Mémoire pour l'obtention de
l'**Habilitation à Diriger des Recherches**
présenté à l'*Université Grenoble-Alpes*

Spécialité : Physique

Measurements, decoherence, and quantum correlations in composite quantum systems

Dominique Spehner

Institut Fourier
&
Laboratoire de Physique et Modélisation des Milieux Condensés,
Université Grenoble Alpes, F-38000 Grenoble, France

Date de soutenance : le 10 avril 2015

Composition du jury :

Stéphane ATTAL (rapporteur)
Roger BALIAN
Daniel BRAUN (rapporteur)
Stephan DE BIEVRE
Benoît DOUÇOT
Alain JOYE
Francesco PETRUCCIONE (rapporteur)

Remerciements

Je tiens tout d'abord à remercier profondément Stéphane Attal, Daniel Braun et Francesco Petruccione qui ont accepté la lourde tâche d'être rapporteurs de ce mémoire, ainsi que Roger Balian, Stephan De Bièvre, Benoît Douçot et Alain Joye pour m'avoir fait l'honneur d'être membres de mon jury. Leur présence à ma soutenance m'encourage énormément à poursuivre dans le chemin que je me suis tracé à l'interface entre la physique et les mathématiques, en évitant de "choisir mon camp", ce qui n'est pas toujours facile. Un grand merci en particulier à Alain qui m'a soutenu, conseillé et encouragé dans cette voie depuis mon arrivée à Grenoble, et ne s'est jamais lassé de me rappeler de temps à autre que je devrais peut-être songer à préparer mon HDR...

Je suis reconnaissant à Fritz Haake, Miguel Orzsag, Sylvain Vogelsberger, Giulia Ferrini, Anna Minguzzi, Frank Hekking et Krzysztof Pawlowski, sans qui les résultats présentés plus loin n'auraient pas pu voir le jour. Les échanges scientifiques pendant nos collaborations m'ont ouvert de nouveaux horizons et m'ont aidé à élargir le champ de mes recherches. D'autres personnes qui m'ont guidé ou avec qui j'ai travaillé dans le passé sur des sujets différents ont également eu un grand rôle dans le développement de mes centres d'intérêt. J'exprime ici ma gratitude à Jean Bellissard, Rolando Rebolledo, Hermann Schulz-Baldes, Alessandra Faggionato, Wojciech De Roeck, Thomas Schucker, Eric Akkermans et Romanas Narevich.

Les travaux décrits dans ce mémoire ont été effectués dans mes deux laboratoires de rattachement, l'Institut Fourier et le LPMMC. Je remercie chaleureusement les enseignant-chercheurs, chercheurs et doctorants de ces deux institutions pour leur bonne humeur, les discussions de tous les jours lors des repas, du café et des autres moments conviviaux, leur disponibilité pour parler de problèmes de mathématiques, de physique et de tout ce qui tourne autour de l'enseignement, et pour les interrogations sans fin sur les différences entre mathématiciens et physiciens. Grâce à l'amitié et au soutien qu'ils m'ont témoignés, je me suis toujours senti à ma place et parfaitement intégré dans les deux laboratoires et suis très fier d'en être membre. Merci à Françoise Truc pour avoir relu et corrigé les fautes d'anglais du premier chapitre. Je suis aussi redevable au personnel administratif de l'Institut Fourier et du LPMMC pour l'aide qu'on m'a apportée dans l'organisation de conférences, de missions et pour résoudre mes problèmes informatiques récurrents. Merci Jean-Daniel pour ta patience.

Tout ce qui va suivre a été rendu possible grâce aux encouragements et au soutien de ma bien-aimée et de mes parents. Je pense beaucoup à eux. Gracias Olga por amarme y por tu apoyo en los momentos dificiles.

Avec le violon, il faut choisir : ou bien tu joues juste, ou bien tu joues tzigane (Boby Lapointe).

Contents

Introduction

The topics discussed in this thesis are at the interface of two fields in quantum mechanics: the theory of open quantum systems and quantum information theory. The first field studies the dynamics of quantum systems coupled to their environment. The second one investigates how quantum systems can be used to accomplish information-processing tasks more efficiently than can be done classically, and what kinds of "quantum correlations" are responsible for that. A major role is played in these two subjects by quantum measurements.

Entanglement is one of the most intriguing features of the quantum world. It usually originates from interactions between the different parts (subsystems) of a composite quantum system. It manifests itself in particular in the experimentally demonstrated violation of the Bell inequalities. Entanglement of a system with its environment or with a measuring apparatus also plays a key role in irreversible dynamical processes like decoherence and the wave packet reduction. For a long time, entanglement has been identified with quantum correlations, that is, correlations which cannot be described by classical probability theory. This identification is fully justified for bipartite systems in pure states $|\Psi\rangle$: then the two subsystems are either entangled or fully decorrelated, the latter situation occurring when $|\Psi\rangle = |\psi\rangle \otimes |\phi\rangle$ is a tensor product of two subsystem states $|\psi\rangle$ and $|\phi\rangle$. In contrast, statistical ensembles of states may exhibit quantum correlations even if the corresponding mixed state has no entanglement. The quantum discord introduced in 2001 by Ollivier and Zurek [184] and Henderson and Vedral [120] is believed to be a more general measure of quantum correlations than entanglement. A full characterization and classification of the different kinds of quantum correlations in bipartite systems is, however, still lacking, and the situation is even worse in multipartite systems composed of three or more subsystems. A way to classify quantum correlations is to look them as resources required to accomplish specific tasks which either cannot be done classically or are less efficiently performed on classical systems. Such tasks may be related to a computational problem or to communication, for instance, sending encrypted information in a secure way [180]. Computational tasks are performed on a quantum computer made of qubits, that is, two-level quantum systems in arbitrary superpositions of $|0\rangle$ and $|1\rangle$ instead of being either in state 0 or 1 as with classical bits. A quantum algorithm is a unitary quantum evolution on a set of qubits followed by a measurement, the outcomes of which should provide the solution of the problem. It has been recently shown [145] that in order to offer an exponential speedup over classical computers, a quantum computation using pure states must necessarily produce multipartite entanglement which is not restricted to blocks of qubits of fixed size as the problem size increases. In contrast, the so-called quantum computation with one qubit (DCQ1) [147] is an algorithm using $(n+1)$ qubits to compute the trace of a $2^n \times 2^n$ unitary matrix exponentially faster than all known classical algorithms, in spite of the fact that the amount of entanglement for any bipartition of the $(n+1)$ qubits is bounded independently of n [71]. The mixed states appearing in the computation have in general a non vanishing quantum discord [72].

Unlike classical systems, quantum systems cannot typically be considered as completely isolated from their environment. This is especially true for large systems, because the energy differences between their nearest levels can be extremely small (excepted for gaped systems), so that even a very weak coupling with the environment may induce level transitions. Theoretical tools to study the dynamics of open quantum systems have been developed in the last century first in nuclear physics and then in quantum optics [65, 46]. As soon as the particles composing the system have interacted in the past with surrounding particles in their environment, and are subsequently entangled with them, the system properties cannot be fully understood independently of the environment [194]. However, one is usually not interested by the environment itself, but only by its action on the system. One then traces out the environmental degrees of freedom in the system-environment wave function $|\Psi_{\mathsf{SE}}\rangle$. This leads to a non-negative trace-one operator $\rho_{\mathsf{S}} = \mathrm{tr}_{\mathsf{E}} |\Psi_{\mathsf{SE}}\rangle\langle\Psi_{\mathsf{SE}}|$ (here tr_{E} stands for the partial trace over the environment E), which can be identified with a density matrix (mixed state) called the reduced state of S. The necessity to describe the state of an open quantum system by a density matrix instead of a wave function does not arise from some lack of knowledge on the system state like in classical statistical physics. Rather, it testifies that S and E are entangled and therefore S has no definite state on its

own. Considering that S and E together form an isolated bipartite system SE, the wave function $|\Psi_{\mathsf{SE}}(t)\rangle$ of SE undergoes a unitary evolution. In contrast, the time evolution of the reduced state $\rho_{\mathsf{S}}(t)$ is irreversible because some information leaks into the environment and is irremediably lost. A widespread theoretical approach to study this irreversible dynamics is to derive and solve first-order (in time) differential equations for $\rho_{\mathsf{S}}(t)$, called master equations [65, 46]. These equations are obtained from the unitary evolution of $|\Psi_{\mathsf{SE}}(t)\rangle$ under suitable approximations, in which the system-environment coupling is almost always treated perturbatively. The most popular master equations are the Lindblad [159] and Bloch-Redfield [201] equations.

The irreversible dynamics of $\rho_{\mathsf{S}}(t)$ is characterized by two kinds of dynamical processes:

(i) *Decoherence*: depending on the coupling between S and E, certain linear superpositions are transformed into statistical mixtures of the same states, such that interferences as well as purely quantum effects like entanglement disappear.

(ii) *Relaxation*: the evolution towards an equilibrium state or an out-of-equilibrium stationary state.

The time scales of these two processes, $\tau_{\rm dec}$ and $\tau_{\rm diss}$, may be very different. The decoherence time $\tau_{\rm dec}$ depends crucially, in addition to the coupling between S and E, on the initial state of S, that is, on the initial linear combination of system states. It typically decreases with the distance between these states, in such a way that $\tau_{\rm dec} \ll \tau_{\rm diss}$ for linear combinations of macroscopically distinguishable states (macroscopic superpositions, also called "Schrödinger cat states" [211]). However, we will see below that for specific system-environment couplings this separation of time scales may not occur.

Since entangled states are linear superpositions of product states, decoherence processes usually destroy entanglement. However, it is possible to engineer artificial environments with the property that the system evolves towards an entangled stationary state [150]. When the different subsystems are coupled to the same degrees of freedom of the environment, the environment-mediated effective interaction between these subsystems may also produce entanglement, even if they are initially disentangled [44].

In this thesis we are interested by the time evolution of the amount of quantum correlations in composite quantum systems coupled to their environment. The impact of decoherence on these correlations is analyzed quantitatively in different situations: systems coupled to a macroscopic measuring apparatus, two qubits coupled to independent reservoirs, and cold atomic gases. We also present a general geometrical framework to quantify entanglement-like or discord-like quantum correlations. In many previous studies, the generation of entanglement by inter-particle interactions in the unitary dynamics of the system is treated separately from decoherence processes. In other words, one imagines that the particles first interact between themselves to get entangled, and then one studies how the resulting entangled state looses its entanglement under the influence of the environment. In this work, we treat the more realistic situation where both entanglement and decoherence processes occur simultaneously. Then the amount of quantum correlations in the system typically reaches a maximum at a certain time, as a result of a competition between the production of quantum correlations by the unitary dynamics and the decrease of quantum correlations due to decoherence effects.

The manuscript is divided into three parts. The first part is devoted to dynamics and relates our results on measurement processes and entanglement losses. More specifically, we will discuss:

1) a specific model and the time scales for a quantum measurement (chapter 1);

2) the time evolution of entanglement for quantum trajectories of two qubits coupled to independent reservoirs (chapter 2);

3) the time evolution of quantum correlations useful for high-precision interferometry in Bose-Josephson junctions (chapter 3).

