sigma}{1+r} \in \mathcal{C}_m \right\} .$$

This Theorem also proves that the negativity does not depend on the choice of the filter, because $\mathcal{R}_A^{NC}(\rho)$ does not. This equivalence is quite intuitive: $\mathcal{N}$ quantifies, in some sense, the negative volume of the P-function of a state ρ; in order to obtain a classical state, such "negative part" must be compensated by adding a classical state σ, which have a completely positive P-function. Since the P-function of σ is normalized, its P-function has unital volume; the coefficient r must be such that $r\sigma$ can "fill up" the negative part of ρ.
Despite the premises, this Resource Theory has some major flaws. To begin with, $\mathcal{N}$ has no manifest continuity or additivity property. So, we can use $\mathcal{N}$ to upper bound conversion rates only on the exact, single-copy scenario; as we already discussed, this is seldom the most interesting case. Strong monotonicity allows for studying probabilistic protocols, too. The problem is that, as we are going to show in a moment, $\mathcal{N}$ is often divergent, making the bound $P^{max}(\rho \to \rho') \leq \mathcal{N}(\rho)/\mathcal{N}(\rho)$ meaningless. More precisely, the following result, contained in [92], hold true.
Proposition 3.7. *Let* $|\psi\rangle$ *be a non-classical pure state having a vanishing overlap with a finite (possibly empty) set of coherent states. Then,* $\mathcal{R}_A^{NC}(\psi) = \infty$.

Proof. Let us write $|\psi\rangle = \sum_{n=0}^{\infty} c_n |n\rangle$. The function $f(\alpha) = e^{|\alpha|^2/2}\langle\psi|\alpha\rangle = \sum_{n=0}^{\infty} \frac{c_n^*}{\sqrt{n!}}\alpha^n$ is a complex entire function of order at most 2 (otherwise, $|\langle\psi|\alpha\rangle|$ would diverge). If $N < \infty$ is the number of zeros of $f(\alpha)$, its Hadamard factorisation [93] becomes $f(\alpha) = e^{a\alpha^2+b\alpha}P_N(\alpha)$, where $P_N(\alpha)$ is a polynomial of degree N and $|a| < \frac{1}{2}$ in order for $|\langle\psi|\alpha\rangle|$ to be bounded. The Husimi Q-function of $|\psi\rangle$ is then

$$\begin{aligned} Q^{\psi}(\alpha) &= e^{-|\alpha|^2}|f(\alpha)|^2 \\ &= e^{-|\alpha|^2}e^{2\Re[a\alpha^2+b\alpha]}|P_N(\alpha)|^2 \\ &= e^{-\mathbf{r}(\alpha)^T A\mathbf{r}(\alpha)+\boldsymbol{\beta}^T\mathbf{r}(\alpha)}|P_N(\alpha)|^2 , \end{aligned} \tag{3.8}$$

with

$$\mathbf{r}(\alpha) = \begin{pmatrix}\Re\alpha \\ \Im\alpha\end{pmatrix}, \quad A = \begin{pmatrix}1-2\Re a & 2\Im a \\ 2\Im a & 1+2\Re a\end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix}2\Re b \\ -2\Im b\end{pmatrix} .$$

It is easy to see that the matrix A has eigenvalues $1 \pm 2|a|$. The Fourier transform of a function in the form (3.8) has again the same structure, but with A^{-1} in place of A. Now, let us suppose for the moment that $a > 0$. In this case, A has an eigenvalue strictly bigger than 1, and hence A^{-1} has one strictly smaller than 1. Thus, $\chi_1^{\psi} = e^{|\alpha|^2}\chi_{-1}^{\psi}$, is necessarily unbounded and, by virtue of Lemma 3.1, we conclude that $\mathcal{R}_A^{NC}(\psi) = \infty$. If instead $a = 0$, we have $A = A^{-1}$. At this point, we have to make a subsequent distinction: if P_N is the trivial polynomial, $|\psi\rangle$ is a coherent state and hence is obviously classical; if P_N is not trivial we end up once again with a divergent χ_1^{ψ} (in this case the divergence is polynomial instead of exponential). □

Among the states which fulfil the hypothesis of the last result there are, for instance, any finite superposition of Fock states and any non-classical Gaussian state. It is worth noticing that the cat state $|\psi_\alpha\rangle$ is a pure state with vanishing overlap with infinite coherent states, but with unbounded χ_1 and hence infinite absolute robustness of non-classicality. In any case, the regularization of P by means of a suitable filter, despite being very suggestive, turns out to be an overcomplication in practice. Indeed, whenever $\mathcal{N}(\rho)$ (and hence $\mathcal{R}_A^{NC}(\rho)$) is finite, it exists a finite r such that the state $\omega = \frac{\rho+r\sigma}{1+r}$ is classical, and hence its P-function P^ω is a (positive) measure. But also σ is a classical state, and hence its P-function P^σ is a (positive) measure as well. This implies that also $P^\rho = (1+r)P^\omega - rP^\sigma$ in a measure, and by Proposition B.3 we see that it has well-defined positive and negative parts, without the need for any filtering. The filter comes in handy only when such decomposition is not well-defined for the non-filtered P-function, but in those cases $\mathcal{N}$ is divergent and it cannot be used for stating any useful result.
Finally, we point out that the existence of states with infinite absolute robustness of non-classicality implies the following interesting result about the geometry of $\mathcal{C}_m$.
Theorem 3.8. *$\mathcal{C}_m$ does not have an internal part (in trace norm).*

3.2.2 Non-classicality and entanglement generation

The generation of entanglement by means of linear optical operations is a well-studied problem in entanglement theory. For example, in [94], a necessary and sufficient condition for the existence of a passive linear operation mapping a given gaussian state into a state having negative partial transpose is given. A formula for the maximum amount of entanglement generated (quantified with the logarithmic negativity) is also reported. Starting from this practical problem, an operationally meaningful Resource Theory of non-classicality can be constructed. Indeed, it is known that only non-classical states can generate entanglement by means of a beam splitter [95, 96, 97], which is the simplest two-port operation in Quantum Optics. More precisely, sending a quantum state ρ and the vacuum through a beam-splitter will generate an entangled state if and only if ρ is non-classical. So, the idea is now to "map" a Resource Theory of non-classicality into a Resource Theory of entanglement, by quantifying the maximum amount of entanglement that can be generated with a given state and free operations. In particular, in [98], the following experimental setup is considered:

- we start from an arbitrary single mode state ρ;
- we append a generic m-mode classical state;
- we perform an arbitrary passive linear unitary on the resulting state;
- we perform destructive measurements on all the modes but 2 (A and B);
- we quantify the entanglement generated between the two modes left with a suitable entanglement monotone.

The goal is to find the maximum (with respect to the chosen monotone) entanglement that can be generated starting from the state ρ. Clearly, such a functional (called **entanglement potential** and denoted with EP) assigning to ρ the maximum attainable entanglement is a monotone under linear optical maps and destructive measurements by construction. Optimizing over all possible entangling protocols would be an unfeasible task, but the problem can be greatly simplified, showing that the optimal protocol consists of only one beam splitter mixing the state ρ with a vacuum ancilla, and sending the output states to A and B. So, we just have to measure the maximum entanglement generated by a beam-splitter, whose transmittivity can be optimized analytically or numerically. The entanglement potential then simply becomes:

$$EP(\rho) = E\left(U_{BS}(\rho \otimes |0\rangle\langle 0|)U_{BS}^\dagger\right) , \tag{3.9}$$

where E is the chosen entanglement monotone.

3.2.3 Non-classicality and Quantum Metrology

A typical problem in Quantum Metrology consists in estimating a physical parameter ϑ whom a quantum operation Λ_ϑ depends upon. A solution is to prepare a quantum system in a state ρ, let it undergo such operation, measure the output and then try to infer the parameter ϑ. The estimated parameter will have a standard deviation $\Delta\vartheta$ associated to it. Now the question has become: how good is a given state ρ for this task, i.e., how small can $\Delta\vartheta$ can be? Classically, the problem of parameters estimation is related to the **Fisher Information**. For a classical probability distribution $p_\vartheta(x)$ depending on a parameter ϑ the Fisher Information (with respect to the parameter ϑ) is defined as follows:

$$F_\vartheta[p_\vartheta] := \int dx\, p_\vartheta(x)\left(\partial_\vartheta \log p_\vartheta(x)\right)^2 . \tag{3.10}$$

Then, the **Cramér-Rao bound** [99] ensures that:

$$(\Delta\vartheta)^2 \geq \frac{1}{F_\vartheta[p_\vartheta]} . \tag{3.11}$$

In a quantum setting, we can associate to any state an infinite number of classical probability distributions via POVMs. This adds an additional layer to the problem: not only the state, but also the POVM can be optimized. We then define the **Quantum Fisher Information** as follows:

$$F_\vartheta^Q[\rho] := \sup_{\mathcal{M} POVM} F_\vartheta\left[\mathcal{P}_\rho^\mathcal{M}\right] . \tag{3.12}$$

Now, it can be proven [100] that, if we adopt the protocol suggested at the beginning of this section, the bound (3.11) holds true for $\Delta\vartheta$ with F_ϑ^Q in place of F_ϑ.
Interestingly enough, one can find the following explicit characterization:

$$F_\vartheta^Q[\rho_\vartheta] = \mathrm{Tr}[\rho_\vartheta \mathcal{L}_\vartheta^2] , \tag{3.13}$$

where we defined $\rho_\vartheta := \Lambda_\vartheta(\rho)$, and the **symmetric logarithmic derivative** is implicitly defined by:

$$\frac{1}{2}\{\mathcal{L}_\vartheta, \rho_\vartheta\} = \partial_\vartheta \rho_\vartheta .$$

If Λ_ϑ is unitary operation generated by the operator ϑX, then one finds:

$$\frac{1}{2}\{\mathcal{L}_\vartheta, \rho_\vartheta\} = i[\rho_\vartheta, X] .$$

Even more surprisingly, there is an explicit expression for the Quantum Fisher Information in terms of the eigenvalues $\{p_j\}_j$ and eigenstates $\{\phi_j\}_j$ of ρ [100]:

$$F_{\vartheta}^{Q}[\rho_{\vartheta}] = 2\sum_{j,k} \frac{(p_j - p_k)^2}{p_j + p_k} \left|\langle\phi_j|X|\phi_k\rangle\right|^2 . \tag{3.14}$$

Finally, in the case of multiple parameters $\{\vartheta_j\}_j$ and observables $\{X_j\}_j$, we introduce the **Quantum Fisher Information Matrix**, defined as follows:

$$\mathbb{F}_{jk}[\rho_{\vartheta}] \coloneqq \frac{1}{2}\operatorname{Tr}\left[\{\mathcal{L}_j, \mathcal{L}_k\}\rho\right] , \tag{3.15}$$

where $\mathcal{L}_j$ is the symmetric logarithmic derivative associated to $e^{i\vartheta_j X_j}$. The following explicit expression can be proven:

$$\mathbb{F}_{jk} = 2\sum_{h,l} \frac{(p_h - p_l)^2}{p_h + p_l} \langle\phi_h|X_j|\phi_l\rangle \langle\phi_l|X_k|\phi_h\rangle . \tag{3.16}$$

In a recent proposal of a Resource Theory of non-classicality [17], applications of this quantum resource in Quantum Metrology has been explored, by exploiting the Quantum Fisher Information in order to build a resource monotone. We start from the Quantum Fisher Information Matrix $\mathbb{F}[\rho]$ associated to a m-mode quantum state ρ and the canonical operators $\hat{\boldsymbol{r}}$, as defined in (1.10). The matrix is computed with all the parameters ϑ_j set to 0. Then, we consider $[\mathbb{F} - 2\mathbb{I}]^+$, which is the positive part (obtained by taking the absolute values of the eigenvalues) of the matrix $\mathbb{F} - 2\mathbb{I}$. Finally, let T be an arbitrary subspace of the global phase space with fixed dimension, and let Tr_T be the trace over such subspace. We defined the quantity:

$$\mathcal{F}_j \coloneqq \frac{1}{4} \max_{T:\,\dim T = k} \operatorname{Tr}_T[\mathbb{F} - 2\mathbb{I}]^+ ,$$

for any $j \leq 2m$. If $f_j[\rho]$ is the j-th biggest eigenvalue of the matrix $[\mathbb{F} - 2\mathbb{I}]^+$, we can write $\mathcal{F}_j[\rho] = \sum_{k=1}^{j} f_k[\rho]$. Note that the operators $e^{i\theta_j \hat{r}_j}$ are just displacement operators, so we are basically quantifying the ability of ρ at sensing small displacements. Now, making use of the sets of free operations defined in section 3.1.2, the following result holds true.
Theorem 3.9. *[17, Theorem 3]*

- *Each one of the f_j is a monotone under P_0.*
- *Each one of the $\mathcal{F}_j$ is a monotone under P_1, and the probability p of going from ρ to ρ' with operations in P_1 is bounded as:*

$$p \leq \frac{\mathcal{F}_j[\rho]}{\mathcal{F}_j[\rho']} .$$

- *$\mathcal{F}_{2m}$ is a strong monotone under $P_{\mathbb{N}}$.*

It is interesting to note that only $\mathcal{F}_{2m}$ is a monotone under the most general set of free operations, i.e., $P_{\mathbb{N}}$, the issue with the other quantities being feed-forward.
The previous Theorem strongly constraints our ability to manipulate states in order to obtain metrological advantages when only certain operations are available. In this sense, non-classicality is a quantum resource for Quantum Metrology. The $\mathcal{F}_j$ can also be used to upper bound success rates of probabilistic processes, but, lacking some form of continuity or additivity, it cannot be exploited for studying approximate or asymptotic conversion rates.