The first chapter is a bit longer than the following ones, the reason being that it deals with fundamental issues in quantum mechanics that I particularly like. In this chapter, we describe a model for a projective quantum measurement in which the measured system is coupled during a finite time with a measuring apparatus, in such a way that it becomes entangled with a single degree of freedom of the apparatus (pointer). The pointer is in turn coupled to infinitely many other degrees of freedom of the macroscopic apparatus, which may be regarded as the environment. This coupling produces decoherence and occurs simultaneously with the system-pointer interaction. At the end of the measurement, a disentangled system-pointer state agreeing with the von Neumann projection postulate is obtained. The entanglement and decoherence times are determined explicitly in this model. This material is based on a joint work with F. Haake (see (1a)-(1c) in the publication list).

The second chapter presents some results obtained with S. Vogelsberger, who did his Ph.D. at the Institut Fourier under my co-supervision (2008-2012), his other supervisor being A. Joye. We consider the simplest

bipartite system composed of two qubits (Hilbert space $\mathbb{C}^2 \otimes \mathbb{C}^2 \simeq \mathbb{C}^4$), coupled to two independent infinite reservoirs. We focus on the Markovian dynamics of the qubits when continuous measurements are performed on the reservoirs. This dynamics consists of a random pure state evolution (quantum trajectory) given by a quantum jump process or a stochastic Schrödinger equation with classical white noises. We show that the mean entanglement of the quantum trajectories decreases exponentially, as opposed to the known complete loss of entanglement after a finite time in the absence of measurements obtained from the master equation for the density matrix [79, 81, 268]. Furthermore, we identify the measurement schemes on the reservoirs that better protect entanglement (see (2a) and (2b) in the publication list).

The third chapter summarizes a series of works in the field of cold atoms, which result from a collaboration with G. Ferrini during her Ph.D. at the LPMMC, with her thesis supervisors A. Minguzzi and F. Hekking, and, more recently, with K. Pawlowski (see (3a)-(3d) in the publication list). These works are devoted to Bose-Josephson junctions formed by clouds of ultracold atoms in Bose-Einstein condensates in two different modes (e.g. two hyperfine energy levels). Because of interactions between atoms, starting from an initial coherent state the unitary dynamics generates quantum correlated states such as spin squeezed states and macroscopic superpositions of coherent states. In experiments, decoherence due to noise and atom losses occurs simultaneously with the formation of these states, thereby limiting the amount of quantum correlations that can be reached. This research program has been initiated during a "délégation au CNRS" at the LPMMC in 2009 and 2010. The main motivation is to find experimentally realizable conditions under which superpositions of coherent states could be not too much degraded by decoherence and could be observed. I enjoyed the close connexion to experiments in this research, which gave me the opportunity to discuss with experimentalists from the groups of M. Oberthaler in Heidelberg and, more recently, P. Treutlein in Basel.

While we focus in the first part on the physical ideas and main results without entering too much in their technical derivations, the second part goes much deeper into mathematical details. One may say that the first part is physically oriented and the second one mathematically oriented. This second part is a (slightly revised version of a) survey article published recently in J. Math. Phys. (see (4a) in the list of publications). This explains the length of this part, made of 9 chapters. The material of these chapters is completely self-contained and independent of the first part. Needless to say, most of the results are not due to the author; the original work by the author is given in Sec. 12.2. We review some topics in quantum information theory related to quantum correlations and to the discrimination of non-orthogonal states, without any reference to a particular physical system and dynamical evolution. The results presented, whose proof are given explicitly, apply to any quantum system with a finite-dimensional Hilbert space. Chapter 12 is devoted to a geometrical measure of discord-like quantum correlations based on the Bures distance on the set of quantum states. We show an explicit link of this measure with a quantum state discrimination problem, pointed out in a joint paper (4c) with M. Orszag. The other topics covered in this second part include generalized and least square measurements (chapter 5), state discrimination (chapter 6), quantum relative entropies (chapter 7), the Bures distance (chapter 8), the quantum Fisher information and the quantum Chernoff bound (chapter 9), bipartite entanglement (chapter 10), and the quantum discord (chapter 11).

Finally, the third (and shorter) part presents some perspectives of our work.

Some of the works I have done in the last ten years are not included here (see the publication list). The reason is that they are either too recent or outside the scope of this manuscript.

Part I

Dynamics

Chapter 1

Models for a quantum measurement

Diese Apparate haben zur Folge, daß sich die "Phase" des Atoms um prinzipiell, unkontrollierbare Beträge ändert, ebenso, wie sich bei einer Bestimmung des Elektronenortes der Impuls ändert [...] Die endgültige Transformationsmatrix e_{nl} [...] nicht mehr durch $\sum_m c_{nm}d_{ml}$ gegeben ist, sondern jedes Glied der Summe hat noch einen unbekannten Phasenfaktor. Wir können also nur erwarten, daß der Mittelwert von $e_{nl}\overline{e}_{nl}$ über alle diese eventuellen Phasenänderungen gleich $\sum_m c_{nm}\overline{c}_{nm}d_{ml}\overline{c}_{ml}$ ist.[1] (W. Heisenberg, 1927) [117].

1.1 Motivations

Since the birth of quantum mechanics, physicists and mathematicians have devoted a lot of efforts to understand and describe theoretically measurement processes in quantum systems (see the monographs and survey articles [253, 258, 40, 102, 194, 277, 210, 6]). The primary motivation of these works was to investigate the foundations of the quantum theory and its interpretation problems, a subject still under debate. The strong perturbation of a quantum system by the measuring apparatus makes measurements in quantum mechanics quite peculiar, this perturbation being related to fundamental principles such as the uncertainty principle and the statistical nature of the quantum predictions [117]. A renewal of interest for measurement processes came in the last decades from experiments which have achieved to store, manipulate, and observe single quantum systems. In these experiments mostly stimulated by applications to quantum information technologies, measurements are performed to monitor the system state. For instance, a single measurement is used to teleport a state or to entangle two subsystems. Repeated measurements lead to the quantum Zeno effect [94, 235] when done directly on the system and to quantum trajectories [199] when done on the environment interacting with it. Beside the possibility to monitor single systems, measurements are crucial in quantum algorithms to get the result of a quantum computation and in quantum communication to extract classical information out of transmitted quantum information. In this chapter, we restrict ourselves to von Neumann projective measurements. More general kinds of measurements are obtained by letting the system interact with an ancilla and performing projective measurements on the latter. These generalized measurements play an important role in quantum information theory and will be considered in chapters 5 and 6. In particular, we will show that they can sometimes be useful to distinguish in a more reliable way non-orthogonal states from a given ensemble.

As one learns during a first year course in quantum theory, the effect of a single run of a measurement is to transform the system state into the projection of this state on an eigenspace of the measured observable, up to normalization. This state transformation is called the reduction of the wave packet. It was considered by the founding fathers of quantum mechanics (in particular von Neumann) as a specific postulate of the theory, completely independent from the "natural" state evolution given by unitary transformations according to the Schrödinger equation [253, 258]. A fundamental question is to know whether such a postulate is actually needed, that is, if the reduction of the wave packet could be derived from Schrödinger's equation applied to the system and measuring apparatus. A positive answer seems to have emerged over the years, the most convincing arguments being in our opinion given in a recent work by Allahverdyan, Balian, and Nieuwenhuizen [6, 7]. To tackle this problem, these authors, as many other authors in the last decades [48, 276, 277, 210, 223, 224], consider a particular model of a system coupled to a macroscopic apparatus and study the corresponding

[1] "The consequence of this apparatus is that the 'phase' of the atom changes by quantities that are uncontrollable in principle, just as the impulse was changed in the determination of the electron's position [...] The final transformation matrix e_{nl} [...] is no longer given by $\sum_m c_{nm}d_{ml}$, and instead each term of the sum will have, in addition, an unknown phase factor. Hence, all we can expect is for the average value of $e_{nl}\overline{e}_{nl}$, over all eventual phase changes, to be equal to $\sum_m c_{nm}\overline{c}_{nm}d_{ml}\overline{c}_{ml}$."

dynamics. A measurement is viewed as a quantum dynamical process originating from some unitary evolution of the measured system and apparatus.

In this first chapter, we focus on the general features and a specific model of an ideal quantum measurement. It is neither our purpose here to discuss refinements or alternative interpretations of the Copenhagen formulation of quantum mechanics like the consistent histories approach or Everett's many-worlds interpretation, nor to describe non-local hidden variable theories like Bohmian mechanics[2]. In Sec. 1.2, we elaborate on the physical meaning of the wave packet collapse, the origin of irreversibility, and the role played by decoherence in measurement processes. We then describe in Sec. 1.3 a particular model proposed by the author and F. Haake (see (1a)-(1c) in the publication list). Most interestingly, we determine explicitly the various time scales of the measurement process. To our knowledge, this has not been done before in a realistic process excepted in the Curie-Weiss model of Refs. [5, 6]. Finally, inspired by the work of Allahverdyan *et al.*, we use in Sec. 1.4 the statistical ensemble interpretation of quantum states, according to which a wave function does not describe the state of a single system but of an ensemble of similarly prepared systems, to explain how in our model individual runs of a measurement yield single outcomes when the value of the pointer macroscopic observable is read by an observer.

1.2 General features of measurement processes

1.2.1 Macroscopic measuring apparatus

In order to measure the value of an observable of a system S, this system must interact during a finite period of time $t_{\rm int}$ with a measuring apparatus M, in such a way that some information on the state of S be transferred to M. If S is classical, the perturbation of its state resulting from this interaction can be neglected, at least for a good enough apparatus (which could in principle be constructed via technical improvements). In contrast, it is never possible to neglect the perturbation made by the apparatus on the state of a quantum system, excepted when it is initially in an eigenstate of the measured observable. For instance, if one sheds light on a particle to measure its position, the photons will give small momentum kicks to the particle in arbitrary directions; the resulting uncertainty Δp in the momentum of the particle satisfies $\Delta x \Delta p \geq \hbar$, where Δx is the precision of the position measurement [118].

In an ideal quantum measurement, the measuring apparatus M must fulfill the following requirements:

1. M is *macroscopic* and possesses a "pointer" variable X with a quasi-classical behavior, to be used as readout of the measurement outcomes. At time $t = 0$ when S starts to interact with M, the value x_0 of X is precisely known.

2. After the S-M interaction has been switched off at time $t = t_{\rm int}$, the eigenvalues s_i of the measured observable S are *perfectly correlated* with the values x_i of the pointer observable X.

3. At the end of the measurement (time $t = t_{\rm meas}$), M should be in one of distinct *stable equilibrium states* with expectation value x_i of X and X having negligible fluctuations on the macroscopic scale. Moreover, the x_i must be *macroscopically distinguishable* for distinct eigenvalues s_i (registration of the result).

One may think of the pointer observable X as the position of the center of mass of the needle of a meter. Thanks to conditions 1 and 3, the value of X will not be perturbed noticeably by an observer looking at the measurement result, so that he does not need to perform a new quantum measurement to obtain it. Furthermore, the measurement outcomes being encoded in stable equilibria of the apparatus, they are registered in a robust and permanent way. This enable to disregard the observer in the measurement process[3]. Conditions 2 and 3 mean that the interaction between S and M provokes a macroscopic change in the state of M. For instance, if X is the position of a meter needle, the expectation values x_i tied up with distinct eigenvalues s_i should be separated by macroscopic distances at time $t_{\rm meas}$. Since S is typically a small system, it can only perturb M weakly. This small perturbation must be either subsequently amplified, so as to lead to macroscopic changes of X, or it must provoke a symmetry breaking in the apparatus initially in a metastable state, so as to drive M into one particular equilibrium. Amplifications of small signals are used e.g. in photo-detectors. Many measurements actually involve a chain $\{\mathsf{M}_n\}_{n=1,\ldots,N}$ of apparatus (cascade): only the first apparatus M_1 (which is not necessarily macroscopic) is in contact with S; each apparatus M_n measures one after the other the

[2]Two reasons for not discussing these approaches here are: (1) I am not familiar with them (2) the main goal of this chapter is to study a concrete model of measurement within the framework of the Copenhagen interpretation and using a statistical physics viewpoint.

[3]For a discussion about the problem of including the observer in the description a measurement, see e.g. [40].

observable X_{n-1} of the previous apparatus; finally, the observer reads the result on the pointer variable X_N of the last apparatus M_N, which satisfies conditions 1-3 above. We will not consider here such complicated chains. Rather, we restrict our discussion to a single apparatus M, assumed to be initially in a metastable state. As emphasized in Refs. [4, 5, 6], the relaxation of M towards one among several equilibrium states can be achieved if M is in a metastable phase of a phase transition. This is the case for instance in a bubble chamber, where a particle moving in an overheated liquid produces local transitions to a stable gaseous phase along its trajectory. One may also think of an apparatus cooled below the critical temperature of a second order phase transition. An example is given by $N \gg 1$ interacting spins identically coupled to a single spin, the latter constituting the measured system (Curie-Weiss model) [5, 6]. Then the pointer observable X is the magnetization of the N spins, which is the order parameter of the transition. In Sec. 1.3, we will consider instead an apparatus M with a single degree of freedom pointer, whose position X is initially at the center $x = 0$ of a symmetric potential up to small thermal fluctuations. The potential is chosen such that an instability occurs when M is coupled to the system, so that the latter can easily drive the pointer outside of the central well.

Because the macroscopic apparatus M is made of atoms, it can be described quantum-mechanically, but statistical physics is required. The initial state of M is given by a density matrix $\rho_{\mathsf{M}}(0)$, associated to a statistical ensemble of states[4]. Inasmuch as S and M have not interacted before $t = 0$, they are not correlated at that time and the bipartite system SM is initially in the product state

$$\rho_{\mathsf{SM}}(0) = \rho_{\mathsf{S}}(0) \otimes \rho_{\mathsf{M}}(0) \ , \tag{1.1}$$

where $\rho_{\mathsf{S}}(0)$ is the state of the system S just before the measurement. Hereafter, we denote by $\Pi_i = \sum_l |\alpha_{il}\rangle\langle\alpha_{il}|$ the eigenprojector with eigenvalue s_i of the measured observable $S = \sum_i s_i \Pi_i$, with $S|\alpha_{il}\rangle = s_i|\alpha_{il}\rangle$ and $\langle\alpha_{il}|\alpha_{jm}\rangle = \delta_{ij}\delta_{lm}$ (the index l accounts for spectral degeneracies). The apparatus state tied up with the eigenvalue s_i at time $t \geq t_{\rm int}$ is denoted by $\rho_{\mathsf{M}|i}(t)$. According to condition 3, each of these states evolve towards an equilibrium with density matrix $\rho_{\mathsf{M}|i}(t_{\rm meas}) = \rho^{\rm eq}_{\mathsf{M}|i}$ and expectation value $x_i = {\rm tr}(\rho^{\rm eq}_{\mathsf{M}|i} X)$ of the pointer observable. Furthermore, the $\rho^{\rm eq}_{\mathsf{M}|i}$ are almost orthogonal,

$${\rm tr}\left(\rho^{\rm eq}_{\mathsf{M}|i}\rho^{\rm eq}_{\mathsf{M}|j}\right) \simeq 0 \quad \text{for } i \neq j. \tag{1.2}$$

1.2.2 Reduction of the wave packet

Within the standard interpretation of quantum mechanics, only the probabilities of the measurement outcomes can be predicted, even if the initial wave function of S is perfectly known (i.e., S is in a pure state). Probabilities are introduced as a fundamental ingredient of the theory, unlike in classical statistical physics where they arise from the impossibility to know in practice the positions and velocities of all particles. It is meaningless to give definite values to outcomes of *unperformed* measurements, this would indeed lead to contradictions with the theory [194]. Local and even some non-local hidden-variable descriptions, which stipulate that the outcomes depend on pre-existing properties of the system independently of the measurement, have been ruled out experimentally by the observation of the violation of the Bell and similar inequalities [15, 105].

In ideal measurements, repeated consecutive measurements of the same observable on a single system always give the same result. For indeed, the signature of imperfect measurements is a nonzero gain of information on the system when one performs the same measurement a second time. A related property of ideal measurements is that if one repeats them twice on a single system, its state does not change during the second measurement. This condition and condition 2 of Sec. 1.2.1 are satisfied if the system-apparatus density matrix is transformed as follows[5]

$$\rho_{\mathsf{S}}(0) \otimes \rho_{\mathsf{M}}(0) \ \longrightarrow \ \rho_{\mathsf{SM}}(t) = \sum_i p_i \rho_{\mathsf{S}|i} \otimes \rho_{\mathsf{M}|i}(t) \quad \text{for } t \geq t_{\rm int} \ , \tag{1.3}$$

with

$$p_i\rho_{\mathsf{S}|i} = \Pi_i \rho_{\mathsf{S}}(0) \Pi_i \quad , \quad p_i = {\rm tr}[\Pi_i \rho_{\mathsf{S}}(0)] \ . \tag{1.4}$$

In fact, the state (1.3) encompasses perfect *classical* correlations[6] between S and M. If as stated previously, $\rho_{\mathsf{M}|i}(t)$ relaxes to an equilibrium $\rho^{\rm eq}_{\mathsf{M}|i}$ in the absence of system-apparatus interaction, then condition 3 is fulfilled for $t_{\rm meas} - t_{\rm int}$ larger than the relaxation time $\tau_{\rm rel}$. To the expense that $\rho_{\mathsf{SM}}(t)$ can be interpreted as a statistical ensemble of states $\rho_{\mathsf{S}|i} \otimes \rho_{\mathsf{M}|i}(t)$ with probabilities p_i, (1.3) is a mathematical formulation of the measurement postulate, namely:

[4] Pure states are produced experimentally by measuring a complete set of commuting observables and post-selecting a set of measurement outcomes. This cannot be done on a macroscopic body, since the number of observables to measure is too large.

[5] Conversely, the invariance of the system state under repeated measurements implies that this state is transformed as $\rho_{\mathsf{S}}(0) \to \sum_i p_i \rho_{\mathsf{S}|i}$ with p_i and $\rho_{\mathsf{S}|i}$ given by (1.4) [21].

[6] See chapter 11 for a definition of classical correlations, as opposed to quantum correlations.

- **Born rules:** the probability to find at the end of the measurement the outcome i (that is, the expectation value x_i of the pointer observable) is equal to $p_i = \mathrm{tr}[\Pi_i \rho_{\mathsf{S}}(0)]$.
- **Reduction of the wave packet:** given that the outcome is i, the system is at time $t = t_{\mathrm{meas}}$ in the state $\rho_{\mathsf{S}|i}$.

The meaning of the Born rules is as follows: if one runs the measurement many times on identical systems initially in the same state $\rho_{\mathsf{S}}(0)$, the fraction of outcomes i will be p_i. The reduction of the wave packet, on the contrary, concerns a single run. It can be interpreted in various ways, depending on the physical meaning given to the collapse

$$\rho_{\mathsf{S}}(0) \otimes \rho_{\mathsf{M}}(0) \;\longrightarrow\; \rho_{\mathsf{S}|i} \otimes \rho^{\mathrm{eq}}_{\mathsf{M}|i} \qquad \text{if the measurement outcome is } i. \tag{1.5}$$

Some authors consider the transformation (1.5) as a proper dynamical process. Instead, we think that (1.5) is simply due to the gain of information when the pointer value is readout at the end of the measurement. This point of view is shared for instance by D. Bohm in his 1951 book. Bohm writes that "the sudden replacement of the statistical ensemble of wave functions by a single wave function represents absolutely no change in the state, but is analogous to the sudden changes in classical probability functions which accompany an improvement of the observer's information" (opus cit. [40], Sec. 22.10)[7]. In mathematical terms, (1.5) is nothing but a conditioning given the event $X = x_i$ (hence the notation $\rho_{\mathsf{S}|i}$ used for the corresponding post-measurement state). Therefore, it makes no sense to look for models exhibiting real collapses as in (1.5). The measurement postulate can be fully established if one is able to

1) show that the dynamical evolution of the density matrix of the system and apparatus is given by (1.3);
2) explain how the final system-apparatus state $\rho_{\mathsf{SM}}(t_{\mathrm{meas}})$ can be interpreted as the state ensemble $\{\rho_{\mathsf{S}|i} \otimes \rho^{\mathrm{eq}}_{\mathsf{M}|i}, p_i\}$.

According to the Bohr interpretation of quantum mechanics, one should adopt a minimalist description in which the state of a given system solely contains the information about this system that we may have access to. Inasmuch as measurements are the only mean by which one can gain information about a system, a state should enable us to compute the probabilities of all measurement outcomes, but does not contain any indication on which precise outcome will occur in a single measurement. Using the language of statistical physics, a state actually describes a thought ensemble of identical systems, from which one has one randomly chosen copy. This statistical interpretation of quantum states is a consequence of the minimalist prescription and of the measurement postulate.

Our contribution concerns the point 1) above. We will indeed present in Sec. 1.3 a physical model of an apparatus for which the transformation (1.3) is realized. Note that this transformation cannot result from a unitary evolution, i.e., it cannot be implemented by a unitary operator acting on the system and apparatus. This is not so surprising because irreversibility (and thus non-unitary dynamics) should be expected from the macroscopic nature of M. Irreversibility is acknowledged by noting that different initial states (1.1) may be transformed into the same final state (take for instance $\rho_{\mathsf{S}}(0)$ not commuting with S and $\rho'_{\mathsf{S}}(0) = \sum_i \Pi_i \rho_{\mathsf{S}}(0) \Pi_i$). Another confirmation is given by determining the von Neumann entropy production[8] $(\Delta S)_{\mathrm{v.N.}}$ between $t = 0$ and $t = t_{\mathrm{meas}}$. A simple calculation yields

$$(\Delta S)_{\mathrm{v.N.}} = S_{\mathrm{v.N.}}\Big(\sum_i p_i \rho_{\mathsf{S}|i}\Big) - S_{\mathrm{v.N.}}\big(\rho_{\mathsf{S}}(0)\big) + \sum_i p_i S_{\mathrm{v.N.}}\big(\rho^{\mathrm{eq}}_{\mathsf{M}|i}\big) - S_{\mathrm{v.N.}}\big(\rho_{\mathsf{M}}(0)\big)\,. \tag{1.6}$$

The difference of the two first terms corresponds to the increase of entropy in the decoherence process

$$\rho_{\mathsf{S}}(0) = \sum_{ij} \Pi_i \rho_{\mathsf{S}}(0) \Pi_j \;\longrightarrow\; \rho_{\mathsf{S}}(t_{\mathrm{meas}}) = \sum_i \Pi_i \rho_{\mathsf{S}}(0) \Pi_i \tag{1.7}$$

in which the off-diagonal terms $i \neq j$ have been erased (see Sec. 1.2.5 below). This sum is strictly positive[9] provided that $\rho_{\mathsf{S}}(0)$ does not commute with the projectors Π_i. The two last terms in (1.6) quantify the average increase of entropy due to the relaxation of the initial apparatus state $\rho_{\mathsf{M}}(0)$ towards its stable equilibria

[7] This point of view was already expressed by Heisenberg in his pioneering paper [117].