4 On the relative entropies of optical non-classicality

4.1 Introduction to the Resource Theory

4.1.1 Motivations

As we already thoroughly discussed, one of the most compelling problems in infinite dimensional Quantum Resource Theories is the study of asymptotic conversion rates. As asymptotic continuity looses meaning in infinite dimension, we cannot apply anymore standard techniques such as Theorem 2.21. Then, we would like to exploit Theorem 2.23, which instead makes sense also on infinite dimensional spaces; unfortunately, it is very hard to build up a resource monotone fulfilling all the required properties.
Since lower semi-continuity and strong super-additivity appear to be crucial properties, it seems reasonable to look for a monotone in the following form:

$$\mathcal{F}(\rho) = \sup_{L\in\mathcal{A}(\mathcal{H})} \inf_{\sigma\in\mathcal{D}_f(\mathcal{H})} f(\rho,\sigma,L)\,, \tag{4.1}$$

for some suitable set $\mathcal{A}(\mathcal{H})$ of operators acting on $\mathcal{H}$ and with $\inf_{\sigma\in\mathcal{D}_f(\mathcal{H})} f(\rho,\sigma,L)$ being continuous in ρ and weakly super-additive in L. Indeed, it is easy to see that the pointwise sup of a continuous function is lower semi-continuous: if L_0 is the operator for which the sup is attained (within an error ϵ) for a state ρ, then:

$$\liminf_{n\to\inf} \sup_{L\in\mathcal{A}(\mathcal{H})} \inf_{\sigma\in\mathcal{D}_f(\mathcal{H})} f(\rho_n,\sigma,L) \geq \liminf_{n\to\inf} \inf_{\sigma\in\mathcal{D}_f(\mathcal{H})} f(\rho_n,\sigma,L_0) = \mathcal{F}(\rho) + \mathcal{O}(\epsilon)\,.$$

As for the super-additivity, we have:

$$\begin{aligned}
&\sup_{L^{AB}\in\mathcal{A}(\mathcal{H}^A\otimes\mathcal{H}^B)} \inf_{\sigma^{AB}\in\mathcal{D}_f(\mathcal{H}^A\otimes\mathcal{H}^B)} f(\rho,\sigma^{AB},L^{AB}) \\
&\qquad\geq \sup_{\substack{L^A\in\mathcal{A}(\mathcal{H}^A)\\ L^B\in\mathcal{A}(\mathcal{H}^B)}} \inf_{\sigma^{AB}\in\mathcal{D}_f(\mathcal{H}^A\otimes\mathcal{H}^B)} f(\rho,\sigma^{AB},L^A\otimes L^B) \\
&\qquad\geq \sup_{L^A\in\mathcal{A}(\mathcal{H}^A)} \inf_{\sigma^A\in\mathcal{D}_f(\mathcal{H}^A)} f(\rho,\sigma^A,L^A) \\
&\qquad\quad + \sup_{L^B\in\mathcal{A}(\mathcal{H}^B)} \inf_{\sigma^B\in\mathcal{D}_f(\mathcal{H}^B)} f(\rho,\sigma^B,L^B)\,.
\end{aligned}$$

However, resource monotones are usually defined in terms of minimisations rather than maximisations (see Proposition 2.2). One might try to solve the problem by considering a functional $\delta(\cdot\,,\cdot)$ which can be written in terms of a sup, and then exchange the sup and the inf. This is routinely done in finite dimensional Quantum Resource Theories thanks to the celebrated Sion's Theorem [101], which however requires some compactness property of the set $\mathcal{D}_f(\mathcal{H})$. Once again, the instrinsic infinite-dimensionality of the problem we are considering conspires against us: indeed, compactness is a very subtle concept in infinite dimensional spaces. But even if we succeed in proving compactness of $\mathcal{D}_f(\mathcal{H})$, the two nested optimizations seem to prevent any efficient computation of $\mathcal{F}(\rho)$. In this chapter, we will be able to solve both problems and prove, for the first time in an infinite-dimensional Quantum Resource Theory, a bound on asymptotic conversion rates. In our Quantum Resource Theory, the set of free operations is as general as possible, i.e., it contains all CPOs. Therefore, our bound remain valid in any Quantum Resource Theory of non-classicality. All the results reported in this chapter are also part of [102]. We mention that the idea of building super-additive monotones in the form (4.1) via Sion's Theorem was introduced in [103].

4.1.2 The monotones

An example of a functional that can be written in terms of a maximisation is D^M (see Lemmas 1.16 and 1.17). We start this section with the following definition.

Definition 4.1. *The* ***measured relative entropy of non-classicality*** *of a quantum state $\rho \in \mathcal{D}(\mathcal{H}_m)$ is defined as follows:*

$$N_r^M(\rho) \coloneqq \inf_{\sigma \in \mathcal{C}_m} D^M(\rho\|\sigma)\,. \tag{4.2}$$

For later convenience, we define also the following auxiliary quantity.

Definition 4.2. *The* ***relative entropy of non-classicality*** *of a quantum state $\rho \in \mathcal{D}(\mathcal{H}_m)$ is defined as follows:*

$$N_r(\rho) \coloneqq \inf_{\sigma \in \mathcal{C}_m} D(\rho\|\sigma)\,. \tag{4.3}$$

The idea of using the relative entropy of non-classicality as a resource monotone was already presented, though not worked out in detail, in [104]. Both the functional we just introduced are proper non-classicality monotones. Indeed, the following result holds true.

Proposition 4.3. *N_r and N_r^M are non-classicality monotones under arbitrary CPOs. Moreover, they possess the following properties:*

- *faithfulness, i.e., $N_r(\rho) = 0$, $N_r^M(\rho) = 0$ if and only if $\rho \in \mathcal{C}_m$;*
- *convexity.*
- *strong monotonicity.*

Proof. Monotonicity follows from Lemma 1.14 and monotonicity under CPTP of D [41], and Proposition 2.2; faithfulness descends from Proposition 2.13 and the fact that $\mathcal{C}_m$ is closed in trace norm by construction; joint convexity of D^M and D ensures the convexity of the monotones. Finally, if we allow for encoding the outcome of any measurement in an incoherent orthonormal basis of "classical flags" $\{|j^C\rangle\}_j$, strong monotonicity follows from standard arguments (see for example [105]). □

Thanks to Lemmas 1.16 and 1.17, we can rewrite N_r^M in terms of the following variational expression:

$$N_r^M(\rho) = \inf_{\sigma\in\mathcal{C}_m} \sup_{0<L\in\mathcal{B}_{sa}(\mathcal{H})} \{\operatorname{Tr}\rho\log L - \log\operatorname{Tr}\sigma L\} \tag{4.4}$$

$$= \inf_{\sigma\in\mathcal{C}_m} \sup_{0<L\in\mathcal{B}_{sa}(\mathcal{H})} \{\operatorname{Tr}\rho\log L + 1 - \operatorname{Tr}\sigma L\} \tag{4.5}$$

As we anticipated, our main goal is to put N_r^M in the form (4.1) by exchanging the sup with the inf. The problem is that in general, for a generic f, only the folowing expression holds true:

$$\sup_{L\in\mathcal{A}(\mathcal{H})} \inf_{\sigma\in\mathcal{D}_f(\mathcal{H})} f(\rho,\sigma,L) \le \inf_{\sigma\in\mathcal{D}_f(\mathcal{H})} \sup_{L\in\mathcal{A}(\mathcal{H})} f(\rho,\sigma,L)\,. \tag{4.6}$$

To obtain an equality one usually relies on the Sion's theorem, whose hypothesis are the following:

- $\mathcal{A}(\mathcal{H})$ is convex;
- $\mathcal{D}_f(\mathcal{H})$ is convex and compact;
- $f(\rho,\sigma,\cdot)$ is concave and upper semi-continuous in L $\forall\sigma$;
- $f(\rho,\cdot,L)$ is convex and lower semi-continuous in σ $\forall L$;

Note that these conditions are topology-dependent: we need to find at least one topology for which they are all satisfied, and we can exchange the sup with the inf. However, this is much harder than it looks, as we are going to show with the following example.
Consider the sequence of coherent states $\{|\alpha_n\rangle\!\langle\alpha_n|\}_n$, with $|\alpha_n| = n$. Obviously, for any quantum state $|\psi\rangle$ it holds $\lim_{n\to\infty} |\langle\alpha_n|\psi\rangle|^2 = 0$; hence, if such sequence has a limit, it must be the null operator $\mathbb{0}$ (at least with any "reasonable" topology, i.e., any topology which makes the inner product continuous). On the other hand, $\operatorname{Tr}[|\alpha_n\rangle\!\langle\alpha_n|\, I] = 1$ $\forall n$, so $\{|\alpha_n\rangle\!\langle\alpha_n|\}_n$ cannot converge to $\mathbb{0}$ in the weak topology. So, it simply does not have any limit with such a choice of a topology. The same argument holds for any of its subsequences, so $\mathcal{C}_m$ cannot be compact in the weak topology. On the other hand, the fact that $\lim_{n\to\infty} |\langle\alpha_n|\psi\rangle|^2 = 0$ implies that $|\alpha_n\rangle \xrightarrow[n\to\infty]{\text{WOT}} \mathbb{0}$, so $\{|\alpha_n\rangle\!\langle\alpha_n|\}_n$ is convergent in the weak operator topology. The problem is that now $f(\rho,\cdot,L)$ cannot be lower semi-continuous in σ, since the function

$$\sigma \mapsto -\operatorname{Tr}[\sigma L]$$

is instead upper semi-continuous in the weak operator topology for some L's. For example, if $L = I$, we have $\lim_{n\to\infty}(-\operatorname{Tr}[|\alpha_n\rangle\!\langle\alpha_n|\, I] = -1)$, but obviously $-\operatorname{Tr}[\mathbb{0}I] = 0$.
The problem cannot be solved by simply considering a "clever" topology which makes at the same time $\mathcal{C}_m$ compact and f upper semi-continuous in σ. Indeed, whichever topology we choose, we will always face the same problem: either $\mathcal{C}_m$ is compact, and hence $\{|\alpha_n\rangle\!\langle\alpha_n|\}_n$ (or a subsequence of its) converges to $\mathbb{0}$ and f is not lower semi-continuous in σ, or f is lower semi-continuous, and hence none of the subsequences of $\{|\alpha_n\rangle\!\langle\alpha_n|\}_n$ can be convergent, and $\mathcal{C}_m$ is not compact. The only solution is to somehow

change the expression itself of our monotone N_r^M. The next section makes a step towards this goal.

4.2 Main results

4.2.1 Restricted optimizations

The following two results ensure that we can restrict the optimizations over L and σ to smaller sets, whenever suitable hypothesis are fulfilled. The first one will play a crucial role in proving our main result, as we will see in the next section. However, it has a modest practical utility in applications, since the restricted sets are still too wide to be explicitly explored. On the other hand, the second result we will present will be used later on for simplifying computations of N_r^M in some specific cases.

Lemma 4.4. *Let us consider the variational expression for $N_r^M(\rho)$. Then:*

- *if $\operatorname{rk}\rho < \infty$ we can assume L to be a finite rank operator;*
- *if $S(\rho) < \infty$ we can assume $L \in \mathcal{T}_{sa}(\mathcal{H}_m)$.*

Proof. Obviously, any time we restrict the set of A the sup can only decrease. So, we have to prove that for any couple of states ρ and σ and for any $L \in \mathcal{B}_{sa}(\mathcal{H}_m)$, $L > 0$ we can find a sequence L_N in the restricted set such that

$$\lim_{N\to\infty} \{\operatorname{Tr}\rho\log L_N - \log\operatorname{Tr}\sigma L_N\} \geq \operatorname{Tr}\rho\log L - \log\operatorname{Tr}\sigma L\,. \tag{4.7}$$

We will denote the basis of eigenstates of ρ with $\{|\phi_j\rangle\}_j$, and the orthogonal projector onto the subspace $\operatorname{span}(\{|\phi_j\rangle\}_{j\leq N})$ with Π^N.

Given a finite rank state $\rho = \sum_{j=1}^M p_j |\phi_j\rangle\!\langle\phi_j|$, we define $L_N := \Pi^N(L)$ and then, for any $N \geq M$, it holds:

$$\begin{aligned}\operatorname{Tr}\rho\log L &= \operatorname{Tr}\Pi^N(\rho)\log L\\ &= \operatorname{Tr}\rho\Pi^N(\log L)\\ &\overset{1}{\leq} \operatorname{Tr}\rho\log\Pi^N(L)\\ &= \operatorname{Tr}\rho\log L_N\,,\end{aligned} \tag{4.8}$$

where in 1 we used Theorem 1.3. Moreover, $\operatorname{Tr}\sigma L_N \xrightarrow[N\to\infty]{} \operatorname{Tr}\sigma L$ for any fixed σ. This proves that we can assume $\operatorname{rk} L < \infty$, and hence also $\operatorname{Tr} L < \infty$. If we use expression (4.4), we can also exploit the scale invariance in L to require $\operatorname{Tr} L = 1$, which means that L is a finite rank state.

For a state $\rho = \sum_j p_j |\phi_j\rangle\!\langle\phi_j|$ with finite entropy we define instead $L_N := \Pi^N(L) + \omega_N$, with $\omega_N = (\mathbb{1} - \Pi^N)(\rho) = \rho - \Pi^N(\rho)$. Then

$$\begin{aligned}\operatorname{Tr}[\rho\log L] &= \operatorname{Tr}[(\Pi^N + \mathbb{1} - \Pi^N)(\rho)L]\\ &= \operatorname{Tr}[\Pi^N(\rho)\log L] + \operatorname{Tr}[\omega_N\log L]\\ &= \operatorname{Tr}[\Pi^N(\rho)\Pi^N(\log L)] + \operatorname{Tr}[\omega_N\log L]\\ &\overset{2}{\leq} \operatorname{Tr}[\Pi^N(\rho)\log\Pi^N(L)] + \operatorname{Tr}[\omega_N\log L]\\ &= \operatorname{Tr}[\rho\log L_N] - \operatorname{Tr}[\omega_N\log\omega_N] + \operatorname{Tr}[\omega_N\log L]\,,\end{aligned} \tag{4.9}$$

where 2 holds again because of Theorem 1.3. Moreover, if $S(\rho) < \infty$ and $|\operatorname{Tr}[\rho \log L]| < \infty$ (which can always be assumed since we are considering the sup over L), it holds:

$$\operatorname{Tr}[\omega_N \log \omega_N] \xrightarrow[N\to\infty]{} 0\,, \qquad \operatorname{Tr}[\omega_N \log L] \xrightarrow[N\to\infty]{} 0\,. \tag{4.10}$$

Hence, we can assume $L \in \mathcal{T}_{sa}(\mathcal{H}_m)$. Again, in (4.4), can assume $L \in \mathcal{D}(\mathcal{H}_m)$ (and also $S(L) < \infty$).

□

Remark 4.5. *Notice that $S(\rho) < \infty$ implies $E(\rho) < \infty$ as well. So, the previous result holds true for any physically meaningful state.*

Proposition 4.6. *If the quantum state $\rho \in \mathcal{D}(\mathcal{H}_m)$ is such that $\rho = \Lambda(\rho)$ under a free operation $\Lambda \in \mathcal{O}_f(\mathcal{H}_m \to \mathcal{H}_m)$, then:*

$$N_r^M(\rho) = \inf_{\sigma\in\Lambda(\mathcal{C}_m)} D^M(\rho||\sigma) = \inf_{\sigma\in\Lambda(\mathcal{C}_m)} \sup_{0<L\in\Lambda^\dagger(\mathcal{B}_{sa}(\mathcal{H}))} \{\operatorname{Tr}[\rho \log L] - \log \operatorname{Tr}[\sigma L]\}\,. \tag{4.11}$$

Proof. We have the following lower bound:

$$\begin{aligned} N_r^M(\rho) &= \inf_{\sigma\in\mathcal{C}_m} D^M(\rho||\sigma) \\ &\overset{1}{\geq} \inf_{\sigma\in\mathcal{C}_m} D^M(\Lambda(\rho)||\Lambda(\sigma)) \\ &= \inf_{\sigma\in\mathcal{C}_m} D^M(\rho||\Lambda(\sigma)) \\ &= \inf_{\sigma\in\Lambda(\mathcal{C}_m)} D^M(\rho||\sigma)\,, \end{aligned} \tag{4.12}$$

where 1 holds because of Lemma 1.14. Since by hypothesis $\Lambda(\mathcal{C}_m) \subseteq \mathcal{C}_m$, it also holds that:

$$N_r^M(\rho) \leq \inf_{\sigma\in\Lambda(\mathcal{C}_m)} D^M(\rho||\sigma)\,, \tag{4.13}$$

which proves the first equality.
Now, note that we can repeat the same procedure with $\Lambda \circ \Lambda$ in place of Λ. Moreover, $\Lambda^\dagger$ is a unital map, and then:

$$\begin{aligned} N_r^M(\rho) &= \inf_{\sigma\in\Lambda\circ\Lambda(\mathcal{C}_m)} \sup_{0<L\in\mathcal{B}_{sa}(\mathcal{H})} \{\operatorname{Tr}[\rho \log L] - \log \operatorname{Tr}[\sigma L]\} \\ &= \inf_{\sigma\in\Lambda(\mathcal{C}_m)} \sup_{0<L\in\mathcal{B}_{sa}(\mathcal{H})} \{\operatorname{Tr}[\Lambda(\rho) \log L] - \log \operatorname{Tr}[\Lambda(\sigma) L]\} \\ &= \inf_{\sigma\in\Lambda(\mathcal{C}_m)} \sup_{0<L\in\mathcal{B}_{sa}(\mathcal{H})} \left\{\operatorname{Tr}[\rho\Lambda^\dagger(\log L)] - \log \operatorname{Tr}[\sigma\Lambda^\dagger(L)]\right\} \\ &\overset{1}{=} \inf_{\sigma\in\Lambda(\mathcal{C}_m)} \sup_{0<L\in\Lambda^\dagger(\mathcal{B}_{sa}(\mathcal{H}))} \{\operatorname{Tr}[\rho \log L] - \log \operatorname{Tr}[\sigma L]\}\,, \end{aligned} \tag{4.14}$$

where 1 holds because of Theorem 1.3. □

Remark 4.7. *All the restrictions we proved so far holds also for the other expression for N_r^M, (4.5), since the usual argument for proving their equivalence still holds with the restricted sets.*

A notable example of a free map which the previous result can be applied to is Δ.

4.2.2 A bound on asymptotic conversion rates

In this section, we complete the program we presented at the beginning of the chapter, and prove the main result of this thesis. We start by proving the following technical result.

Lemma 4.8. *The cone*

$$\mathcal{C}_{m,+} \coloneqq \{\lambda\sigma : \lambda \geq 0,\, \sigma \in \mathcal{C}_m\} \subset \mathcal{T}_{sa}(\mathcal{H}_m) \tag{4.15}$$

generated by the set of classical states is closed with respect to the weak topology on* $\mathcal{T}_{sa}(\mathcal{H}_m)$. *Moreover, the set*

$$\widetilde{\mathcal{C}}_m \coloneqq \mathcal{C}_{m,+} \cap \{T \in \mathcal{T}_{sa}(\mathcal{H}_m) :\ \|T\|_1 \leq 1\} = \operatorname{conv}(\mathcal{C}_m \cup \{0\}) \tag{4.16}$$

of sub-normalised classical states is weak compact.*

Proof. Let us consider the single-mode case for simplicity. The generalization to m modes is straightforward. Let us define the following set of operators:

$$S \coloneqq \left\{ \sum_{\mu,\nu=1}^{n} \psi_\mu^* \psi_\nu \, e^{\frac{1}{2}|\alpha_\mu - \alpha_\nu|^2} \, \lambda^{\hat{a}^\dagger \hat{a}} \mathcal{D}(\alpha_\mu - \alpha_\nu) \lambda^{\hat{a}^\dagger \hat{a}} :\ n \in \mathbb{N},\, \boldsymbol{\psi} \in \mathbb{C}^n,\, \boldsymbol{\alpha} \in \mathbb{C}^n,\, \lambda \in [0,1) \right\}.$$

Note that $S \subset \mathcal{K}_{sa}(\mathcal{H})$, since every operator in the set is a finite linear combination of operators of the form $\lambda^{\hat{a}^\dagger \hat{a}} \mathcal{D}(\alpha_\mu - \alpha_\nu) \lambda^{\hat{a}^\dagger \hat{a}} = e^{\log \lambda \hat{a}^\dagger \hat{a}} \mathcal{D}(\alpha_\mu - \alpha_\nu) e^{\log \lambda \hat{a}^\dagger \hat{a}}$, which are clearly compact (in fact, even trace class) as long as $\lambda \in [0,1)$.
The **dual cone** of S is defined as follows:

$$S^* \coloneqq \{T \in \mathcal{T}_{sa}(\mathcal{H}) :\ \mathrm{Tr}[TK] \geq 0 \ \forall\ K \in \mathcal{S}\}\,.$$

The functional

$$\mathcal{T}_{sa}(\mathcal{H}) \ni T \mapsto \mathrm{Tr}[TK]$$

is weak* continuous $\forall\ K \in \mathcal{K}_{sa}(\mathcal{H})$, by definition of the weak* topology. Hence, any set

$$S_K^* = \{T \in \mathcal{T}_{sa}(\mathcal{H}) :\ \mathrm{Tr}[TK] \geq 0\}$$

is weak* closed in $\mathcal{T}_{sa}(\mathcal{H})$ for any compact operator K. In turns, S^* is closed because it is the intersection of closed sets.

It is elementary to see that $|\beta\rangle\!\langle\beta| \in \mathcal{S}^*$ for every $\beta \in \mathbb{C}^m$, because

$$
\begin{aligned}
&\langle\beta| \sum_{\mu,\nu=1}^{n} \psi_\mu^* \psi_\nu \, e^{\frac{1}{2}|\alpha_\mu-\alpha_\nu|^2} \, \lambda^{\hat{a}^\dagger\hat{a}} \mathcal{D}(\alpha_\mu-\alpha_\nu) \lambda^{\hat{a}^\dagger\hat{a}} |\beta\rangle \\
&\quad = \sum_{\mu,\nu=1}^{n} \psi_\mu^* \psi_\nu \, e^{\frac{1}{2}|\alpha_\mu-\alpha_\nu|^2} \, \langle\beta| \lambda^{\hat{a}^\dagger\hat{a}} \mathcal{D}(\alpha_\mu-\alpha_\nu) \lambda^{\hat{a}^\dagger\hat{a}} |\beta\rangle \\
&\quad \overset{1}{=} \sum_{\mu,\nu=1}^{n} \psi_\mu^* \psi_\nu \, e^{\frac{1}{2}|\alpha_\mu-\alpha_\nu|^2} e^{-(1-\lambda^2)|\beta|^2} \, \langle\lambda\beta| \mathcal{D}(\alpha_\mu-\alpha_\nu) |\lambda\beta\rangle \\
&\quad \overset{2}{=} e^{-(1-\lambda^2)|\beta|^2} \sum_{\mu,\nu=1}^{n} \psi_\mu^* \psi_\nu \, e^{\lambda((\alpha_\mu-\alpha_\nu)\beta^* - (\alpha_\mu-\alpha_\nu)^*\beta)} \\
&\quad = e^{-(1-\lambda^2)|\beta|^2} \sum_{\mu,\nu=1}^{n} \psi_\mu^* \, e^{\lambda(\alpha_\mu\beta^* - \alpha_\mu^*\beta)} \, \psi_\nu \, e^{\lambda(\alpha_\nu^*\beta - \alpha_\nu\beta^*)} \\
&\quad = e^{-(1-\lambda^2)|\beta|^2} \left| \sum\nolimits_{\mu=1}^{n} \psi_\mu^* \, e^{\lambda(\alpha_\mu\beta^* - \alpha_\mu^*\beta)} \right|^2 \\
&\quad \geq 0 \, .
\end{aligned}
$$

In 1 we used:

$$
\begin{aligned}
e^{\log\lambda \hat{a}^\dagger\hat{a}} \, |\beta\rangle &= e^{-\frac{|\beta|^2}{2}} \sum_{n=1}^{\infty} \frac{(\lambda\beta)^n}{\sqrt{n!}} \, |n\rangle \\
&= e^{-(1-\lambda^2)|\beta|^2} e^{-\frac{|\lambda\beta|^2}{2}} \sum_{n=1}^{\infty} \frac{(\lambda\beta)^n}{\sqrt{n!}} \, |n\rangle \\
&= e^{-(1-\lambda^2)|\beta|^2} \, |\lambda\beta\rangle \ ,
\end{aligned}
$$

while in 2 we used (1.19). Obviously, any convex combination of coherent states, i.e. any classical state, is in S^* as well. Therefore, we conclude that $\mathcal{C}_{m,+} \subseteq \mathcal{S}^*$.
Let us now prove the opposite inclusion. Pick $T \in \mathcal{T}_{sa}(\mathcal{H})$ such that $\mathrm{Tr}[TK] \geq 0$ for all $K \in S$; then

$$
\begin{aligned}
0 &\leq \liminf_{\lambda\to 1^-} \sum_{\mu,\nu=1}^{n} \psi_\mu^* \psi_\nu \, e^{\frac{1}{2}|\alpha_\mu-\alpha_\nu|^2} \ \mathrm{Tr}\left[T \, \lambda^{a^\dagger a} \mathcal{D}(\alpha_\mu-\alpha_\nu) \lambda^{\hat{a}^\dagger\hat{a}} \right] \\
&= \sum_{\mu,\nu=1}^{n} \psi_\mu^* \psi_\nu \, e^{\frac{1}{2}|\alpha_\mu-\alpha_\nu|^2} \ \mathrm{Tr}\left[T \, \mathcal{D}(\alpha_\mu-\alpha_\nu) \right] \\
&= \sum_{\mu,\nu=1}^{n} \psi_\mu^* \psi_\nu \, e^{\frac{1}{2}|\alpha_\mu-\alpha_\nu|^2} \, \chi_0^T(\alpha_\mu-\alpha_\nu) \\
&= \sum_{\mu,\nu=1}^{n} \psi_\mu^* \psi_\nu \, \chi_1^T(\alpha_\mu-\alpha_\nu)
\end{aligned}
$$

for all $\alpha \in \mathbb{C}^n$ and $\psi \in \mathbb{C}^n$, where χ_1^T is the characteristic function of T (being trace-class, it is proportional to a state). Hence, for all $\alpha \in \mathbb{C}^n$ the matrix $\left(\chi_1^T(\alpha_\mu-\alpha_\nu)\right)_{\mu,\nu=1,\dots,n}$ is positive semi-definite. Applying the classical Bochner theorem, we see that the function

$\chi_1^T(\alpha)$ is the Fourier transform of a positive measure, and hence T is proportional to a classical state.
We have then proved that $\mathcal{C}_{m,+} = S^*$, and hence that is weak* closed. Since the unit ball of $\mathcal{T}_{sa}(\mathcal{H})$ is weak* compact by the Banach-Alaoglu theorem [106], $\widetilde{\mathcal{C}}_m$ is the intersection of a weak* closed and a weak* compact set, and hence it is itself weak* compact. □

Now we have all the elements to exchange the order of sup and inf in N_r^M.
Theorem 4.9. *For any quantum state $\rho \in \mathcal{D}(\mathcal{H}_m)$ such that $S(\rho) < \infty$, the measured relative entropy of non-classicality can be rewritten as follows:*

$$\begin{aligned} N_r^M(\rho) &= \sup_{0<L\in\mathcal{B}_{\mathrm{sa}}(\mathcal{H}_m)} \left\{ \operatorname{Tr}[\rho \log L] - \log \sup_{\boldsymbol{\alpha}\in\mathbb{C}^m} \langle \boldsymbol{\alpha}|L|\boldsymbol{\alpha}\rangle \right\} \\ &= \sup_{0<L\in\mathcal{B}_{\mathrm{sa}}(\mathcal{H}_m)} \left\{ \operatorname{Tr}[\rho \log L] + 1 - \sup_{\boldsymbol{\alpha}\in\mathbb{C}^m} \langle \boldsymbol{\alpha}|L|\boldsymbol{\alpha}\rangle \right\}. \end{aligned} \tag{4.17}$$

Proof. First of all, let us rewrite

$$\begin{aligned} N_r^M(\rho) &= \inf_{\sigma\in\mathcal{C}_m} D^M(\rho\|\sigma) \\ &= \inf_{\sigma\in\mathcal{C}_m,\,\lambda\in[0,1]} \left\{ D^M(\rho\|\sigma) - \log\lambda \right\} \\ &= \inf_{\sigma\in\mathcal{C}_m,\,\lambda\in[0,1]} D^M(\rho\|\lambda\sigma) \\ &= \inf_{\sigma\in\widetilde{\mathcal{C}}_m} D^M(\rho\|\sigma)\,, \end{aligned}$$

where $\widetilde{\mathcal{C}}_m$ is defined in (4.16). With the help of Lemma 4.4, we continue by introducing the variational representation for the measured relative entropy:

$$\begin{aligned} N_r^M(\rho) &= \inf_{\sigma\in\widetilde{\mathcal{C}}_m} \sup_{0<L\in\mathcal{T}_{sa}(\mathcal{H}_m)} \{\operatorname{Tr}[\rho\log L] + 1 - \operatorname{Tr}[\sigma L]\} \\ &= \inf_{\sigma\in\widetilde{\mathcal{C}}_m} \sup_{0<L\in\mathcal{T}_{sa}(\mathcal{H}_m)} f(\rho,\sigma,L)\,, \end{aligned}$$

where

$$f(\rho,\sigma,L) \coloneqq \operatorname{Tr}[\rho\log L] + 1 - \operatorname{Tr}[\sigma L]\,.$$

Recall that Lemma 4.4 ensures that we can assume $L \in \mathcal{T}_{sa}(\mathcal{H})$, and thus obviously $L \in \mathcal{K}_{sa}(\mathcal{H})$. Now:

(i) $\widetilde{\mathcal{C}}_m$ is weak* compact by Lemma 4.8, and manifestly convex;

(ii) $\{L \in \mathcal{T}_{sa}(\mathcal{H}_m) : L > 0\}$ is convex;

(iii) $f(\rho,\cdot,L)$ is a convex (actually, linear) function on $\widetilde{\mathcal{C}}_m$ for every fixed compact operator $L > 0$; by definition of weak* topology it is also weak* continuous;

(iv) $f(\rho,\sigma,\cdot)$ is a concave function on $\{L \in \mathcal{T}_{sa}(\mathcal{H}_m) : L > 0\}$ for all $\sigma \in \widetilde{\mathcal{C}}_m$, because the operatorial logarithm is operator concave; it is also upper semi-continuous with respect to the trace norm topology since L is trace-class and hence we can rewrite:

$$\operatorname{Tr}[\rho\log L] = -D(\rho\|L) - S(\rho)\,,$$

with $D(\rho\|\cdot)$ being lower semi-continuous in L for any ρ.