[8] For a definition and properties of the von Neumann and other entropies, see chapter 7.

[9] In fact, using (1.4) and (1.7) one finds that $S_{\mathrm{v.N.}}\big(\rho_{\mathsf{S}}(t_{\mathrm{meas}})\big) - S_{\mathrm{v.N.}}\big(\rho_{\mathsf{S}}(0)\big)$ is equal to the conditional entropy

$$S_{\mathrm{v.N.}}\big(\rho_{\mathsf{S}}(0) || \rho_{\mathsf{S}}(t_{\mathrm{meas}})\big) = \mathrm{tr}\left[\rho_{\mathsf{S}}(0)\big(\ln \rho_{\mathsf{S}}(0) - \ln \rho_{\mathsf{S}}(t_{\mathrm{meas}})\big)\right],$$

which is known to be strictly positive if $\rho_{\mathsf{S}}(0) \neq \rho_{\mathsf{S}}(t_{\mathrm{meas}})$ (see chapter 7).

$\rho^{\rm eq}_{\mathsf{M}|i}$ which have maximal entropies[10] $S_{\rm v.N.}(\rho^{\rm eq}_{\mathsf{M}|i}) > S_{\rm v.N.}(\rho_{\mathsf{M}}(0))$. The positivity of $(\Delta S)_{\rm v.N.}$ means that the measurement process erases some information in the system and apparatus. However, if the final pointer value x_i is known, the final state being as in (1.5) with probability p_i, the entropy balance for the system is negative,

$$\sum_i p_i S_{\rm v.N.}\big(\rho_{\mathsf{S}|i}\big) - S_{\rm v.N.}\big(\rho_{\mathsf{S}}(0)\big) \;\leq\; 0\,. \tag{1.8}$$

As it should be the case, the measurement increases our knowledge about the system, in spite of the loss of information in the system and apparatus in the absence of readout of the pointer value. A rigorous proof of (1.8) can be found in Ref. [156].

1.2.3 Density matrices and state ensembles

We now come back to the point 2) raised above. It is important here to stress the difference between density matrices and statistical ensembles of states, i.e., collections of quantum states $\rho_i \geq 0$ (with $\mathrm{tr}\,\rho_i = 1$) to which one attaches some probabilities $p_i > 0$ (with $\sum_i p_i = 1$). A density matrix is associated to the ensemble $\{\rho_i, p_i\}$ if $\rho = \sum_i p_i\rho_i$, that is, the ensemble defines a convex decomposition of ρ. It is convenient to think of this decomposition as a (fictitious) *state preparation*. More precisely, let us imagine that some experimental device ("black box") prepares many identical systems with a fraction p_i of them in state ρ_i. An observer receives one of such systems chosen at random. He only knows that he has a chance p_i to get a system in state ρ_i. It is an obvious mathematical fact that a density matrix admits infinitely many convex decompositions (excepted of course if it is a pure state)[11]. This statement has a fundamental physical interpretation in the quantum theory, which has strong implications and whose importance is disregarded in many textbooks. Indeed, let us consider two state preparations $\{\rho_i, p_i\}_{i=1}^m$ and $\{\sigma_j, q_j\}_{j=1}^p$ made by two different devices operating on identical systems, which correspond to the same density matrix $\rho = \sum_i p_i\rho_i = \sum_j q_j\sigma_j$. Since ρ contains all knowledge about the outcome probabilities of any measurement according to Born's rules, an observer receiving many copies of the system has absolutely no way to know from which of the two devices they come from. In other words, the full information that one can collect on a system via measurements is encoded in ρ and does not require the knowledge of the ensemble involved in the state preparation[12]. A measurement process including the reduction of the wave packet must be viewed as a state preparation. More precisely, if the measurement is repeated many times on identical copies and one selects all runs giving the pointer value x_i, one thereby prepares the system and apparatus in the state $\rho_{\mathsf{SM}|i} = \rho_{\mathsf{S}} \otimes \rho^{\rm eq}_{\mathsf{M}|i}$ with probability p_i. A polarizer transmitting light linearly polarized along the x direction is an example of such a state post-selection: when a single photon with polarization along the unit vector $\mathbf{n}$ is sent to the polarizer, it is either absorbed with probability $(\mathbf{n}\cdot\mathbf{e}_y)^2$ or transmitted with probability $(\mathbf{n}\cdot\mathbf{e}_x)^2$, and in the latter case it has a polarization $\mathbf{e}_x$ at the exit of the polarizer.

Therefore, to derive the measurement postulate one has not only to show that the density matrix of the system and apparatus is transformed according to (1.3), but also to prove that the system-apparatus dynamics prepares the state ensemble $\{\rho_{\mathsf{SM}|i}, p_i\}$ and that the readout of the pointer variable provides us complete information about this ensemble. In other words, in order to explain that individual runs of a measurement lead to well-defined outcomes i and to the collapse (1.5) if outcome i occurs, the sole knowledge of the density operator $\rho_{\mathsf{SM}}(t_{\rm meas})$ is not enough. One has in addition to argue that the measurement process is equivalent to a preparation of SM in the states $\rho_{\mathsf{SM}|i}$ with probabilities p_i given by Born's rules. It is indeed not granted from (1.3) that if one knows the outcome i of a single run, the corresponding state that emerges should be $\rho_{\mathsf{SM}|i}$, instead of some state from another ensemble with the same density matrix $\rho_{\mathsf{SM}}(t_{\rm meas})$. A related deeper question is to understand how classical probabilities may emerge from a quantum evolution. We will come back to these issues in Sec. 1.4.

1.2.4 Origin of the irreversibility of a measurement

As stated in the introduction, macroscopic bodies cannot be considered as isolated from their environment. This is in particular the case for the macroscopic apparatus M. Hence statistical physics is not only required because the initial state of M is a statistical mixture, but also because one must take into account the coupling of M

[10] In order to avoid any bias in the measurement, one should in fact require that all equilibria $\rho^{\rm eq}_{\mathsf{M}|i}$ have the same entropy.

[11] Actually, given two orthogonal eigenstates $|\psi_k\rangle$ and $|\psi_l\rangle$ of ρ with eigenvalues p_k and p_l, one has

$$p_k|\psi_k\rangle\langle\psi_k| + p_l|\psi_l\rangle\langle\psi_l| = p_+|\psi_+\rangle\langle\psi_+| + p_-|\psi_-\rangle\langle\psi_-| \quad \text{with} \quad \sqrt{p_\pm}|\psi_\pm\rangle = \sqrt{p_k}|\psi_k\rangle \pm \mathrm{e}^{\mathrm{i}\varphi}\sqrt{p_l}|\psi_l\rangle\,,$$

where φ is an arbitrary phase.

[12] This "quantum ambiguity" explains why meaningful definitions of entanglement based on the entanglement of the pure states $|\Psi_i\rangle$ in a convex decomposition $\rho = \sum_i p_i|\Psi_i\rangle\langle\Psi_i|$ must involve a minimization over all such decompositions (see Sec. 2.3).

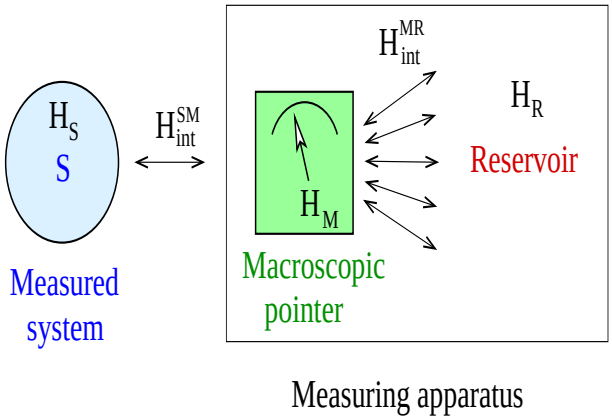

Figure 1.1: Model for a quantum measurement: the system S interacts with the pointer of a measuring apparatus which is itself coupled to an infinite reservoir formed by the microscopic degrees of freedom of the apparatus.

with its environment. This environment may consist of the inaccessible microscopic degrees of freedom of the apparatus itself. It is convenient to redefine the apparatus in such a way that it contains only the macroscopic degrees of freedom, the other ones being included in the environment which acts as an infinite reservoir R (this separation is of course arbitrary, but the only important point in what follows is that R is infinite; this is always the case for a macroscopic apparatus in the thermodynamic limit). The irreversible dynamics of the system and apparatus results from a transfer of information from SM to R caused by the coupling between M and R. As time evolves, this transferred information is spread over infinitely many degrees of freedom of R and cannot be retrieved. Therefore, a realistic quantum measurement model must take into account both the system-apparatus and the apparatus-reservoir couplings. The three-partite system SMR can be assumed to be isolated, its dynamics being reversible and given by the Schrödinger equation. As it has been recalled in the introduction, the system-apparatus reduced state is defined as the partial trace over the reservoir of the density matrix $\rho_{\mathsf{SMR}}(t)$ of SMR,

$$\rho_{\mathsf{SM}}(t) = \mathrm{tr}_{\mathsf{R}}[\rho_{\mathsf{SMR}}(t)] \quad , \quad \rho_{\mathsf{SMR}}(t) = \mathrm{e}^{-\mathrm{i}tH_{\mathsf{SMR}}}\rho_{\mathsf{SMR}}(0)\,\mathrm{e}^{\mathrm{i}tH_{\mathsf{SMR}}} \; . \tag{1.9}$$

The Hamiltonian H_{SMR} is the sum of the free Hamiltonians H_{S}, H_{M}, and H_{R} of S, M, and R, and of the system-apparatus and apparatus-reservoir couplings $H_{\mathrm{int}}^{\mathsf{SM}}$ and $H_{\mathrm{int}}^{\mathsf{MR}}$,

$$H_{\mathsf{SMR}} = H_{\mathsf{S}} \otimes 1_{\mathsf{M}} \otimes 1_{\mathsf{R}} + 1_{\mathsf{S}} \otimes H_{\mathsf{M}} \otimes 1_{\mathsf{R}} + 1_{\mathsf{S}} \otimes 1_{\mathsf{M}} \otimes H_{\mathsf{R}} + H_{\mathrm{int}}^{\mathsf{SM}} \otimes 1_{\mathsf{R}} + 1_{\mathsf{S}} \otimes H_{\mathrm{int}}^{\mathsf{MR}} \; . \tag{1.10}$$

Hereafter, 1_{A} denotes the identity operator acting on the Hilbert space of the system indicated by the subscript A (which will not always appear explicitly). For brevity, we will sometimes omit these identity operators when the Hilbert space is clear from the context, e.g. we identify H_{S} with $H_{\mathsf{S}} \otimes 1_{\mathsf{M}} \otimes 1_{\mathsf{R}}$. The interaction between S and R does not play an important role in the measurement[13] and can be neglected.