Since all assumptions of Sion's minimax theorem are satisfied, we can exchange infimum and supremum, and write

$$\begin{aligned}
N_r^M(\rho) &\overset{1}{=} \sup_{0<L\in\mathcal{T}_{sa}(\mathcal{H}_m)} \inf_{\sigma\in\widetilde{\mathcal{C}}_m} f(\rho,\sigma,L) \\
&= \sup_{0<L\in\mathcal{T}_{sa}(\mathcal{H}_m)} \inf_{\sigma\in\widetilde{\mathcal{C}}_m} \{\mathrm{Tr}[\rho\log L] + 1 - \mathrm{Tr}[\sigma L]\} \\
&\overset{2}{=} \sup_{0<L\in\mathcal{T}_{sa}(\mathcal{H}_m)} \left\{\mathrm{Tr}[\rho\log L] + 1 - \sup_{\sigma\in\mathcal{C}_m} \mathrm{Tr}[\sigma L]\right\} \\
&\overset{3}{=} \sup_{0<L\in\mathcal{T}_{sa}(\mathcal{H}_m)} \left\{\mathrm{Tr}[\rho\log L] + 1 - \sup_{\boldsymbol{\alpha}\in\mathbb{C}^m} \langle\boldsymbol{\alpha}|L|\boldsymbol{\alpha}\rangle\right\} \\
&\overset{4}{=} \sup_{0<L\in\mathcal{B}_{sa}(\mathcal{H}_m)} \left\{\mathrm{Tr}[\rho\log L] + 1 - \sup_{\boldsymbol{\alpha}\in\mathbb{C}^m} \langle\boldsymbol{\alpha}|L|\boldsymbol{\alpha}\rangle\right\} .
\end{aligned}$$

Here, the identity in 1 is Sion's theorem; 2 comes from the fact that it is always convenient to choose $\mathrm{Tr}\,\sigma = 1$; 3 is due to the fact the extreme points of $\mathcal{C}_m$ are coherent states [85], and the sup of a convex function over a convex set is alway attained on an extreme point of the set; finally, in 4 we extended the sup to all $0 < L \in \mathcal{B}_{\mathrm{sa}}(\mathcal{H}_m)$, which can only increase the result, and used the fact that in general equation (4.6) holds true. The alternative expression for N_r^M can be obtain with the usual argument. □

Remark 4.10. *The restrictions introduced in Proposition 4.6 are still valid in the alternative expression of N_r^M: we can just repeat all the steps above with the restricted sets for σ and L, which are obviously convex.*

Remark 4.11. *N_r^M is lower semi-continuous in the trace norm topology, since we just wrote it as a* sup *of a manifestly trace norm continuous (in ρ) objective function.*

Thanks to Theorem 4.9, we can prove another fundamental property of N_r^M.

Proposition 4.12. *N_r^M is strongly super-additive.*

Proof. The proof is straightforward:

$$\begin{aligned}
&N_r^M(\rho^{AB}) \\
&= \sup_{L^{AB}\in\mathcal{B}_{sa}(\mathcal{H}_m^A\otimes\mathcal{H}_{m'}^B)} \left\{\mathrm{Tr}[\rho^{AB}\log L^{AB}] + 1 - \sup_{\boldsymbol{\alpha}^{\boldsymbol{AB}}\in\mathbb{C}^{m+m'}} \langle\boldsymbol{\alpha}^{AB}|L^{AB}|\boldsymbol{\alpha}^{AB}\rangle\right\} \\
&\geq \sup_{\substack{L^A\in\mathcal{B}_{sa}(\mathcal{H}_m^A)\\ L^B\in\mathcal{B}_{sa}(\mathcal{H}^B q_m)}} \left\{\mathrm{Tr}[\rho^{AB}\log(L^A\otimes L^B)] - \log \sup_{\substack{\boldsymbol{\alpha}^A\in\mathbb{C}^m\\ \boldsymbol{\alpha}^B\in\mathbb{C}^{m'}}} \langle\boldsymbol{\alpha}^A|L^A|\boldsymbol{\alpha}^A\rangle\,\langle\boldsymbol{\alpha}^B|L^B|\boldsymbol{\alpha}^B\rangle\right\} \\
&\geq \sup_{\substack{L^A\in\mathcal{B}_{sa}(\mathcal{H}_m^A)\\ L^B\in\mathcal{B}_{sa}(\mathcal{H}_m^B)}} \left\{\mathrm{Tr}[\rho^{AB}\log(L^A\otimes L^B)] - \log \sup_{\substack{\boldsymbol{\alpha}^A\in\mathbb{C}^m\\ \boldsymbol{\alpha}^B\in\mathbb{C}^{m'}}} \langle\boldsymbol{\alpha}^A|L^A|\boldsymbol{\alpha}^A\rangle\,\langle\boldsymbol{\alpha}^B|L^B|\boldsymbol{\alpha}^B\rangle\right\} \\
&\geq \sup_{L^A\in\mathcal{B}_{sa}(\mathcal{H}_m^A)} \left\{\mathrm{Tr}[\rho^{AB}(\log(L^A)\otimes I^B)] - \log \sup_{\boldsymbol{\alpha}^A\in\mathbb{C}^m} \langle\boldsymbol{\alpha}^A|L^A|\boldsymbol{\alpha}^A\rangle\right\} \\
&\quad + \sup_{L^B\in\mathcal{B}_{sa}(\mathcal{H}_{m'}^B)} \left\{\mathrm{Tr}[\rho^{AB}(I^A\otimes\log(L^B))] - \log \sup_{\boldsymbol{\alpha}^B\in\mathbb{C}^{m'}} \langle\boldsymbol{\alpha}^B|L^B|\boldsymbol{\alpha}^B\rangle\right\} .
\end{aligned}$$

Thus,

$$N_r^M(\rho^{AB}) \geq N_r^M(\rho^A) + N_r^M(\rho^B)\,,$$

with $\rho^A = \mathrm{Tr}_B\,\rho^{AB}$ and $\rho^B = \mathrm{Tr}_A\,\rho^{AB}$. □

Up to now we proved lower semi-continuity and strongly super-additivity of N_r^M. In order to apply Theorem 2.23, we need also weak additivity. Unfortunately, it is not clear whether N_r^M is weakly additive or not, so we have to consider its regularized version

$$N_r^{M,\infty}(\rho) := \lim_{n\to\infty} \frac{1}{n} N_r^M(\rho^{\otimes n})\,,$$

which inherits all the nice properties of N_r^M, but is also weakly addditive. Now we have all the ingredients to finally prove our main result: a bound on asymptotic conversion rates.

Theorem 4.13. *For any two finite entropy states $\rho \in \mathcal{D}(\mathcal{H}_m)$ and $\rho' \in \mathcal{D}(\mathcal{H}_{m'})$, the maximal asymptotic transformation rate $\widetilde{R}(\rho \to \rho')$ and the asymptotic transformation rate $R(\rho \to \rho')$ under classicality-preserving transformations are bounded by*

$$R(\rho \to \rho') \leq \widetilde{R}(\rho \to \rho') \leq \frac{N_r^{M,\infty}(\rho)}{N_r^{M,\infty}(\rho')} \leq \frac{N_r(\rho)}{N_r^M(\rho')}\,. \tag{4.18}$$

Proof. The bound with $N_r^{M,\infty}$ comes directly from Theorem 2.21. Now, strong super-additivity of N_r^M implies $N_r^M \leq N_r^{M,\infty}$, while Proposition 2.16 and $N_r \geq N_r^M$ together imply $N_r \geq N_r^{M,\infty}$. This complete the proof. □

Note that the bound with N_r and N_r^M, despite being aesthetically less appealing, is much more useful from a practical point of view than the other one. Indeed, N_r is written in terms of a inf, while N_r^M is written in terms of a sup: by plugging suitable ansatzes in their expressions, we can upper bound their ratio. So, the bound (4.18) can be, at least in principle, approximated up to any arbitrary precision.
It is noteworthy to mention that the same machinery can be replicated for other infinite dimensional Resource Theories. For example, in the case of Quantum Thermodynamics, the set of free states contains just one state, and then it is obviously compact. This allows us, for instance, to extend Theorem 2.24 to the infinite-dimensional case. As for the Resource Theory of infinite-dimensional coherence, we can just repeat all the steps we went through for non-classicality. On the contrary, this construction does not work for infinite-dimensional entanglement: the reason is that the extremal points of $\mathcal{S}(\mathcal{H})$ are not tensor product states, which is a crucial fact we used in Proposition 4.12.

4.2.3 Bounds on the monotones

In this section, we develop some techniques to upper bound N_r and N_r^M. In particular, they are both finite for finite energy states, as the following result shows. This implies that the bound (4.18) is never trivial for physically meaningful states.

Theorem 4.14. *Let $\rho \in \mathcal{D}(\mathcal{H}_m)$ be an m-mode state with irreducible photon number $N := N_0(\rho)$, where the r.h.s. is defined by* (1.24). *Then*

$$N_r(\rho) \leq m\, g(N/m)\,, \tag{4.19}$$

where $g(x) := (x+1)\log(x+1) - x\log x$.

Proof. We start by noticing that since N_r is invariant under displacement, we can assume w.l.o.g. that $\mathrm{Tr}[\rho\hat{a}_j] = 0$ for all $j \in \{1,\ldots,m\}$, so that $N(\rho) = N_0(\rho) = N$. Now, consider the $m\times m$ covariance matrix W defined by $W_{jk} := \mathrm{Tr}[\rho\,\hat{a}_j^\dagger\hat{a}_k]$. Since for an arbitrary $m\times m$ unitary matrix U we can induce the transformation $\hat{a}_j \mapsto \sum_k U_{jk}^*\hat{a}_k$ by passive symplectic unitaries, which leave the relative entropy of non-classicality invariant, we can w.l.o.g. assume that $W_{jj} = \mathrm{Tr}[\rho\,\hat{a}_j^\dagger\hat{a}_j] = N/m$. To see why, remember that given a Hermitian matrix X with spectrum $\lambda \in \mathbb{R}^m$ and a vector $\mu \prec \lambda$, where $\prec$ denotes majorisation, we can always find a unitary U such that $\mathrm{diag}(UXU^\dagger) = \mu$ (this is known as Schur–Horn theorem, see for example [107]). In our case, the vector of constant entries $\frac{\mathrm{Tr}\,W}{m}\,e = \frac{N}{m}\,e$, where $e := (1,\ldots,1)^\intercal \in \mathbb{R}^m$, is definitely majorised by the spectrum of W, which implies our claim. Therefore, hereafter we assume that $\mathrm{Tr}[\rho\,\hat{a}_j^\dagger\hat{a}_j] = N/m$. Now, let us consider a one-mode thermal state:

$$\tau_\nu := \frac{1}{\nu+1}\sum_{n=0}^{\infty}\left(\frac{\nu}{\nu+1}\right)^n |n\rangle\!\langle n| = \frac{1}{\nu+1}\left(\frac{\nu}{\nu+1}\right)^{\hat{a}^\dagger\hat{a}}. \tag{4.20}$$

For an m-tuple $\vec{\nu} = (\nu_1,\ldots,\nu_m) \in [0,\infty)^m$, set $\sigma(\vec{\nu}) := \tau_{\nu_1}\otimes\ldots\otimes\tau_{\nu_m}$. Then

$$\begin{aligned}N_r(\rho) &\leq \inf_{\nu_1,\ldots,\nu_m\geq 1} D\left(\rho\,\|\sigma(\vec{\nu})\right)\\ &= \inf_{\nu_1,\ldots,\nu_m\geq 1}\left\{-S(\rho) - \sum\nolimits_j\left(-\log(\nu_j+1) + \frac{N}{m}\log\left(\frac{\nu_j}{\nu_j+1}\right)\right)\right\}\\ &= -S(\rho) + m\,g\left(N/m\right),\end{aligned}$$

where we used the variational representation

$$g(x) = \inf_{\nu\geq 1}\left\{\log(1+\nu) - x\log\left(\frac{\nu}{\nu+1}\right)\right\},$$

whose proof is elementary. □

The next result gives another, independent, upper bound for the measured relative entropy of nonclassicality.

Theorem 4.15. *Let* $\rho \in \mathcal{D}(\mathcal{H}_m)$ *be a m-mode state with finite Wehrl entropy:*

$$S_W(\rho) = -\int d^{2m}\alpha\, Q^\rho(\boldsymbol{\alpha})\log Q^\rho(\boldsymbol{\alpha}) < \infty\,.$$

Then the measured relative entropy of non-classicality is finite and in particular

$$-\log||Q^\rho||_\infty \leq N_r^M(\rho) + S(\rho) + m\log\pi \leq S_W(\rho)\,. \tag{4.21}$$

Proof. Let us assume $m = 1$. Since ρ has finite Wehrl entropy by hypothesis, it has also finite von Neumann entropy. By virtue of Proposition 4.4, we can assume L to be a density operator and write

$$\begin{aligned}
N_r^M(\rho) &= \sup_{\omega\in\mathcal{D}(\mathcal{H})} \left\{ \operatorname{Tr}\rho\log\omega - \log\sup_\beta \langle\beta|\,\omega\,|\beta\rangle \right\} \\
&= \sup_{\omega\in\mathcal{D}(\mathcal{H})} \left\{ -S(\rho) - D(\rho\|\omega) - \log\sup_\beta \pi^m Q^\omega(\beta) \right\} \\
&\le \sup_{\omega\in\mathcal{D}(\mathcal{H})} \left\{ -D^M(\rho\|\omega) - \log\|Q^\omega\|_\infty \right\} - S(\rho) - m\log\pi \\
&\overset{1}{\le} \sup_{\omega\in\mathcal{D}(\mathcal{H})} \left\{ -D_{KL}(Q^\rho\|Q^\omega) - \log\|Q^\omega\|_\infty \right\} - S(\rho) - m\log\pi \\
&\le \sup_{\omega\in\mathcal{D}(\mathcal{H})} \left\{ S_W(\rho) + \int d^2\alpha\, Q^\rho(\alpha)\log\|Q^\omega\|_\infty - \log\|Q^\omega\|_\infty \right\} - S(\rho) - m\log\pi \\
&= S_W(\rho) - S(\rho) - m\log\pi\,.
\end{aligned}$$

Here 1 holds because in the definition of D^M we can choose a specific POVM, and in particular

$$\mathbb{1} = \frac{1}{\pi}\int d^2\alpha\, |\alpha\rangle\!\langle\alpha|\,,$$

which physically corresponds to a heterodyne detection [9].

At the same time we can just choose $\omega = \rho$:

$$N_r^M(\rho) \ge \operatorname{Tr}\rho\log\rho - \log||Q^\rho||_\infty - m\log\pi = -S(\rho) - m\log\pi - \log||Q^\rho||_\infty\,,$$

which proves the claim. □

The following result comes directly from the expression of the Q-function of Gaussian states, which are gaussian functions.

Corollary 4.16. *Let ρ be an arbitrary m-mode Gaussian state with quantum covariance matrix V. Then*

$$\frac{1}{2}\log\det(V+\mathbb{1}) \le N_r^M(\rho) + S(\rho) \le \frac{1}{2}\log\det(V+\mathbb{1}) + m\log e\,.$$

4.2.4 Approximation by truncation

In this section we prove that, even though N_r and N_r^M are not continuous in general, they are so on spectral truncations of any state.

Lemma 4.17. *Let $\rho,\sigma\in\mathcal{D}(\mathcal{H}_m)$ be two m-mode states, and set $\epsilon := \frac{1}{2}\|\rho-\sigma\|_1$. Assume that $E := N(|\rho-\sigma|) < \infty$, where $N(A) := \sum_j \operatorname{Tr}[A\,\hat a_j^\dagger \hat a_j]$ is the mean photon number of the operator $A \ge 0$. Then*

$$|N_r(\rho) - N_r(\sigma)| \le \epsilon\, m\, g\left(\frac{E}{\epsilon\, m}\right) + (1+\epsilon)H_2\left(\frac{\epsilon}{1+\epsilon}\right), \tag{4.22}$$

where $H_2(p) := -p\log p - (1-p)\log(1-p)$ is the binary entropy function.