We can now formulate the fourth requirement characterizing ideal measurements [259]:

4. *The coupling between* S *and* M *commutes with all eigenprojectors of* S*, that is, for all orthonormal eigenbasis* $\{|\alpha_{il}\rangle\}$ *of* S*,* $[H_{\mathrm{int}}^{\mathsf{SM}}, \Pi_{il} \otimes 1] = 0$ *with* $\Pi_{il} = |\alpha_{il}\rangle\langle\alpha_{il}|$.

This condition ensures that if S is initially in an eigenstate $|\alpha_{il}\rangle$ of S, then its state will be unchanged. In other words, an ideal measurement does not modify the measured observable. In principle, it is necessary that the whole Hamiltonian (1.10) commutes with the projectors Π_{il}, so that one should also require $[H_{\mathsf{S}}, \Pi_{il}] = 0$. However, in practice it is sufficient that the time T_{S} of the evolution of the eigenvectors $|\alpha_{il}\rangle$ of S under the dynamics generated by H_{S} be much larger than the time duration t_{meas} of the measurement. Hence the system Hamiltonian H_{S} can be completely disregarded in the measurement process. Under these assumptions, (1.9) yields $\rho_{\mathsf{SMR}}(t) = \Pi_{il} \otimes \rho_{\mathsf{MR}}(t)$ whenever $\rho_{\mathsf{SMR}}(0) = \Pi_{il} \otimes \rho_{\mathsf{MR}}(0)$.

An example of coupling satisfying our fourth condition is

$$H_{\mathrm{int}}^{\mathsf{SM}} = gS \otimes P \; , \tag{1.11}$$

[13] Actually, it turns out that for a system S strongly coupled to M, decoherence caused by a direct coupling of S with the reservoir R has a much smaller effect than decoherence due to the indirect coupling between S and R mediated by the apparatus M [224].

where g is a coupling constant and P an observable of the apparatus. For instance, let us consider a spin one-half, labelled with the index 0, from which one wants to measure the z-component $S = \sigma_z^{(0)}$. The apparatus is composed of $N \gg 1$ spins one-half labelled from 1 to N (Curie-Weiss model). A coupling of the form (1.11) is $H_{\rm int}^{\sf SM} = -(g/\mu)\sigma_z^{(0)} \otimes M_z$, where $M_z = \mu \sum_{n=1}^N \sigma_z^{(n)}$ is the total magnetization of the apparatus (M_z is the pointer observable, see above) and μ is the magnetic moment of the spins. Depending on the spin values $s_\pm = \pm 1/2$, $H_{\rm int}^{\sf SM}$ acts as an external effective magnetic field $B_{\rm eff} = \mp g M_z/2$ provoking a symmetry breaking with a spontaneous magnetization in the critical phase [5, 6]. Another example, which will be analyzed in Sec. 1.3, is to define P as the total momentum of the apparatus. Then $H_{\rm int}^{\sf SM}$ generates translations of the center-of-mass position X by a distance $g s_i$ depending on the eigenvalue s_i.

For the coupling (1.11), the Gibbs state associated to the system-apparatus Hamiltonian $H_{\sf SM} = H_{\sf M} + H_{\rm int}^{\sf SM}$ with the constraints of fixed energy (micro-canonical ensemble) and fixed values of all observables commuting with S is of the form given by the post-measurement state (1.3) [6]. This Gibbs state is defined as $\rho_{\sf SM}^{\rm eq} = Z_{\sf SM}^{-1} {\rm e}^{-\beta H_{\sf SM} - \sum_i Y_i}$ with $Y_i = \sum_{l,m} y_{ilm} |\alpha_{il}\rangle\langle\alpha_{im}|$. Similarly, an easy calculation gives the state of $\sf SM$ at time t in the absence of coupling with the reservoir,

$$\rho_{\sf SM}^{\rm ent}(t) = \sum_{i,j} \Pi_i \rho_{\sf S}(0) \Pi_j \otimes {\rm e}^{-{\rm i}t H_{\sf M}^{(i)}} \rho_{\sf M}(0) {\rm e}^{{\rm i}t H_{\sf M}^{(j)}} \quad , \quad H_{\sf M}^{(i)} = H_{\sf M} + g s_i P \, , \tag{1.12}$$

where we have neglected the system Hamiltonian $H_{\sf S}$. The state (1.12) is in general entangled. It differs from the classical state (1.3) by the presence of the off-diagonal terms $i \neq j$ in the sum (coherences). To see entanglement more clearly, let us assume that S has a non-degenerate spectrum and that $\sf S$ and $\sf M$ are initially in pure states $|\psi(0)\rangle$ and $|\mu(0)\rangle$ (as stressed above, the latter assumption is not realistic as far as $\sf M$ is concerned). Then $\sf SM$ is at time $t_{\rm meas}$ in the pure state

$$|\Psi_{\sf SM}^{\rm ent}\rangle = \sum_i c_i |\alpha_i\rangle |\mu_i^{\rm eq}\rangle \quad , \quad c_i = \langle\alpha_i|\psi(0)\rangle \quad , \quad |\mu_i^{\rm eq}\rangle = {\rm e}^{-{\rm i}(t_{\rm meas}-t_{\rm int})H_{\sf M}} {\rm e}^{-{\rm i}t_{\rm int} H_{\sf M}^{(i)}} |\mu(0)\rangle \, . \tag{1.13}$$

If $|\psi(0)\rangle$ is not an eigenstate of S and $|\mu(0)\rangle$ is not an eigenstate of ${\rm e}^{{\rm i}t_{\rm int} H_{\sf M}^{(j)}} {\rm e}^{-{\rm i}t_{\rm int} H_{\sf M}^{(i)}}$ for all i and j, then $|\Psi_{\sf SM}^{\rm ent}\rangle$ cannot be written as a product state and is thus entangled[14].

Under the condition 3 of Sec. 1.2.1, the apparatus states $|\mu_i^{\rm eq}\rangle$ are macroscopically distinguishable. Hence (1.13) is a superposition of macroscopically distinct states ("Schrödinger cat state"). Such superpositions are known to be very fragile under a coupling to an environment, namely, they are transformed after an extremely short time $\tau_{\rm dec}$ into a statistical mixture,

$$|\Psi_{\sf SM}^{\rm ent}\rangle\langle\Psi_{\sf SM}^{\rm ent}| \;\longrightarrow\; \rho_{\sf SM}(t_{\rm meas}) = \sum_i |c_i|^2 |\alpha_i\rangle\langle\alpha_i| \otimes |\mu_i^{\rm eq}\rangle\langle\mu_i^{\rm eq}| \, . \tag{1.14}$$

Therefore, thanks to the decoherence process (1.14) we obtain the desired irreversible transformation (1.3) in a quantum measurement. Note that this process gets rid of the quantum correlations between $\sf S$ and $\sf M$ contained in the off-diagonal ($i \neq j$) terms in (1.12).

1.2.5 Decoherence

Before presenting a careful analysis of the dynamics of $\sf SM$ given by (1.9) and (1.10), which will be done in Sec. 1.3.3, let us explain the decoherence process (1.14) by means of a simplified textbook-like model. In this model, one keeps only the apparatus-reservoir coupling $H_{\rm int}^{\sf MR}$ in the total Hamiltonian (1.10) [276]. We take

$$H_{\rm int}^{\sf MR} = X \otimes B \quad , \quad B = N^{-\frac{1}{2}} \sum_{\nu=1}^N B_\nu \, , \tag{1.15}$$

where B_ν are operators acting on single degrees of freedom[15] of $\sf R$. The initial state of $\sf SMR$ is $\rho_{\sf SMR}(0) = |\Psi_{\sf SM}^{\rm ent}\rangle\langle\Psi_{\sf SM}^{\rm ent}| \otimes \rho_{\sf R}$, where $|\Psi_{\sf SM}^{\rm ent}\rangle$ is the system-apparatus entangled state (1.13) and $\rho_{\sf R}$ is the reservoir initial state. We assume for simplicity that $\langle B_\mu\rangle_{\sf R} = {\rm tr}(B_\mu \rho_{\sf R}) = 0$. For any i, the modulus square $|\langle x|\mu_i^{\rm eq}\rangle|^2$ of the apparatus wave function is assumed to present a sharp peak at the value x_i. The reduced system-apparatus state at time τ then reads

$$\rho_{\sf SM}(\tau) = {\rm tr}_{\sf R}[\rho_{\sf SMR}(\tau)] \simeq \sum_{i,j} c_i c_j^* K_{ij}(\tau) |\alpha_i\rangle\langle\alpha_j| \otimes |\mu_{{\sf M}|i}^{\rm eq}\rangle\langle\mu_{{\sf M}|j}^{\rm eq}| \tag{1.16}$$

[14] Actually, under these assumptions the vectors $|\mu_i^{\rm eq}\rangle$ are not all collinear and the reduced apparatus state ${\rm tr}_{\sf S} |\Psi_{\sf SM}^{\rm ent}\rangle\langle\Psi_{\sf SM}^{\rm ent}| = \sum_i |c_i|^2 |\mu_i^{\rm eq}\rangle\langle\mu_i^{\rm eq}|$ has a non-zero entropy, which means that $|\Psi_{\sf SM}^{\rm ent}\rangle$ is entangled (see Sec. 2.3).

[15] One may think of B_μ as self-adjoint linear combinations of creation and annihilation operators of a photon in mode ν.

with the decoherence factor $K_{ij}(\tau) = \langle \mathrm{e}^{-\mathrm{i}\tau(x_i - x_j)B} \rangle_{\mathsf{R}}$. Since the operators B_μ act on distinct degrees of freedom, they commute between themselves and can be considered as independent random variables. The central limit theorem implies that for $N \gg 1$,

$$K_{ij}(\tau) = \exp\left\{ -\frac{\tau^2}{\tau_{\rm dec}(x_i,x_j)^2} \right\} \quad , \quad \tau_{\rm dec}(x_i,x_j) = \frac{1}{\langle B^2 \rangle_{\mathsf{R}}^{\frac{1}{2}} |x_i - x_j|} \tag{1.17}$$

(here $\langle B^2 \rangle_{\mathsf{R}}$ is the limit of $N^{-1} \sum_\nu \langle B_\nu^2 \rangle_{\mathsf{R}}$). Defining the decoherence time $\tau_{\rm dec} = \max_{i \neq j} \tau_{\rm dec}(x_i, x_j)$, one concludes that the system and apparatus are approximately in the statistical mixture given in the right-hand side of (1.14) at times $\tau \gtrsim \tau_{\rm dec}$.