Proof. The argument of [108, Lemma 7] carries over to our case. We find two states $\Delta, \Delta' \in \mathcal{D}(\mathcal{H}_m)$ such that

$$\rho - \sigma = \epsilon \left(\delta - \delta'\right), \quad |\rho - \sigma| = \epsilon \left(\delta + \delta'\right).$$

In particular, $\epsilon N(\delta) = N(\epsilon\delta) \leq N(|\rho - \sigma|) \leq E$. Let us construct the state

$$\omega := \frac{1}{1+\epsilon}\,\rho + \frac{\epsilon}{1+\epsilon}\,\delta' = \frac{1}{1+\epsilon}\,\sigma + \frac{\epsilon}{1+\epsilon}\,\delta\,.$$

Then, on the one hand

$$\begin{aligned} N_r(\omega) &\overset{1}{\leq} \frac{1}{1+\epsilon}\,N_r(\sigma) + \frac{\epsilon}{1+\epsilon}\,N_r(\delta) \\ &\overset{2}{\leq} \frac{1}{1+\epsilon}\,N_r(\sigma) + \frac{\epsilon}{1+\epsilon}\,m\,g\left(\frac{E}{\epsilon\,m}\right). \end{aligned}$$

Here, the estimate in 1 comes from convexity, while that in 2 is an application of Theorem 4.14. On the other hand, we can write

$$\begin{aligned} N_r(\omega) &= N_r\left(\frac{1}{1+\epsilon}\,\rho + \frac{\epsilon}{1+\epsilon}\,\delta'\right) \\ &\overset{3}{\geq} \frac{1}{1+\epsilon}\,N_r(\rho) + \frac{\epsilon}{1+\epsilon}\,N_r(\delta') - H_2\left(\frac{\epsilon}{1+\epsilon}\right) \\ &\geq \frac{1}{1+\epsilon}\,N_r(\rho) - H_2\left(\frac{\epsilon}{1+\epsilon}\right), \end{aligned} \tag{4.23}$$

where the inequality in 3 follows e.g. from [109, Proposition 5.24]. Putting all together we see that

$$N_r(\rho) - N_r(\sigma) \leq \epsilon\,m\,g\left(\frac{E}{\epsilon\,m}\right) + (1+\epsilon)H_2\left(\frac{\epsilon}{1+\epsilon}\right).$$

Together with the corresponding inequality with ρ and σ exchanged, this yields the claim. □

Remark 4.18. *Note that the same hold for N_r^M. We can use the same upper bound since $N_r^M \leq N_r$, and also the same lower bound because D_{KL} fulfills an inequality similar to that we used in (4.23). Therefore, the following Corollary hold for N_r^M as well.*

Corollary 4.19. *Let $\rho, \sigma \in \mathcal{D}(\mathcal{H}_m)$ be two m-mode commuting states with finite photon number $N(\rho), N(\sigma) \leq E$. Then*

$$|N_r(\rho) - N_r(\sigma)| \leq \epsilon\,m\,g\left(\frac{2E}{\epsilon\,m}\right) + (1+\epsilon)H_2\left(\frac{\epsilon}{1+\epsilon}\right). \tag{4.24}$$

In particular, denoting with $\rho = \sum_k p_k\,|e_k\rangle\!\langle e_k|$ the spectral decomposition of ρ, the sequence of spectral truncations $\rho_n := \left(\sum_{k\leq n} p_k\right)^{-1} \sum_{k\leq n} p_k\,|e_k\rangle\!\langle e_k|$ satisfies

$$N_r(\rho) = \lim_{n\to\infty} N_r(\rho_n)\,. \tag{4.25}$$

Proof. Thanks to Lemma 4.17, in order to prove (4.24) it suffices to show that $N(|\rho-\sigma|) \leq 2E$. Indeed, if $\rho = \sum_k p_k\,|e_k\rangle\!\langle e_k|$ and $\sigma = \sum_k q_k\,|e_k\rangle\!\langle e_k|$ then $|\rho - \sigma| = \sum_k |p_k - q_k|\,|e_k\rangle\!\langle e_k| \leq \sum_k (p_k + q_k)\,|e_k\rangle\!\langle e_k| = \rho + \sigma$, so that $N(|\rho - \sigma|) \leq N(\rho) + N(\sigma) \leq 2E$.

To deduce (4.25), note that $[\rho, \rho_n] = 0$, with $\epsilon_n := \frac{1}{2}\|\rho - \rho_n\|_1 \xrightarrow[n\to\infty]{} 0$. Also, eventually $N(\rho_n) \leq 2N(\rho) =: 2N$, so that

$$|N_r(\rho) - N_r(\rho_n)| \leq \epsilon_n m\, g\left(\frac{4N}{\epsilon_n m}\right) + (1+\epsilon_n) H_2\left(\frac{\epsilon_n}{1+\epsilon_n}\right) \xrightarrow[n\to\infty]{} 0\,,$$

where we used the well-known fact that $\lim_{\epsilon\to 0^+} \epsilon\, g(\delta/\epsilon) = 0$ for all $\delta > 0$. □

4.3 Applications

4.3.1 Relation with other Resource Theories

We want to briefly consider the relation between our Resource Theory of non-classicality and those presented in section 3.2. To begin with, N_r (and hence also N_r^M) gives a lower bound for $\log(1 + \mathcal{R}_G)$ [43, 92], and hence also $\log(1 + \mathcal{R}_A)$, with $\mathcal{R}_A$ being equal to the P-negativity defined in section 3.2.1. So, our monotones give a lower bound for $\mathcal{N}$. However, as we already pointed out, $\mathcal{N}$ is often divergent, so it cannot be used as a sensible upper bound for our monotones. On the contrary, our monotones are always finite for finite energy states.
If we choose $E = E_r$ (the relative entropy of entanglement) in equation 3.9, N_r (for a single mode state) gives instead an upper bound for the entanglement potential EP. Indeed:

$$\begin{aligned} N_r(\rho) &= \inf_{\sigma\in\mathcal{C}_1} D(\rho\|\sigma) \\ &\overset{1}{=} \inf_{\sigma\in\mathcal{C}_1} D(\rho\otimes|0\rangle\!\langle 0|\,\|\sigma\otimes|0\rangle\!\langle 0|) \\ &\overset{2}{\geq} \inf_{\sigma\in\mathcal{C}_1} D(U_{BS}(\rho\otimes|0\rangle\!\langle 0|)U_{BS}^\dagger \| U_{BS}(\sigma\otimes|0\rangle\!\langle 0|)U_{bs}^\dagger) \\ &\overset{3}{\geq} \inf_{\sigma'\in\mathcal{S}(\mathcal{H}_2)} D(U_{BS}(\rho\otimes|0\rangle\!\langle 0|)U_{BS}^\dagger\|\sigma')\,. \end{aligned}$$

In 1 we used the weak additivity of D; in 2 we used its monotonicity under CPTP; in 3 we used the fact that classical states do not get entangled by linear optical operations, and hence $U_{BS}(\sigma\otimes|0\rangle\!\langle 0|)U_{bs}^\dagger \in \mathcal{S}(\mathcal{H})$.
Finally, unfortunately we know no explicit relations between our monotones and the metrological monotone defined in section 3.2.3.

4.3.2 Fock diagonal states

We start this section with two results which follow trivially from Propositions 4.6 and the fact that Δ is a CPO.
Corollary 4.20. *Let $\rho \in \mathcal{D}(\mathcal{H})$ be a Fock-diagonal quantum state. Then:*

$$N_r^M(\rho) = \inf_{\sigma\in\mathcal{C}_m^{FD}} \sup_{\substack{0<L\in\mathcal{B}(\mathcal{H}) \\ L \text{ Fock-diagonal}}} [\mathrm{Tr}[\rho\log L] - \log\mathrm{Tr}[\sigma L]]$$

Remark 4.21. *If ρ is Fock-diagonal, also σ can be taken to be so, and hence they commute. Then, $N_r(\rho) = N_r^M(\rho)$ for any Fock-diagonal state ρ.*

Now we specialize Proposition 4.20 to the case $\operatorname{rk}\rho < \infty$.

Proposition 4.22. *Let ρ be a single-mode finite rank Fock-diagonal state. Then:*

$$N_r^M(\rho) = \sup_{\widetilde{L}} \left\{ \operatorname{Tr}\rho \log \widetilde{L} - \log \sup_{\alpha\in[0,\sqrt{M}]} \langle\alpha|\, \widetilde{L}\, |\alpha\rangle \right\}, \tag{4.26}$$

where we used the shorthand notation $\widetilde{L}$ to denote hermitian, positive and Fock diagonal operators satisfying $\operatorname{supp}(L) = \operatorname{supp}(\rho)$*, and $M = \max\{n : \langle n|\,\rho\,|n\rangle \neq 0\}$.*

Proof. Let Π be the projector over $\operatorname{supp}(\rho)$. We have that

$$\begin{aligned}
N_r^M(\rho) &\overset{1}{=} \sup_{\substack{0<L\in\mathcal{B}_{sa}(\mathcal{H})\\ L\,\text{Fock-diagonal}}} \left\{ \operatorname{Tr}[\rho\log L] - \log\sup_{\alpha\in\mathbb{C}} \langle\alpha|L|\alpha\rangle \right\} \\
&= \sup_{\substack{0<L\in\mathcal{B}_{sa}(\mathcal{H})\\ L\,\text{Fock-diagonal}}} \left\{ \operatorname{Tr}[\Pi(\rho)\log L] - \log\sup_{\alpha\in\mathbb{C}} \langle\alpha|L|\alpha\rangle \right\} \\
&\overset{2}{=} \sup_{\substack{0<L\in\mathcal{B}_{sa}(\mathcal{H})\\ L\,\text{Fock-diagonal}}} \left\{ \operatorname{Tr}[\rho\log \Pi(L)] - \log\sup_{\alpha\in\mathbb{C}} \langle\alpha|L|\alpha\rangle \right\} \\
&\overset{3}{=} \sup_{\widetilde{L}} \left\{ \operatorname{Tr}[\rho\log \widetilde{L}] - \log\sup_{\alpha\in\mathbb{C}} \langle\alpha|\widetilde{L}|\alpha\rangle \right\} \\
&\overset{4}{=} \sup_{\widetilde{L}} \left\{ \operatorname{Tr}[\rho\log \widetilde{L}] - \log\sup_{\alpha\in[0,M]} \langle\alpha|\widetilde{L}|\alpha\rangle \right\}
\end{aligned}$$

Here:

- 1 is a combination of Theorem 4.9, Corollary 4.20 and Remark 4.10;
- 2 comes as usual from Theorem 1.3;
- 3 holds because $\Pi(L)$ is always in the same form as $\widetilde{L}$, and obviously $\langle\alpha|\Pi(L)|\alpha\rangle \leq \langle\alpha|L|\alpha\rangle$, so it is always convenient to take $\Pi(L)$ in pace of L;
- 4 holds because for $|\alpha|^2 > M$ the function

  $$\operatorname{Tr}\Delta(|\alpha\rangle\langle\alpha|)\tilde{L} = e^{-|\alpha|^2}\sum_{n=0}^{M} \frac{|\alpha|^{2n}\ell_n}{n!}, \tag{4.27}$$

  where the ℓ_n are the eigenvalues of $\widetilde{L}$, becomes monotonically decreasing in α.

□

Remark 4.23. *From Corollary 4.19 we know that*

$$N_r^M\!\left(\frac{\Pi^N(\rho)}{\operatorname{Tr}\Pi^N(\rho)}\right) \underset{N\to\infty}{\to} N_r^M(\rho) \tag{4.28}$$

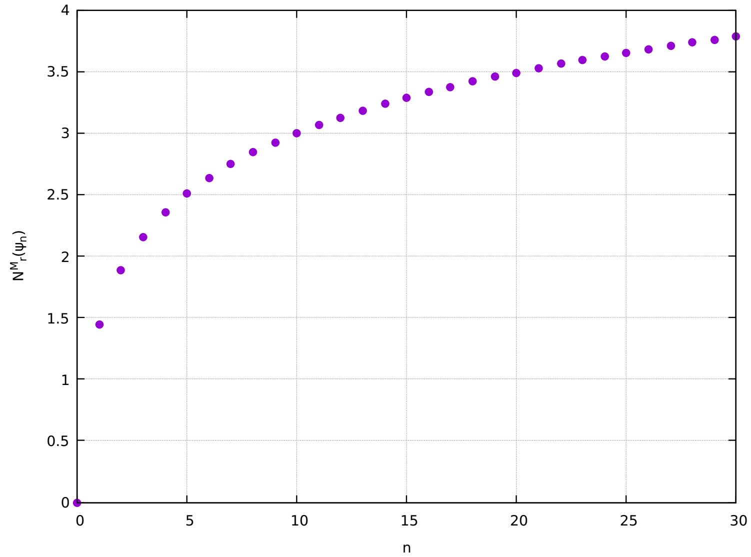


Figure 4.1: (Measured) relative entropy of Nonclassicality for a Fock state, for different values of n.

for any Fock diagonal state, where Π^N *is the orthogonal projector onto* $\mathrm{span}(\{|n\rangle\}_{n\leq N})$*. Therefore, in principle we can use Proposition 4.22 to approximate numerically* $N_r^M(\rho)$ *for any FD state* ρ *with arbitrary precision.*

The simplest example of Fock diagonal states is given by Fock states themselves. For a Fock state $\psi_n = |n\rangle\langle n|$ we can take $L = |n\rangle\langle n|$, and thus:

$$N_r^M(\psi_n) = -\log \sup_{\alpha\in\mathbb{C}} |\langle n|\alpha\rangle|^2 = \log\left(\frac{n!e^n}{n^n}\right) \underset{n\gg 1}{\approx} \frac{1}{2}\log 2\pi n\,.$$

The result is plotted in figure 4.1.