The Gaussian decay of $K_{ij}(\tau)$ is characteristic of the universal short time regime [43] in which $\tau_{\rm dec}$ is much smaller than the time scales of evolution under the respective Hamiltonians of SM and R. It is worth noting that the decoherence time decreases with the distance $d = |x_i - x_j|$ separating the components of the superposition (1.13). This is a generic feature occurring in most decoherence processes, albeit it is not a general one[16]. Let us also point out that for more realistic translation invariant apparatus-reservoir couplings[17], the decoherence time is found to decrease like d^{-2} at small distances and to saturate to a finite value $\tau_{\rm dec}(\infty) > 0$ at large distances [97, 126]. However, even in this case, for typical system-reservoir coupling strengths, $\tau_{\rm dec}$ is ridiculously small when d reaches a macroscopic magnitude[18]. It thus makes sense when dealing with macroscopic superpositions to neglect the Hamiltonians H_{SM} and H_{R}, which generate state evolutions on larger time scales than $\tau_{\rm dec}$. This justifies a posteriori that one can retain only the interaction between SM and R in the total Hamiltonian [43].

We emphasize that the decoherence factor $K_{ij}(\tau)$ quantifies in some sense the overlap between the time-evolved conditional states $\rho_{\mathsf{R}|i}(\tau)$ and $\rho_{\mathsf{R}|j}(\tau)$ of the reservoir, with $\rho_{\mathsf{R}|i}(\tau) = \mathrm{e}^{-\mathrm{i}\tau x_i B} \rho_{\mathsf{R}} \mathrm{e}^{\mathrm{i}\tau x_i B}$. This is more clearly seen when R is initially in a pure state $|\phi_{\mathsf{R}}\rangle$. In this (unrealistic) case, the three-partite system SMR is at time τ in the pure state

$$|\Psi^{\rm ent}_{\mathsf{SMR}}(\tau)\rangle = \sum_i c_i |\Psi_{\mathsf{SMR}|i}(\tau)\rangle \quad , \quad |\Psi_{\mathsf{SMR}|i}(\tau)\rangle = |\alpha_i\rangle |\mu^{\rm eq}_{\mathsf{M}|i}\rangle |\phi_{\mathsf{R}|i}(\tau)\rangle \tag{1.18}$$

with $|\phi_{\mathsf{R}|i}(\tau)\rangle = \mathrm{e}^{-\mathrm{i}\tau x_i B} |\phi_{\mathsf{R}}\rangle$. Then $K_{ij}(t)$ coincides with the scalar product $\langle \phi_{\mathsf{R}|i}(\tau) | \phi_{\mathsf{R}|j}(\tau) \rangle$. Consequently, decoherence comes from the establishment of perfect correlations between the components of the superposition (1.13) and nearly orthogonal states $\rho_{\mathsf{R}|i}(\tau)$ of the reservoir at times $\tau \gtrsim \tau_{\rm dec}$. This process is analogous to a measurement in which the state of SM would be measured by the reservoir, as evidenced by the formal similarity between (1.13) and (1.18) and the almost orthogonality of the $\rho_{\mathsf{R}|i}(\tau)$ as in (1.2). Such an analogy is only formal, because the aforementioned orthogonality is due to intricated phases in the large Hilbert space of R. The states $\rho_{\mathsf{R}|i}(\tau)$ are not associated to macroscopic distinct behaviors of the reservoir, hence a readout of the outcome is impossible. Like in a perfect measurement, each component of the superposition is left invariant by the decoherence process ($\rho_{\mathsf{SM}}(\tau) \simeq \rho_{\mathsf{SM}}(0)$ if all c_i but one vanish), meaning that the apparatus states $|\mu^{\rm eq}_{\mathsf{M}|i}\rangle$ are robust against the coupling with R via the Hamiltonian (1.15). Using the analogy between decoherence and measurements, Zurek [275] has proposed to call "pointer states" the states of an open quantum system which are weakly affected by the dynamics in the time interval $[0, \tau_{\rm dec}]$, whereas in this time interval superpositions of such states are strongly affected and decohere into statistical mixtures[19].

It is often claimed that the pointer states are determined by the coupling with the reservoir. This statement is not true in general. For instance, as it has been shown in [233], if one replaces the position operator X by the momentum P in the Hamiltonian (1.15), the macroscopic superposition $|\Psi^{\rm ent}_{\mathsf{SM}}\rangle$ evolves at short times to the same state $\rho_{\mathsf{SM}}(\tau)$ as in (1.16), albeit with a decoherence factor $K_{ij}(\tau)$ decaying like $\mathrm{e}^{-(\tau/\tau_{\rm dec})^3}$ and a larger decoherence time $\tau_{\rm dec}$. This comes from the apparatus Hamiltonian H_{M}, which transforms the off-diagonal contributions involving states with macroscopically distinct positions into terms involving states with distinct momenta, the latter being washed out rapidly by the coupling $H^{\mathsf{MR}}_{\rm int} = gP \otimes B$ (the diagonal contributions are also affected but at longer time scales). Therefore, the pointer states $|\mu^{\rm eq}_{\mathsf{M}|i}\rangle$ of the apparatus are determined by the apparatus Hamiltonian H_{M} rather than by the precise form of the apparatus-bath coupling. This is

[16] We will see in chapter 3 an example of decoherence process having an impact on a superposition of coherent states which is independent of the distance between these coherent states.

[17] Such translation-invariant coupling Hamiltonians can be approximated by (1.15) for distances d smaller than the mean wavelength of the photons/phonons in the reservoir (dipole approximation).

[18] This is due to the factor $\hbar$ in the numerator of the expression (1.17) of $\tau_{\rm dec}$, which is here (and in all what follows) set to unity!

[19] This terminology has an insightful physical meaning but is not mathematically precise and should be used with some care. In fact, the pointer basis is ambiguously defined when both the system Hamiltonian and the system-reservoir coupling play a role in the dynamics (i.e., when $\tau_{\rm dec}$ is of the same order of magnitude than the time scale of the free evolution of SM).

consistent with condition 3 of Sec. 1.2.2, according to which the final states of M should be equilibrium states. In contrast, in view of (1.16) the system state becomes $\rho_{\mathsf{S}}(\tau) \simeq \sum_i |c_i|^2 |\alpha_i\rangle\langle\alpha_i|$ at time $\tau \gg \tau_{\rm dec}$, independently of the system Hamiltonian H_{S}. The basis in which this state is diagonal is selected by the system-apparatus coupling $H_{\rm int}^{\mathsf{SM}}$ satisfying condition 4 above.

Let us point out that the central limit theorem implies a stronger result than the Gaussian decay of the decoherence factor $K_{ij}(\tau)$. Namely, the expectation value $\langle \mathrm{e}^{-\mathrm{i}\tau(x_i-x_j)B} O\rangle_{\mathsf{R}}$ also decays like a Gaussian with the rate $\tau_{\rm dec}^{-1}$ for any local observable[20] O of the three-partite system SMR.

We end this subsection by some comments related to the point 2) raised in Sec. 1.2.2. Although it is easy to obtain the system-apparatus density matrix $\rho_{\mathsf{SM}}(t_{\rm meas})$ in the right-hand side of (1.14) from the total microscopic state $\rho_{\mathsf{SMR}}(t_{\rm meas})$ by tracing out the microscopic degrees of freedom of the reservoir, it is not clear how to associate a state ensemble to $\rho_{\mathsf{SM}}(\tau)$. Due to the vanishing of $\langle \mathrm{e}^{-\mathrm{i}\tau(x_i-x_j)B} O\rangle_{\mathsf{R}}$ for any local observable O of SMR, it is impossible in practice to distinguish at times $\tau \gtrsim \tau_{\rm dec}$ the entangled state (1.18) from an incoherent mixture of the states $|\Psi_{\mathsf{SMR}|i}(\tau)\rangle$ with probabilities $|c_i|^2$. For indeed, the expectation value of any local observable O will be the same in the two states and measurements of O give no information about the terms $c_i c_j^* |\Psi_{\mathsf{SMR}|i}(\tau)\rangle\langle\Psi_{\mathsf{SMR}|j}(\tau)|$ for $i \neq j$ in the expansion of $|\Psi^{\rm ent}_{\mathsf{SMR}}(\tau)\rangle\langle\Psi^{\rm ent}_{\mathsf{SMR}}(\tau)|$. However, the identification of $\rho_{\mathsf{SM}}(\tau)$ with a statistical mixture amounts to replace a linear superposition in a much larger space by classical probabilities, i.e., to ignore quantum correlations that cannot be measured but nevertheless exist. We thus have to face a much more subtle situation than in classical statistical physics, where one usually ignores some microscopic degrees of freedom without being obliged to ignore at the same time some *fundamental* correlations[21]. Even by invoking ignorance about inaccessible quantum correlations, it is not clear how one could replace them by classical probabilities and say that SM is in one of the states $|\alpha_i\rangle|\mu^{\rm eq}_{\mathsf{M}|i}\rangle$ with probability $|c_i|^2$.

1.2.6 Example: the Stern-Gerlach apparatus

To illustrate the above considerations, it is instructive to consider a simple textbook example of quantum measurement in which the measured system is the spin degree of freedom of a moving atom with spin one half, the measured observable $S = \sigma_z/2$ is the spin along (Oz), and the pointer variable X is the atomic position [40]. The atom moves on the line $y = z = 0$ and crosses a magnet with an inhomogeneous magnetic field. This field is approximated in the vicinity of the atom trajectory by $\mathbf{B}(z) \simeq (B_z(0) + \partial_z B_z(0) z)\,\mathbf{e}_z$. The corresponding Stern-Gerlach apparatus is represented in Fig. 1.2. The system-apparatus coupling Hamiltonian has the form (1.11),

$$H_{\rm int}^{\mathsf{SM}} = \mu_B \partial_z B_z(0) S \otimes Z \;, \tag{1.19}$$

where Z is the atomic position operator along (Oz) and μ_B the Bohr magneton. The constant part in $\mathbf{B}(z)$ contributes to the system Hamiltonian $H_{\mathsf{S}} = \mu_B B_z(0) S$, which commutes with S since we ignore the x and y-components of the magnetic field. In a first approximation, the atom moves freely in the vacuum. The apparatus Hamiltonian is $H_{\mathsf{M}} = P^2/(2M)$, where M is the atomic mass. The reservoir is the screen and eventually the scatterers in the imperfect vacuum between the magnet and the screen. Even if the apparatus is just composed of the orbital atomic degree of freedom and is therefore not macroscopic, its positions on the screen correlated to the spin up and down may be separated by large distances, provided that the distance L between the magnet and the screen is much larger than the length ℓ of the magnet. If one includes the screen, a macroscopic apparatus is obtained. Before the entrance in the magnet, the atom is in an arbitrary spin linear superposition $|\psi(0)\rangle = c_\uparrow|\uparrow\rangle + c_\downarrow|\downarrow\rangle$ of eigenstates of σ_z and its wave packet $|\mu(0)\rangle$ has a sharply defined momentum $\mathbf{p}(0) = p_x \mathbf{e}_x$ in the x direction. We denote by Δp_x, Δp_y and Δp_z the momentum uncertainties. We neglect spin-orbit coupling, so that spin and position are initially uncorrelated and the atom is initially in the state (1.1).