Another example of Fock diagonal states is represented by Fock state mixed with classical Fock diagonal noise. For instance, a realistic case is represented by thermal noise. We will call such states **noisy Fock states**:

$$\rho_{n,\nu}(p) = p\,|n\rangle\langle n| + (1-p)\tau_\nu \tag{4.29}$$

In principle, we can approximate the exact value of $N_r^M(\rho_{n,\nu}(p))$ with arbitrary precision for any n and ν, as pointed out in Remark 4.23. Let us first consider the simpler case $T\to 0$, which is a good approximation in certain regimes, for instance optical frequencies at room temperature. The state then becomes $\rho_{n,0}(p) = p\,|n\rangle\langle n| + (1-p)\,|0\rangle\langle 0|$, and thanks to Proposition 4.22 we can assume L to be in the form $L = \ell\,|n\rangle\langle n| + |0\rangle\langle 0|$ (we already exploited the scale invariance). Now we have only to perform two nested optimizations over one real parameter each:

$$N_r^M(\rho_{n,0}(p)) = \sup_{\ell}\left\{p\log\ell - \log \max_{\beta\in[0,\sqrt{n}]}(e^{-\beta^2}(1+\ell\beta^{2n}/n!))\right\}. \tag{4.30}$$

For $n\leq 4$ the maximization can even be carried out analitically, since the inner maximisation reduces to solving a n-th order algebraic equation. For example, for $n=1$ one simply finds $\beta = \sqrt{p}$, $\ell = 1/(1-p)$ and $N_r^M(\rho_{1,0}(p)) = p + (1-p)\log(1-p)$. For a finite temperature one has to consider truncations of ρ and perform numerical optimizations until some tolerance threshold is achieved. The results for different values of ν and n are reported in Figures 4.2 and 4.3. Note that from the plots one can observe that the content of non-classicality never reaches 0 for a (nontrivial) noisy state. This is because Fock diagonal states have a unbounded χ_1, and hence infinite absolute robustness of non-classicality. Hence, their non-classicality content cannot be destroyed by classical noise. By means of noisy Fock states, we are also able to construct an explicit example of a non-trivial protocol which (asymptotically) saturates our bound 4.18. Let us consider the state $\rho_{n,0}(p)$ for $0 < p < 1$ and the following protocol, implemented with only linear optics, destructive measurements and feed forward:

i) we send $\rho_{n,0}(p)$ through a beam splitter with transmittivity $t = \epsilon \ll \frac{1}{n}$ and a vacuum ancilla $|0\rangle\langle 0|$ at the other port, obtaining the following entangled state [110]:

$$\begin{aligned} \rho^{out} =& (1-p)\,|00\rangle\langle 00| \\ &+ p\sum_{m,m'=0}^{n} (\epsilon-1)^{2n-m-m'}\epsilon^{m+m'}\sqrt{\binom{n}{m}\binom{n}{m'}}\,|m, n-m\rangle\langle m', n-m'|\ ; \end{aligned} \tag{4.31}$$

ii) we perform photon counting on the ancillary mode: we measure 0 photons with probability $1 - p + p(1-\epsilon) = 1 - p\epsilon$, 1 photon with probability $p(1-\epsilon)^{n-1}\epsilon$ and all the other outcomes with probabilities $\mathcal{O}(\epsilon^2)$ (the condition $\epsilon \ll \frac{1}{n}$ ensure that we can neglect the possibility of detecting more than 1 photon despite the growing combinatorial factor in (4.31));

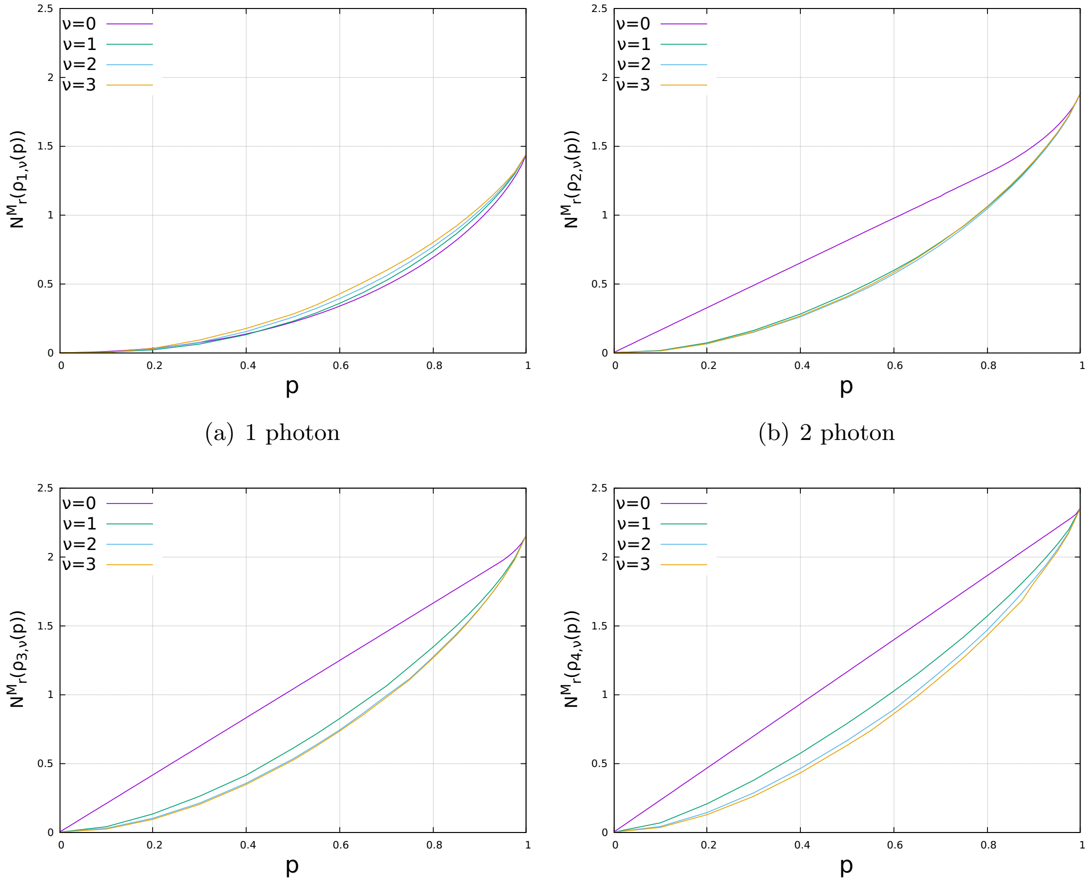
iii) If we measure 0 photons, we are left once again with the initial state $\rho_{n,0}(p)$ in the principal mode, and we re-send the state through the beam splitter and repeat the procedure; if we measure 1 photon, we are left with $\rho^{fin} = \psi_{n-1}$ in the principal mode

iv) with sufficiently many iterations we obtain with probability (almost) p the state ψ_{n-1}.

By convexity of the monotone we also know that $N_r^M(\rho_{n,0}(p)) \leq pN_r^M(\psi_n)$, and hence $R(\rho_{n,0}(p) \to \psi_{n-1}) \leq p\frac{N_r^M(\psi_n)}{N_r^M(\psi_{n-1})}$. But we just proved that the optimal transformation rate is $R(\rho_{n,0}(p) \to \psi_{n-1}) \geq p$, which implies that $N_r^M(\rho_{n,0}(p) \geq pN_r^M(\psi_{n-1})$. By combining the results we find that, for large n, we have $R(\rho_{n,0}(p) \to \psi_{n-1}) \approx p$, which is saturated by the protocol we presented. Moreover, we also proved that for large n we have $N_r^M(\rho_{n,0}(p)) \approx p\log n$, which is confirmed by the numerical analysis reported in Figure 4.3.

4.3.3 Cat states

In this section, we will consider applications of our Resource Theories involving cat states $\psi_\alpha = |\psi_\alpha\rangle\langle\psi_\alpha|$. To begin with, notice that the parameter α can always be taken



(a) 1 photon (b) 2 photon

(c) 3 photon (d) 4 photon

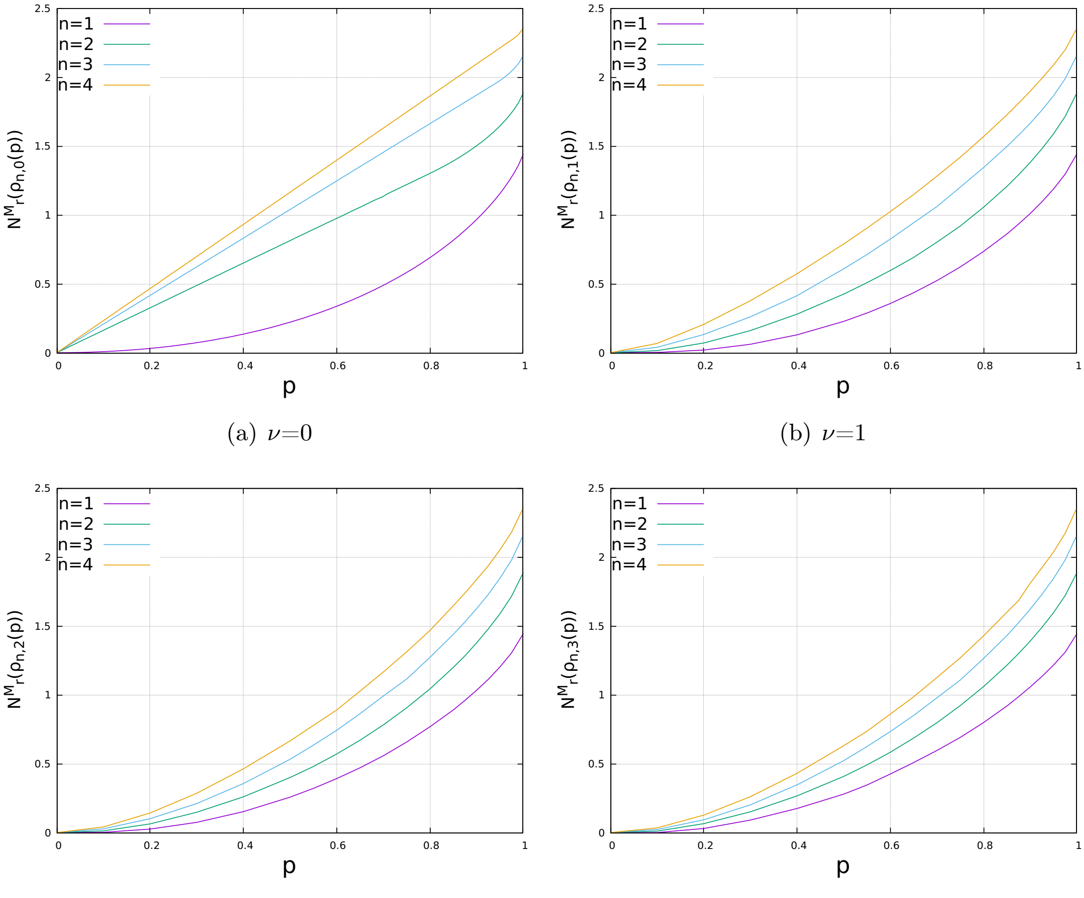
Figure 4.2: Nonclassicality for noisy Fock states: varying ν at fixed n.

to be real, since rotations in phase space are implemented by phase shifters, which let N_r^M invariant. From now on, we will always assume $\alpha \in \mathbb{R}$. Any cat state with $\alpha \in \mathbb{R}$ is invariant under reflections in phase space with respect to the real or imaginary axis. The totally symmetrizing map with respect to these symmetries, which we will denote with Λ^S, is obviously a projection and a free operation, since it sends coherent states into manifestly classical states.
From Proposition 4.6 and Remark 4.10 we see that

$$\begin{aligned} N_r^M(\psi_\alpha) &= \inf_{\sigma \in \Lambda^S(\mathcal{C})} \sup_{\substack{0 < L \in \Lambda^S(\mathcal{B}_{sa}(\mathcal{H})) \\ \mathrm{rk}\, L < \infty}} \{ \langle \psi_\alpha | \log L | \psi_\alpha \rangle - \log \mathrm{Tr}\, \sigma L \} \\ &= \sup_{\substack{0 < L \in \Lambda^S(\mathcal{B}_{sa}(\mathcal{H})) \\ \mathrm{rk}\, L < \infty}} \left\{ \langle \psi_\alpha | \log L | \psi_\alpha \rangle - \log \sup_{\beta \in \mathbb{C}} \langle \beta | L | \beta \rangle \right\} . \end{aligned} \tag{4.32}$$

An upper bound can be found by plugging an ansatz for σ in the first line of 4.32. The simplest choice which makes the result finite is

$$\sigma = \frac{1}{2} (|\alpha\rangle\langle\alpha| + |-\alpha\rangle\langle -\alpha|) \tag{4.33}$$



(a) ν=0 (b) ν=1 (c) ν=2 (d) ν=3

Figure 4.3: Nonclassicality for noisy Fock states: varying n at fixed ν.

In this case we get

$$
\begin{aligned}
N_r^M(\psi_\alpha) &\leq \sup_{\substack{0<L\in\Lambda^G(\mathcal{B}_{sa}(\mathcal{H}))\\ \mathrm{rk}\,L<\infty}} \left\{ \langle\psi_\alpha|\log L\,|\psi_\alpha\rangle - \log\langle\alpha|\,L\,|\alpha\rangle \right\} \\
&\overset{1}{=} \sup_{\substack{0<L\in\Lambda^G(\mathcal{B}_{sa}(\mathcal{H}))\\ L\in\mathrm{span}(|\alpha\rangle,|-\alpha\rangle)}} \left\{ \langle\psi_\alpha|\log L\,|\psi_\alpha\rangle - \log\langle\alpha|\,L\,|\alpha\rangle \right\} .
\end{aligned}
$$

Here 1 holds once again because of Theorem 1.3 (we projected onto $\mathrm{span}(|\alpha\rangle, |-\alpha\rangle)$). Notice that thanks to the scale invariance and the phase space symmetry, the operator L can be parametrized by a single real parameter:

$$
L = |\alpha\rangle\!\langle\alpha| + |-\alpha\rangle\!\langle-\alpha| + \ell(|\alpha\rangle\!\langle-\alpha| + |-\alpha\rangle\!\langle\alpha|),\; -1 \leq \ell \leq 1\,. \tag{4.34}
$$

In principle one can improve this esteem by considering more general classical states in the form

$$
\sigma = \frac{p_0}{2}(|\alpha\rangle\!\langle\alpha| + |-\alpha\rangle\!\langle-\alpha|) + \sum_{k=1}^{N} \frac{p_k}{2}(|\beta_k\rangle\!\langle\beta_k| + |-\beta_k\rangle\!\langle-\beta_k|)\,, \tag{4.35}
$$

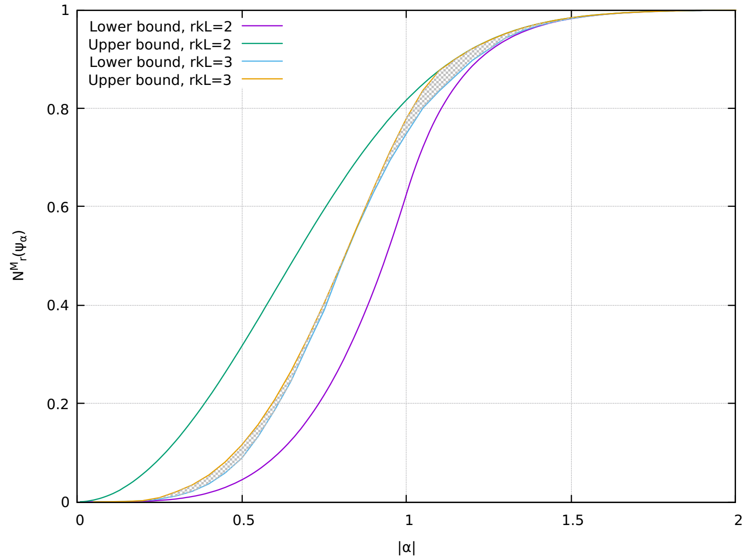


Figure 4.4: Bounds for the measured relative entropy of non-classicality for a single mode cat state. The shaded region represents the narrowest bound obtained.

and repeating the same steps with $L \in \text{span}(|\alpha\rangle, |-\alpha\rangle, \{|\beta_k\rangle\}_k)$. In Figure 4.4 we reported two different upper bounds, obtained with σ as in (4.33) and

$$\sigma = \frac{q}{2}(|\alpha\rangle\!\langle\alpha| + |-\alpha\rangle\!\langle-\alpha|) + (1-q)\,|0\rangle\!\langle 0| \,, \tag{4.36}$$

which allow to assume L to have rank 2 and 3 respectively.

A lower bound for $N_r^M(\psi_\alpha)$ can be easily computed by setting a maximum rank for L in the second line of (4.32) and then optimising numerically. In particular, for $\text{rk}\, L = 2$, we get

$$N_r^M(\psi_\alpha) \geq \sup_L \left\{ \langle\psi_\alpha| \log L \,|\psi_\alpha\rangle - \log \sup_{\beta\in\mathbb{C}} \langle\beta|\, L\, |\beta\rangle \right\},$$

with L being in the form (4.34). For $\text{rk}\, L = 3$, in order to preserve the symmetry, we have necessarily $L \in \text{span}(|\alpha\rangle, |-\alpha\rangle, |0\rangle)$. In Figure 4.4 we reported two different lower bounds, obtained with $\text{rk}\, L = 2$ and $\text{rk}\, L = 3$.

The ansatz (4.36) can be used to upper bound $N_r(\psi_\alpha)$ as well, by just plugging it into $D(\psi_\alpha\|\cdot)$. Then, the result, together with the lower bounds for $N_r^M(\psi_\alpha)$, can be used for upper bounding the rate of the following state conversion (concentration of cat states):

$$|\psi_\alpha\rangle \mapsto |\psi_{\sqrt{2}\alpha}\rangle \,.$$

The result, given in terms of the efficiency r of the conversion, is compared with the rate obtained with a protocol proposed in [111], and reported in Figure 4.5. Our bound is much bigger than the efficiency of the protocol, but this is not surprising: the protocol represents a non-asymptotic and exact transformation, making use of only ancillary coherent states, beam splitters and photodetection. The framework we considered is just much more general.

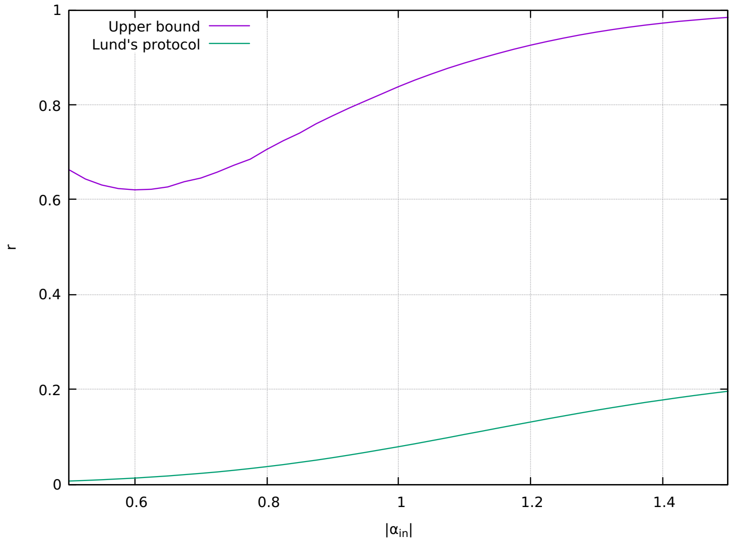


Figure 4.5: Efficiency for the concentration of cat states.

Conclusions

We introduced a new Resource Theory of optical non-classicality based on the measured relative entropy of non-classicality. Our resource monotone displays a number of interesting properties, which are usually hard to obtain in infinite-dimensional Quantum Resource Theories, or even in some finite-dimensional ones: lower semi-contintuity, strong super-additivity and finiteness on finite energy quantum states. Thanks to this properties, we were able to prove an upper bound on asymptotic conversion rates, which is never trivial for physically meaningful states. As we already pointed out, this is the first similar result in an infinite-dimensional setting. We studied several applications of our Resource Theory, and we proved a number of additional results which help in the computation or estimation of the resource content of states: in particular, we presented a method for numerically computing, up to any arbitrary precision, the resource content of Fock diagonal states. We also obtained universal upper and lower bounds for our resource monotone, differing by an additive constant in the case of gaussian states. Finally, we applied our results to several experimentally relevant scenarios, such as the purification of Fock states and the concentration of cat states. Once again, we stress the fact that many ideas developed in this thesis can be helpful in order to study other infinite-dimensional Quantum resource Theories, such as Entanglement Theory and Quantum Thermodynamics in infinite dimension. These should be considered as future directions of research.

Appendices

A | Some topology

Consider a possibly infinite-dimensional Hilbert space $\mathcal{H}$. Several fundamental classes of operators on $\mathcal{H}$ are defined as follows.

- $\mathcal{T}_{sa}(\mathcal{H})$: the space of **trace class** (self-adjoint) operators on $\mathcal{H}$, i.e., operators for which a base-independent trace can be defined.
- $\mathcal{T}^{+}_{(sa)}$: the cone of positive semidefinite (self-adjoint) trace class operators on $\mathcal{H}$;
- $\mathcal{D}(\mathcal{H})$: the set of density operators on $\mathcal{H}$;
- $\mathcal{K}_{(sa)}$: the Banach space of **compact** (self-adjoint) operators on $\mathcal{H}$, i.e., operators whose singular values approach 0;
- $\mathcal{K}^{+}_{(sa)}$: the cone of positive semidefinite (self-adjoint) compact operators on $\mathcal{H}$;
- $\mathcal{B}_{(\mathrm{sa})}$: the Banach space of bounded (self-adjoint) operators on $\mathcal{H}$, equipped with the operator norm.

Clearly, one has that $\mathcal{T}(\mathcal{H}) \subseteq \mathcal{K}(\mathcal{H}) \subseteq \mathcal{B}(\mathcal{H})$, with equality if and only if $\mathcal{H}$ is finite-dimensional. Also, the duality relation $\mathcal{T}(\mathcal{H})^* = \mathcal{B}(\mathcal{H})$ holds at the level of Banach spaces.
Two norms that can be defined on a Hilbert space are:

- $\|\cdot\|_\infty$: the **operator norm**, defined as $\|A\|_\infty \coloneqq \sup_{|\psi\rangle\in\mathcal{H}} \langle\psi|A|\psi\rangle$;
- $\|\cdot\|_1$: the **trace norm**, defined as: $\|A\|_1 = \mathrm{Tr}\sqrt{A^\dagger A}$.

There are several topologies that one can define on these spaces [112]. Here is a quick guide.

- The **weak operator topology** (**WOT**) on $\mathcal{B}(\mathcal{H})$ (and hence on respective subspaces) is the coarsest topology that makes all functionals $A \mapsto \langle\psi|A|\psi\rangle$ continuous, for all $|\psi\rangle \in \mathcal{H}$.
- The **weak* topology** (**W*T**) on $\mathcal{T}(\mathcal{H})$ is the coarsest topology that makes all functionals $T \mapsto \mathrm{Tr}[TK]$ continuous, for all $K \in \mathcal{K}(\mathcal{H})$.
- The **weak topology** (**WT**) on $\mathcal{T}(\mathcal{H})$ is the coarsest topology that makes all functionals $T \mapsto \mathrm{Tr}[TA]$ continuous, for all $A \in \mathcal{B}(\mathcal{H})$.
- The **trace norm topology** (**TNT**) on $\mathcal{T}(\mathcal{H})$ is the one induced by the trace norm $\|\cdot\|_1$.

- The **operator norm topology** (**ONT**) on $\mathcal{B}(\mathcal{H})$ is the one induced by the operator norm $\|\cdot\|_\infty$.

When we write a limit, we specify the topology by writing the respective acronym above the arrow. For example, "$\xrightarrow[n\to\infty]{\text{TNT}}$" means convergence in trace norm for $n \to \infty$. A consequence of [113, Lemma 9] is the following.

Lemma A.1 ('SWOT' convergence lemma [114, Lemma 4.3])**.** *For a net $(\omega_\alpha)_\alpha \subseteq \mathcal{T}_{sa}^+$ of positive semidefinite trace class operators, if $\omega_\alpha \xrightarrow[\alpha]{\text{wot}} \omega \in \mathcal{T}_{sa}^+$ in the weak operator topology, and moreover $\operatorname{Tr}\omega_\alpha \xrightarrow[\alpha]{} \operatorname{Tr}\omega$, then $\omega_\alpha \xrightarrow[\alpha]{\text{n}} \omega$ in norm.*

Corollary A.2. *The weak topology and the norm topology coincide on $\mathcal{T}_{sa}^+$. They also coincide with the weak operator topology on $\mathcal{D}(\mathcal{H})$.*

The norm topology does not coincide with the weak operator topology on $\mathcal{T}_{sa}^+$. For instance, the sequence of Fock states $(|n\rangle\langle n|)_n$ converges to 0 in the weak operator topology, but it is clearly not convergent in the norm topology.

B | Some distribution Theory

A **distribution** [115] T is a linear functional defined on the space of **test functions**, i.e., smooth functions on $\mathbb{R}$ (the generalisation to more general domains is straightforward) with compact support: $T : C_c^\infty(\mathbb{R}) \to \mathbb{R}$. We want to introduce two fundamental concepts related to distributions.
The **weak sign** at point $x_0 \in \mathbb{R}$ of a distribution T is defined as follows.
Definition B.1. *T is weakly positive (negative) in x_0 if it exists an open neighbourhood $A_0 \subset \mathbb{R}$ of x_0 such that for any non-negative test function $\varphi_{A_0} \in \mathbb{C}_c^\infty(A_0)$ with support in A_0 it holds:*

$$T(\varphi_{A_0}) \geq 0 \, (\leq 0) \, .$$

In the following, for ease of terminology, we will drop the attribute "weak" when no ambiguities can arise. An example of a positive distribution is the usual δ function, while it suffices to put a minus in front of it in order to obtain a negative one. Note that, contrarily to standard real-valued functions, distributions does not have to be either positive or negative at each point: derivatives of the δ function are straightforward counterexamples.
Another fundamental property of distributions is the **order**.
Definition B.2. *The order of a distribution T is the smallest integer $m \in \mathbb{N}$ such that for any compact set $K \subset \mathbb{R}$ it exists a constant $C_K < \infty$ such that:*

$$|T(\varphi_K)| \leq C_K \sup_{x \in K} \sup_{\alpha \leq m} |\partial^\alpha \varphi(x)| \, ,$$

for any test function $\varphi_K \in C_c^\infty(K)$ with support in K.

Clearly, an example of a distribution of order m is the m-th derivative of the δ function. A distribution of order 0 is also called a **measure**. Note that the order can be $+\infty$ as well.
It is not a coincidence that, among all the derivatives of the δ function, only the δ function itself, which is a measure, has a definite sign. Indeed, the following result holds true.
Proposition B.3. *A distribution $T : C_c^\infty(\mathbb{R}) \to \mathbb{R}$ is a measure if and only if it can be decomposed as $T = T_+ - T_-$, with $T_\pm$ being positive distributions on $C_c^\infty(S_\pm)$, and $S_+ \cup S_- = \mathbb{R}$ and $S_+ \cap S_- = \emptyset$.*

Proof. We start start by assuming that such a decomposition is possible. Let us consider T_+ first, and let $K \subset \mathbb{R}$ be an arbitrary compact set. By hypothesis, for any $x \in K$ it exists an open set A_x such that $T_+(\varphi) \geq 0$ for any non-negative test function $\varphi \in C_c^\infty(A_x)$. Clearly, $K \subset \bigcup_{x \in K} A_x$; being K compact, by definition it exists a finite set of indices J

such that $K \subset \bigcup_{j\in J} A_{x_j} =: A$, with $x_j \in K$ for any $j \in J$ [112]. It is easy to construct a non-negative test function $f \in C_c^\infty(A)$ such that $f(x) = 1\,\forall\, x \in K$ and $f(x) = 0\,\forall\, x \notin A$. Now, Let $\varphi_K \in C_c^\infty(K)$ be an arbitrary test function. We define the two auxiliary test functions $g_\pm \in C_c^\infty(A)$ as follows:

$$g_\pm \coloneqq \|\varphi\|_\infty f \pm \varphi\,.$$

Clearly, both g_+ and g_- are non-negative. So, it holds that:

$$0 \leq T_+(g_\pm) = \|\varphi\|_\infty T_+(f) \pm T(\varphi)\,,$$

which implies that

$$|T_+(\varphi)| \leq C_K \|\varphi\|_\infty\,,$$

where we defined $C_K \coloneqq T_+(f)$, meaning that T_+ is indeed a measure. A similar procedure can be carried out for T_-. The fact that the sum of two measures is again a measure follows trivially from triangular inequality:

$$|T(\varphi)| \leq |T_+(\varphi)| + |T_-(\varphi)| \leq C_{K,+}\|\varphi\|_\infty + C_{K,-}\|\varphi\|_\infty \leq 2\max\{C_{K,+}, C_{K,-}\}\|\varphi\|_\infty\,.$$

The other direction of the proof is just the Hahn decomposition Theorem [116]. □

Note that the fact that a measure can always be decomposed in a positive and a negative part does not implies that it has a definite sign at any point, a counterexample being (3.5).

Bibliography

[1] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge: Cambridge University Press, 2010.

[2] J. Preskill. *Lecture Notes for Physics 219: Quantum Information and Computation*. 2015.

[3] M. M. Wilde. *Quantum Information Theory*. 2nd. Cambridge University Press, 2017.

[4] G. Mackey. *Mathematical Foundations of Quantum Mechanics*. Benjamin, 1963.

[5] F. Hansen and G. K. Pedersen. "Jensen's Operator Inequality". In: *Bull. Lond. Math. Soc.* 35.4 (2003), pp. 553–564.

[6] J. R. Klauder and E. C. G. Sudarshan. *Fundamentals of Quantum Optics*. Benjamin, NY, 1968.

[7] S. Barnett and P. M. Radmore. *Methods in Theoretical Quantum Optics*. Oxford Series in Optical and Imaging Sciences. Clarendon Press, 2002.

[8] W. P. Schleich. *Quantum optics in phase space*. John Wiley & Sons, 2011.

[9] A. Serafini. *Quantum Continuous Variables: A Primer of Theoretical Methods*. CRC Press, Taylor & Francis Group, 2017.

[10] Claude Cohen-Tannoudji, Bernanrd Diu, and Franck Laloë. *Quantum Mechanics (Vol. 1)*. Wiley-VCH, 1977.

[11] H. F. Baker. "Alternants and Continuous Groups". In: *P. Lond. Math Soc.* s2-3.1 (1905), pp. 24–47.

[12] E. Schrödinger. "Der stetige Übergang von der Mikro- zur Makromechanik". In: *Naturwissenschaften* 14.28 (1926), pp. 664–666.

[13] J. R. Klauder. "The action option and a Feynman quantization of spinor fields in terms of ordinary c-numbers". In: *Ann. Phys. (N. Y.)* 11.2 (1960), pp. 123–168.

[14] R. J. Glauber. "Coherent and Incoherent States of the Radiation Field". In: *Phys. Rev.* 131 (6 1963), pp. 2766–2788.

[15] E. C. G. Sudarshan. "Equivalence of Semiclassical and Quantum Mechanical Descriptions of Statistical Light Beams". In: *Phys. Rev. Lett.* 10 (7 1963), pp. 277–279.

[16] K. C. Tan et al. "Quantifying the Coherence between Coherent States". In: *Phys. Rev. Lett.* 119 (19 2017), p. 190405.

[17] B. Yadin et al. "Operational Resource Theory of Continuous-Variable Nonclassicality". In: *Phys. Rev. X* 8 (4 2018), p. 041038.