The time spent in the magnet is $t_{\rm int} = M\ell/p_x$. The crossing of the magnet entangles the atomic spin and position,

$$(c_\uparrow|\uparrow\rangle + c_\downarrow|\downarrow\rangle)|\mu(0)\rangle \;\longrightarrow\; c_\uparrow|\uparrow\rangle|\mu_+(t_{\rm int})\rangle + c_\downarrow|\downarrow\rangle|\mu_-(t_{\rm int})\rangle \;, \tag{1.20}$$

where $|\mu_\pm(t_{\rm int})\rangle = e^{\mp \mathrm{i}\mu_B \partial_z B_z(0) Z t_{\rm int}/2}|\mu(0)\rangle$ is a wave packet with sharply defined momentum $\mathbf{p}_\pm = p_x \mathbf{e}_x \pm \mu_B \partial_z B_z(0) t_{\rm int}\,\mathbf{e}_z/2$. At the exit of the magnet, the two wave packets $|\mu_\pm(t)\rangle$ separate from each other as

[20] By local observable we mean here an observable which does not act on a part of R containing infinitely many of degrees of freedom.

[21] Let us quote H.D. Zeh in this respect: "Identifying the [system-apparatus] superposition with an ensemble of states (represented by a statistical operator ρ) which merely leads to the same *expectation values* $\langle O\rangle = \mathrm{tr}(O\rho)$ for an *axiomatically limited set of observables* O (such as local ones) would obviously beg the question. This insufficient argument is nonetheless found widely in the literature (cf Haag 1992). It would be equivalent to a quantum mechanical state space smaller than required by a general superposition principle." (opus cit. from [102], chapter 2). See also [277] for a different perspective.

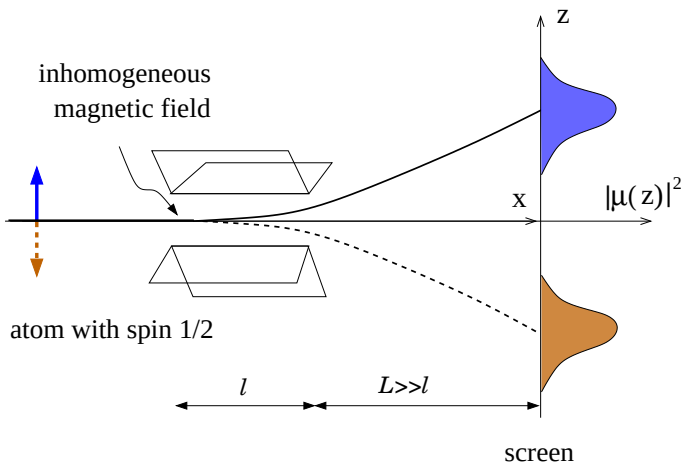

Figure 1.2: *Stern-Gerlach apparatus for measuring the spin* 1/2 *of an atom.*

they propagate. If one assumes a perfect vacuum, the state of the atom just before it hits the screen (time $t_{\rm meas} \simeq ML/p_x \gg t_{\rm int}$) is given by (1.13) with $|\mu_\pm^{\rm eq}\rangle$ having mean position $z_\pm = \pm\mu_B \partial_z B_z(0) t_{\rm int} t_{\rm meas}/(2M) = \pm d/2$. According to conditions 1-3 of Sec. 1.2.1, the distance d between the two wave packet centers should be macroscopic. Moreover, the position uncertainty Δz of the wave packets should be much smaller than d. Taking into account the spreading of each wave packet, it is easy to show that this is the case when [40]

$$\Delta p_z \ll \mu_B \partial_z B_z(0) t_{\rm int} \ . \tag{1.21}$$

This condition means that the peaks in momentum of the two wave packets can be resolved at the exit the magnet. The motion of the centers of the wave packets after the exit of the magnet is well described by classical mechanics.

The last stage of the measurement is the decoherence process. As a result of the interaction of the atom with the molecules in the screen (or with any device measuring the position of the atom), all information about the coherences between the two wave packets is transferred to the screen degrees of freedom and irremediably lost. The linear superposition $|\Psi_{\sf SM}^{\rm ent}\rangle$ is then transformed into the density matrix $\rho_{\sf SM}(t_{\rm meas})$ as in (1.3). Let us stress again that the two states are quite different. For an atom in state $|\Psi_{\sf SM}^{\rm ent}\rangle$, one could in principle reconstruct a localized wave packet and the initial spin state $|\psi(0)\rangle$ by recombining the two beams. For instance, if $c_\uparrow = c_\downarrow = 1/\sqrt{2}$ this would lead to an eigenstate of σ_x with eigenvalue 1. In contrast, for an atom in the statistical mixture $\rho_{\sf SM}(t_{\rm meas})$, the same state would be obtained by recombining the two beams. This state has a vanishing expectation of σ_x whatever the values of $c_\uparrow$ and $c_\downarrow$ are.

1.3 Determination of the time scales in a measurement process

1.3.1 Our model of measuring apparatus

In the dynamical analysis of Secs. 1.2.4-1.2.6, we have made the unrealistic assumption that the system-apparatus and apparatus-reservoir couplings are switched on and off one after the other. We would like now to study what happens when both couplings are present simultaneously between times $t = 0$ and $t_{\rm int}$. In this situation, the system-apparatus interaction does not produce macroscopic superpositions such as (1.13), because the superposition is transformed into a statistical mixture simultaneously as it is produced.

We consider a specific model of measuring apparatus M, designed to measure a given observable S with discrete spectrum of the system S. As in Sec. 1.2.4, we identify M with the pointer, assumed here to have a single degree of freedom with associated Hilbert space $\mathcal{H}_{\mathsf{M}} = L^2(\mathbb{R})$. All the other degrees of freedom of the measuring device are included in the reservoir R, which has Hilbert space $\mathcal{H}_{\mathsf{R}} = \otimes_{\nu=1}^N \mathcal{H}_\nu$, $\mathcal{H}_\nu$ being the Hilbert space associated to the νth degree of freedom. All formulas below are valid in the limit $N \gg 1$, taken before any other limit (such as large time limits) as it is the rule in statistical physics. The pointer observable is the position X, acting as a multiplication operator in $\mathcal{H}_{\mathsf{M}}$. The pointer Hamiltonian reads

$$H_{\mathsf{M}} = \frac{P^2}{2M} + V(X) \ , \tag{1.22}$$

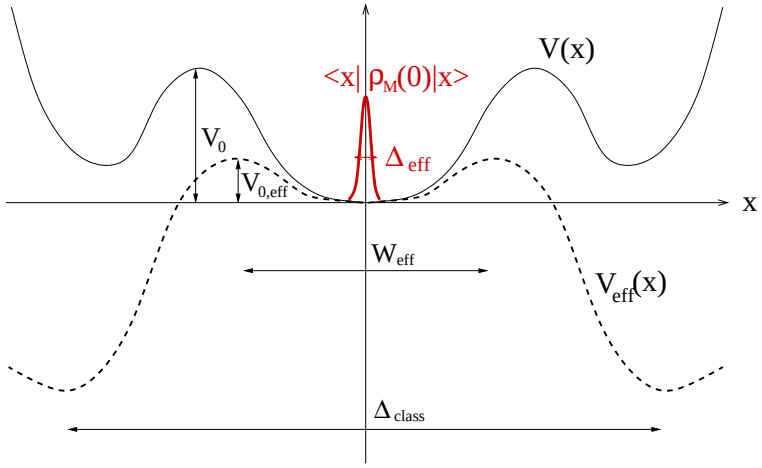

Figure 1.3: Pointer potential $V(x)$ (plain line) and effective potential $V_{\rm eff}(x)$ (broken lines) in arbitrary units. The widths W and $W_{\rm eff}$ of the potential barriers around $x = 0$ of the two potentials are much larger than the width $\Delta_{\rm eff} \approx \Delta_{\rm th}$ of the density of pointer position density $\langle x|\rho_P(0)|x\rangle$, represented in green in the figure.

where P is the momentum canonically conjugated to X and M the pointer mass. The potential $V(x)$ is even and has a local minimum at $x = 0$, i.e., $V'(0) = 0$ and $V''(0) > 0$. The height of the two potential barriers surrounding this minimum is much larger than the thermal energy β^{-1}, in such a way that M has a well-defined rest state at $x = 0$ (see Fig. 1.3). The system-apparatus coupling is of the form (1.11). It is chosen such that (i) it does not change the measured observable S (condition 4 of Sec. 1.2.4); (ii) it is capable of shifting the pointer position by an amount proportional to S, so as to tie up each eigenvalue s_i of S with a specific pointer position; (iii) it involves a large coupling constant g, so that different eigenvalues end up in one-to-one correspondence with pointer states separated by large distances.

Finally, we choose the apparatus-reservoir coupling[22] of the form (1.15), with operators B_ν acting on the Hilbert space $\mathcal{H}_\nu$ and satisfying $\mathrm{tr}(B_\nu \rho_{\mathsf{R}}) = 0$. The additivity of B in contributions B_ν acting on single degrees of freedom will enable us to invoke the quantum central limit theorem (QCLT) [183, 247]. We make no specific assumption on the reservoir Hamiltonian H_{R} excepted that there should be no long-range correlations in the free reservoir Gibbs state $\rho_{\mathsf{R}} = Z_{\mathsf{R}}^{-1} e^{-\beta H_{\mathsf{R}}}$ (more precisely, $\mathrm{tr}(B_\mu B_\nu \rho_{\mathsf{R}})$ must decay to zero faster than $1/|\mu - \nu|$ for $|\mu - \nu| \gg 1$ [247]). Such long-range correlations would invalidate the QCLT, which basically shows that R behaves as a collection of harmonic oscillators linearly coupled to M in the limit $N \to \infty$. This implies Gaussian statistics (Wick theorem) for the time-correlation functions associated to B with respect to ρ_{R}. The $2n$-point correlation functions are then fully determined in terms of the two-point correlator

$$K_{\mathsf{R}}(t) = \langle B(t)B\rangle_{\mathsf{R}} = \mathrm{tr}\big(B(t)B\rho_{\mathsf{R}}\big) \quad , \quad B(t) = e^{\mathrm{i}H_{\mathsf{R}}t} B e^{-\mathrm{i}H_{\mathsf{R}}t} \, , \tag{1.23}$$

and the $(2n+1)$-point correlation functions vanish.