[18] S. L. Braunstein and P. van Loock. "Quantum information with continuous variables". In: *Rev. Mod. Phys.* 77 (2 2005), pp. 513–577.

[19] A. Ferraro, S. Olivares, and M. G. A. Paris. "Gaussian states in continuous variable quantum information". In: *Preprint arXiv:quant-ph/0503237* (2005).
[20] X.-B. Wang et al. "Quantum information with Gaussian states". In: *Phys. Rep.* 448.1 (2007), pp. 1–111.
[21] C. Weedbrook et al. "Gaussian quantum information". In: *Rev. Mod. Phys.* 84 (2 2012), pp. 621–669.
[22] M. H. Stone. "On One-Parameter Unitary Groups in Hilbert Space". In: *Ann. Math.* 33 (3 1932), pp. 643–648.
[23] C.-K. Li1, R. Roberts, and X. Yin. "Decomposition of unitary matrices and quantum gates". In: *Int. J. Quantum Inf.* 11 (1 2013).
[24] M. Reck et al. "Experimental Realization of Any Discrete Unitary Operator". In: *Phys. Rev. Lett.* 73.1 (1994), pp. 58–63.
[25] E. P. Wigner. *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra.* Academic Press, 1959.
[26] P. Kok et al. "Linear Optical Quantum Computing with Photonic Qubits". In: *Rev. Mod. Phys.* 79 (1 2007), pp. 135–174.
[27] M. G. A. Paris. "Displacement operator by beam splitter". In: *Phys. Lett. A* 217.2 (1996), pp. 78–80.
[28] G. B. Folland. *Harmonic Analysis in Phase Space.* Princeton University Press, 1989.
[29] E. Wigner. "On the Quantum Correction For Thermodynamic Equilibrium". In: *Phys. Rev.* 40 (5 1932), pp. 749–759.
[30] S. Kullback. *Information Theory and Statistics.* Dover Publications, 1959.
[31] S. Kullback and R. A. Leibler. "On Information and Sufficiency". In: *Ann. Math. Statist.* 22.1 (1951), pp. 79–86.
[32] I. Csiszár. "Information-Type Measures of Difference of Probability Distributions and Indirect Observations". In: *Studia Sci. Math. Hungarica* 2 (1967), pp. 299–318.
[33] S. Kullback. "A lower bound for discrimination information in terms of variation (Corresp.)" In: *IEEE Trans. Inf. Theory* 13.1 (1967), pp. 126–127.
[34] M. S. Pinsker. *Information and Information Stability of Random Variables and Processes.* Ed. by A. Feinstein. Holden-Day series in time series analysis. Holden-Day, 1964.
[35] M. J. Donald. "On the relative entropy". In: *Commun. Math. Phys.* 105.1 (1986), pp. 13–34.
[36] F. Hiai and D. Petz. "The proper formula for relative entropy and its asymptotics in quantum probability". In: *Comm. Math. Phys.* 143.1 (1991), pp. 99–114.
[37] M. Berta, O. Fawzi, and M. Tomamichel. "On variational expressions for quantum relative entropies". In: *Lett. Math. Phys.* 107.12 (2017), pp. 2239–2265.
[38] G. de Palma. "The Wehrl entropy has Gaussian optimizers". In: *Lett. Math. Phys.* 108.1 (2018), pp. 97–116.
[39] H. Umegaki. "Conditional expectation in an operator algebra. IV. Entropy and information". In: *Kodai Math. Sem. Rep.* 14.2 (1962), pp. 59–85.
[40] G. Lindblad. "Entropy, information and quantum measurements". In: *Commun. Math. Phys.* 33.4 (1973), pp. 305–322.
[41] A. Wehrl. "General properties of entropy". In: *Rev. Mod. Phys.* 50 (2 1978), pp. 221–260.

[42] D. Petz. “Sufficient subalgebras and the relative entropy of states of a von Neumann algebra”. In: *Commun. Math. Phys.* 105.1 (1986), pp. 123–131.
[43] E. Chitambar and G. Gour. “Quantum resource theories”. In: *Rev. Mod. Phys.* 91 (2 2019), p. 025001.
[44] C. H. Bennett et al. “Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels”. In: *Phys. Rev. Lett.* 70 (13 1993), pp. 1895–1899.
[45] C. H. Bennett and S. J. Wiesner. “Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states”. In: *Phys. Rev. Lett.* 69 (20 1992), pp. 2881–2884.
[46] S. L. Braunstein and H. J. Kimble. “Dense coding for continuous variables”. In: *Phys. Rev. A* 61 (4 2000), p. 042302.
[47] A. K. Ekert. “Quantum cryptography based on Bell’s theorem”. In: *Phys. Rev. Lett.* 67 (6 1991), pp. 661–663.
[48] C. H. Bennett. “Quantum cryptography using any two nonorthogonal states”. In: *Phys. Rev. Lett.* 68 (21 1992), pp. 3121–3124.
[49] C. H. Bennett. “Quantum cryptography: public key distribution and coin tossing”. In: *Proc. IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India.* 1984, pp. 175–179.
[50] T. C. Ralph. “Continuous variable quantum cryptography”. In: *Phys. Rev. A* 61 (1 1999), p. 010303.
[51] A. Einstein, B. Podolsky, and N. Rosen. “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?” In: *Phys. Rev.* 47 (10 1935), pp. 777–780.
[52] N. Bohr. “Can Quantum-Mechanical Description of Physical Reality be Considered Complete?” In: *Phys. Rev.* 48 (8 1935), pp. 696–702.
[53] N. Brunner et al. “Bell nonlocality”. In: *Rev. Mod. Phys.* 86.2 (2014), p. 419.
[54] J. S. Bell. “On the Einstein-Podolsky-Rosen paradox”. In: *Physics* 1.3 (1964), pp. 195–200.
[55] A. Aspect, J. Dalibard, and G. Roger. “Experimental Tests of Realistic Local Theories via Bell’s Theorem”. In: *Phys. Rev. Lett.* 47 (7 1981), pp. 460–463.
[56] A. Aspect, J. Dalibard, and G. Roger. “Experimental Test of Bell’s Inequalities Using Time-Varying Analyzers”. In: *Phys. Rev. Lett.* 49 (25 1982), pp. 1804–1807.
[57] R. F. Werner. “Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model”. In: *Phys. Rev. A* 40 (8 1989), pp. 4277–4281.
[58] R. Horodecki et al. “Quantum entanglement”. In: *Rev. Mod. Phys.* 81 (2 2009), pp. 865–942.
[59] M. B. Plenio and S. Virmani. “An Introduction to Entanglement Measures”. In: *Quantum Inf. Comput.* 7.1 (2007), pp. 1–51.
[60] V. Vedral and M. B. Plenio. “Entanglement Measures and Purification Procedures”. In: *Phys. Rev. A* 57 (3 1998), pp. 1619–1633.
[61] V. Vedral et al. “Quantifying Entanglement”. In: *Phys. Rev. Lett.* 78 (12 1997), pp. 2275–2279.
[62] J. Åberg. “Quantifying superposition”. In: *Preprint arXiv:quant-ph/0612146* (2006).
[63] T. Baumgratz, M. Cramer, and M. B. Plenio. “Quantifying Coherence”. In: *Phys. Rev. Lett.* 113 (14 2014), p. 140401.

[64] L. Lami. “Completing the Grand Tour of asymptotic quantum coherence manipulation”. In: *Preprint arXiv:1902.02427* (2019).
[65] A. Streltsov, G. Adesso, and M. B. Plenio. “Colloquium: Quantum coherence as a resource”. In: *Rev. Mod. Phys.* 89 (4 2017), p. 041003.
[66] T. Theurer et al. “Resource Theory of Superposition”. In: *Phys. Rev. Lett.* 119 (23 2017), p. 230401.
[67] A. Winter and D. Yang. “Operational Resource Theory of Coherence”. In: *Phys. Rev. Lett.* 116 (12 2016), p. 120404.
[68] J. Goold et al. “The role of quantum information in thermodynamics: a topical review”. In: *J. Phys. A* 49.14 (2016), p. 143001.
[69] F. G. S. L. Brandão et al. “The Resource Theory of Quantum States Out of Thermal Equilibrium”. In: *Phys. Rev. Lett.* 111.25 (2013).
[70] N. Yunger Halpern and D. T. Limmer. “Fundamental limitations on photoisomerization from thermodynamic resource theories”. In: *Preprint arXiv:1811.06551* (2018).
[71] C. Palazuelos. “Superactivation of Quantum Nonlocality”. In: *Phys. Rev. Lett.* 109.19 (2012).
[72] P. Faist, M. Berta, and F. Brandão. “Thermodynamic Capacity of Quantum Processes”. In: *Phys. Rev. Lett.* 122 (20 2019), p. 200601.
[73] S. Meznaric, S. R. Clarck, and Datta A. “Quantifying the Nonclassicality of Operations”. In: *Phys. Rev. Lett.* 110 (7 2013).
[74] T. Theurer et al. “Quantifying Operations with an Application to Coherence”. In: *Phys. Rev. Lett.* 122 (19 2019), p. 190405.
[75] B. Synak-Radtke and M. Horodecki. “On asymptotic continuity of functions of quantum states”. In: *J. Phys. A* 39.26 (2006), pp. L423–L437.
[76] A. Peres. “Separability criterion for density matrices”. In: *Phys. Rev. Lett.* 77 (8 1996), pp. 1413–1415.
[77] M. Horodecki, P. Horodecki, and R. Horodecki. “mixed states: necessary and sufficient conditions”. In: *Phys. Lett. A* 223.1–2 (1996), pp. 1–8.
[78] M. J. Donald and M. Horodecki. “Continuity of Relative Entropy of Entanglement”. In: *Phys. Lett. A* 264.4 (1999), pp. 257–260.
[79] M. Horodecki, P. Horodecki, and R. Horodecki. “Mixed-State Entanglement and Distillation: Is there a “Bound” Entanglement in Nature?” In: *Phys. Rev. Lett.* 80 (24 1998), pp. 5239–5242.
[80] M. B. Plenio. “Logarithmic Negativity: A Full Entanglement Monotone That is not Convex”. In: *Phys. Rev. Lett.* 95 (9 2005), p. 090503.
[81] G. Gour. “Quantum Resource Theories in the Single-Shot Regime”. In: *Phys. Rev. A* 65 (6 2017).
[82] B. Yadin et al. “Quantum Processes Which Do Not Use Coherence”. In: *Phys. Rev. X* 6 (4 2016), p. 041028.
[83] L. Lami, B. Regula, and G. Adesso. “Generic Bound Coherence under Strictly Incoherent Operations”. In: *Phys. Rev. Lett.* 122 (15 2019), p. 150402.
[84] E. Chitambar and G. Gour. “Comparison of incoherent operations and measures of coherence”. In: *Phys. Rev. A* 94 (5 2016), p. 052336.
[85] A. Bach and U. Lüxmann-Ellinghaus. “The simplex structure of the classical states of the quantum harmonic oscillator”. In: *Commun. Math. Phys.* 107.4 (1986), pp. 553–560.

[86] K. C. Tan and H. Jeong. “Nonclassical light and metrological power: An introductory review”. In: *AVS Quantum Science* 1.1 (2019), p. 014701.
[87] J. Sperling. “Characterizing Maximally Singular Phase-Space Distributions”. In: *Phys. Rev. A* 94.1 (2016).
[88] S. Bochner. “Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse”. In: *Math. Ann.* 108.1 (1933), pp. 378–410.
[89] R. Simon et al. “Necessary and Sufficient Classicality Conditions on Photon Number Distributions”. In: *arXiv:quant-ph/9709030* (1997).
[90] K. C. Tan, S. Choi, and H. Jeong. “Negativity of Quasiprobability Distributions as a Measure of Nonclassicality”. In: *Phys. Rev. Lett.* 124 (11 2020), p. 110404.
[91] T. Kiesel and W. Vogel. “Nonclassicality Filters and QuasiProbabilities”. In: *Phys. Rev. A* 82 (3 2010).
[92] B. Regula et al. , in preparation.
[93] Steven George Krantz. *Handbook of Complex Variables*. Springer, 1999.
[94] M. M. Wolf, J. Eisert, and M. B. Plenio. “Entangling Power of Passive Optical Elements”. In: *Phys. Rev. Lett.* 90 (4 2003), p. 047904.
[95] Y. Aharonov et al. “A quantum characterization of classical radiation”. In: *Ann. Phys. (N. Y.)* 39.3 (1966), pp. 498–512.
[96] M. S. Kim et al. “Entanglement by a beam splitter: Nonclassicality as a prerequisite for entanglement”. In: *Phys. Rev. A* 65 (3 2002), p. 032323.
[97] X.-B. Wang. “Theorem for the beam-splitter entangler”. In: *Phys. Rev. A* 66 (2 2002), p. 024303.
[98] J. K. Asbóth, J. Calsamiglia, and H. Ritsch. “Computable Measure of Nonclassicality for Light”. In: *Phys. Rev. Lett.* 94 (17 2005), p. 173602.
[99] R. Cramér. *Mathematical Methods of Statistic*. Princeton University Press, 1946.
[100] J. Joo, W. J. Munro, and T. P. Spiller. “Quantum Estimation for Quantum Technology”. In: *int. j. quant. inf.* 7 (1 2009), pp. 125–137.
[101] M. Sion. “On general minimax theorems”. In: *Pacific J. Math.* 8.1 (1958), pp. 171–176.
[102] G. Ferrari et al. , in preparation.
[103] M. Berta et al. “Amortized channel divergence for asymptotic quantum channel discrimination”. In: *Preprint arXiv:1808.01498* (2018).
[104] F. G. S. L. Brandão and M. B. Plenio. “Entanglement theory and the second law of thermodynamics”. In: *Nat. Phys.* 4 (2008), 873 EP -.
[105] M. Piani. “Relative Entropy of Entanglement and Restricted Measurements”. In: *Phys. Rev. Lett.* 103 (16 2009), p. 160504.
[106] W. Rudin. *Functional Analysis*. McGraw-Hill, 1991.
[107] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.
[108] A. Winter. “Tight Uniform Continuity Bounds for Quantum Entropies: Conditional Entropy, Relative Entropy Distance and Energy Constraints”. In: *Commun. Math. Phys.* 347.1 (2016), pp. 291–313.
[109] M. Ohya and D. Petz. *Quantum Entropy and Its Use*. Theoretical and Mathematical Physics. Springer Berlin Heidelberg, 2004.
[110] M.S. Kim et al. “Entanglement by a beam splitter: Nonclassicality as a prerequisite for entanglement”. In: *Phys. Rev. A* 65.3 (2002).

[111] A. P. Lund et al. "Conditional Production of Superpositions of Coherent States with Inefficient Photon Detection". In: *Phys. Rev. A* 70.2 (2004).
[112] M. Reed and B. Simon. *Methods of Modern Mathematical Physics, Volume 1: Functional Analysis*. Academic Press, 1981.
[113] A. Winter. "Coding theorem and strong converse for quantum channels". In: *IEEE Trans. Inf. Theory* 45.7 (1999), pp. 2481–2485.
[114] E. B. Davies. "Quantum stochastic processes". In: *Commun. Math. Phys.* 15.4 (1969), pp. 277–304.
[115] G. Grubb. *Distributions and Operators*. Springer, 2009.
[116] P. R. Halmos. *Measure Theory*. Springer, 1950.