We shall make use of a few general properties and assumptions on the correlator (1.23). Firstly, $K_{\mathsf{R}}(t)$ and its real part $\mathrm{Re}\, K_{\mathsf{R}}(t)$ are of positive type, i.e., they have non-negative Fourier transforms. Since $K_{\mathsf{R}}(t)^* = K_{\mathsf{R}}(-t)$, the Fourier transforms $\widehat{\mathrm{Re}\, K_{\mathsf{R}}}(\omega)$ and $(\widehat{\mathrm{Im}\, K_{\mathsf{R}}})(\omega)$ of the real and imaginary parts of $K_{\mathsf{R}}(t)$ are even and odd functions of ω, respectively. We assume that these functions admit derivatives of sufficiently high orders, in such a way that $K(t)$ decays rapidly to zero as $t \to \infty$. Furthermore, we suppose that $(\widehat{\mathrm{Im}\, K_{\mathsf{R}}})(\omega) \sim -\mathrm{i}\,\widehat{\gamma}\,\omega^m$ as $\omega \to 0$ with $m = 1, 3, \ldots$ a positive odd integer and $\widehat{\gamma} > 0$. By analogy with the case of a reservoir of harmonic oscillators linearly coupled to M, we speak of *Ohmic damping* when $m = 1$ and of *super-Ohmic damping* when $m > 1$ [256]. The small frequency behavior $(\widehat{\mathrm{Re}\, K_{\mathsf{R}}})(\omega) \sim 2\,\widehat{\gamma}\,\omega^{m-1}/\beta$ of the Fourier transform of $\mathrm{Re}\, K_{\mathsf{R}}(t)$ can

[22] For a translation-invariant apparatus, a more physical coupling Hamiltonian would be

$$H_{\rm int}^{\mathsf{MR}} = N^{-\frac{1}{2}} \sum_q e^{iqX} \otimes (B_q + B_{-q}^\dagger)$$

with q the momentum of the mode q of the reservoir and the operators B_q acting on this mode transformed as $B_q \to B_q e^{\mathrm{i}qa}$ under space translations by distances a. For small separations between the pointer positions, this Hamiltonian can be approximated by the Hamiltonian (1.15) which is not translation invariant but is easier to handle (see also the discussion in Sec. 1.2.5). Our assumptions on the apparatus-reservoir interaction are here dictated by technical rather than by fundamental reasons. A straightforward generalization of the foregoing results to couplings of the form $H_{\rm int}^{\mathsf{MR}} = X^\alpha \otimes B$ with $\alpha \in \mathbb{N}^\star$ is given in [224].

be deduced from that of $\mathrm{Im}\, K_{\mathsf{R}}(t)$ thanks to the Kubo-Martin-Schwinger (KMS) relation

$$(\widehat{\mathrm{Re}\, K_{\mathsf{R}}})(\omega) = \mathrm{i}\frac{(\widehat{\mathrm{Im}\, K_{\mathsf{R}}})(\omega)}{\tanh(\beta\omega/2)} \,. \tag{1.24}$$

The KMS relation holds for all frequencies. It follows from the properties of the Gibbs state ρ_{R} [42] entering in the correlator (1.23).

Before going further, it is important to identify the different time scales of the model described above:

(1) the characteristic time T_{M} for the motion of the pointer under its Hamiltonian H_{M}, defined as the period $T_{\mathsf{M}} = 2\pi(M/V''(0))^{1/2}$ of oscillations around the minimum of the potential $V(x)$;

(2) the characteristic time[23] T_{S} for the evolution of the measured observable S under the Hamiltonian H_{S};

(3) the system-pointer interaction time t_{int};

(4) the decoherence time τ_{dec} (to be estimated below);

(5) the reservoir correlation time τ_{R}, thermal time β, and characteristic time t_{R} of evolution of B under the reservoir Hamiltonian H_{R}. The first and last times are respectively the smallest time such that $K_{\mathsf{R}}(t) \simeq 0$ for $t \gtrsim \tau_{\mathsf{R}}$ and the largest time such that $K_{\mathsf{R}}(t) \simeq K(0)$ for $|t| \lesssim t_{\mathsf{R}}$. Note that $t_{\mathsf{R}} \leq \beta \leq \tau_{\mathsf{R}}$, with equality $\tau_{\mathsf{R}} = \beta$ at low enough temperature.

As discussed in Sec. 1.2.4, in an ideal measurement the free dynamics of S remains ineffective on the measured observable S during the measurement. Furthermore, the pointer time T_{M} is a classical time, which should be much larger than all quantum time scales in the model. We thus assume the following separation of time scales

$$\tau_{\mathrm{dec}}, t_{\mathrm{int}} \ll T_{\mathsf{S}} \qquad , \qquad \tau_{\mathrm{dec}}, t_{\mathrm{int}}, \beta \ll T_{\mathsf{M}} \,. \tag{1.25}$$

1.3.2 Initial state

We start at time $t = 0$ with a system-apparatus-reservoir state

$$\rho_{\mathsf{SMR}}(0) = \rho_{\mathsf{S}}(0) \otimes \rho_{\mathsf{MR}}(0) \tag{1.26}$$

with no correlations between S and the apparatus and reservoir. In contrast, the two latter are initially correlated and in a metastable local thermal equilibrium[24] $\rho_{\mathsf{MR}}(0) = Z_{\mathsf{MR}}^{-1} e^{-\beta(H_{\mathsf{M}}+H_{\mathsf{R}}+H_{\mathrm{int}}^{\mathsf{MR}})}$ in which the apparatus is localized near $x = 0$. By invoking the high-temperature limit $\beta \ll T_{\mathsf{M}}$ and the Gaussian statistics of B (as implied by the QCLT), one finds that the initial reduced density matrix $\rho_{\mathsf{M}}(0) = \mathrm{tr}_{\mathsf{R}}[\rho_{\mathsf{MR}}(0)]$ of the pointer is given in the position representation by (see below)

$$\langle x|\rho_{\mathsf{M}}(0)|x'\rangle \propto \mathrm{e}^{-\beta(V_{\mathrm{eff}}(x)+V_{\mathrm{eff}}(x'))/2}\, \mathrm{e}^{-2\pi^2(x-x')^2/\lambda_{\mathrm{th}}^2} \,, \tag{1.27}$$

where $\lambda_{\mathrm{th}} = 2\pi(\beta/M)^{1/2}$ is the thermal de Broglie wavelength. The pointer potential appears renormalized by the apparatus-reservoir interaction as

$$V_{\mathrm{eff}}(x) = V(x) - \gamma_0 x^2 \,, \qquad \gamma_0 = \int_{-\infty}^{0} \mathrm{d}t\, \mathrm{Im}\, K_{\mathsf{R}}(t) \;\geq 0 \,. \tag{1.28}$$

For local stability of the apparatus, the apparatus-reservoir coupling must be weak enough so that $V''_{\mathrm{eff}}(0) > 0$; we even bound the latter curvature finitely away from zero by, say, $V''_{\mathrm{eff}}(0) > V''(0)/2$, i.e.,

$$\gamma_0 < \frac{1}{4} V''(0) \,. \tag{1.29}$$

This makes sure that the initial density of pointer positions

$$\langle x|\rho_{\mathsf{M}}(0)|x\rangle \propto \exp\left(-\frac{x^2}{2\Delta_{\mathrm{eff}}^2}\right) \quad , \quad |x| \lesssim \Delta_{\mathrm{eff}} \tag{1.30}$$

[23] This time is by definition the largest time t such that $\mathrm{tr}[S^{(0)}(t_1)\dots S^{(0)}(t_n)\rho_{\mathsf{S}}(0)] \simeq \mathrm{tr}[S^n \rho_{\mathsf{S}}(0)]$ for $|t_1|, \dots, |t_n| \leq t$, with $S^{(0)}(t) = e^{\mathrm{i}H_{\mathsf{S}}t} S e^{-\mathrm{i}H_{\mathsf{S}}t}$.

[24] This formula for $\rho_{\mathsf{MR}}(0)$ should not be taken too seriously. In fact $\rho_{\mathsf{MR}}(0)$ is not a true equilibrium and the Hamiltonian $H_{\mathsf{M}} + H_{\mathsf{R}} + H_{\mathrm{int}}^{\mathsf{MR}}$ is unbounded from below.

has a single peak at $x = 0$ with a renormalized width $\Delta_{\rm eff} = (\beta V''_{\rm eff}(0))^{-1/2}$ of the order of the bare thermal fluctuation $\Delta_{\rm th} = (\beta V''(0))^{-1/2}$. By virtue of the inequality $\gamma_0 \leq \beta K_{\mathsf{R}}(0)/2$ (which follows from the KMS relation (1.24), see [224], Appendix C), one finds that the stability condition (1.29) is fulfilled when $\Delta_{\rm th}^2 K_{\mathsf{R}}(0)\beta^2 < 1/2$.

Let V_0 and $W \approx (V_0/V''(0))^{1/2}$ be the height and width of the two potential barriers of $V(x)$ surrounding the local minimum at $x = 0$. If $V(x) = {\rm o}(x^2)$ at large distances $|x| \gtrsim W$, the effective potential $V_{\rm eff}(x)$ is unstable. The pointer initial state (1.27) is then a metastable local thermal equilibrium. In order to be able to prepare the apparatus in such a local equilibrium, the height $V_{0,\rm eff}$ of the potential barriers of the effective potential $V_{\rm eff}(x)$ around $x = 0$ must be large compared with the thermal energy[25]. Thanks to (1.29), this is the case provided that the bare potential $V(x)$ satisfies the same requirement[26], i.e., $V_0 \gg \beta^{-1}$. Interestingly, $V(x)$ can be chosen such that the two potential barriers of $V_{\rm eff}(x)$ are separated by a mesoscopic distance $W_{\rm eff} \approx (V_{0,\rm eff}/V''_{\rm eff}(0))^{1/2} \gg \Delta_{\rm eff}$ which is small compared with the macroscopic readout scale $\Delta_{\rm class}$, as illustrated in Fig. 1.3. The system-apparatus interaction then just has to get the pointer out of the well, leaving the subsequent displacement growth to the action of the effective potential. The instability of this potential thus provides the amplification mechanism necessary to fulfill condition 3 of Sec. 1.2.1. For a macroscopic apparatus at high temperature ($\beta \ll T_{\mathsf{M}}$), the different length scales are ordered as

$$\lambda_{\rm th} \ll \Delta_{\rm th} \approx \Delta_{\rm eff} \ll W_{\rm eff} \ll \Delta_{\rm class} \ . \tag{1.31}$$

To fix ideas, for $T_{\mathsf{M}} = 1\,{\rm s}$, $M = 1\,{\rm g}$, and a temperature of 1 K one has $\lambda_{\rm th} \approx 10^{-21}\,{\rm m}$ and $\Delta_{\rm th} \approx 10^{-11}\,{\rm m}$.

1.3.3 Dynamics

We now proceed to determine the time-evolved density matrix (1.9) of the system and apparatus. At time $t \ll T_{\mathsf{S}}, T_{\mathsf{M}}$, the kernel of this matrix is found to be

$$\langle x|\rho_{\mathsf{SM}}(t)|x'\rangle = \sum_{i,j} \Pi_i \rho_{\mathsf{S}}(0) \Pi_j \langle x_i(t)|\rho_{\mathsf{M}}(0)|x'_j(t)\rangle \ {\rm e}^{-D_t(x_i(t),x'_j(t);s_i,s_j)} {\rm e}^{-{\rm i}\phi_t} \tag{1.32}$$

with a non-negative decoherence exponent D_t given by

$$D_t(x,x';s_i,s_j) = \frac{1}{2}\int_0^t {\rm d}\tau_1 \int_0^t {\rm d}\tau_2 \, {\rm Re}\, K_{\mathsf{R}}(\tau_1-\tau_2)\big(x'_j(-\tau_1) - x_i(-\tau_1)\big)\big(x'_j(-\tau_2) - x_i(-\tau_2)\big) \ \geq \ 0 \tag{1.33}$$

and a real phase ϕ_t irrelevant for decoherence (the non-negativity of D_t follows trivially from the non-negativity of the Fourier transform of ${\rm Re}\, K_{\mathsf{R}}(t)$). The quantities $x_i(t)$ and $x'_j(t)$ in (1.32) and (1.33) are the shifted pointer positions

$$x_i(t) = x - g s_i t \quad , \quad x'_j(t) = x' - g s_j t \ . \tag{1.34}$$

Entanglement and decoherence contribute separately in this remarkably simple system-apparatus state; they lead respectively to the shifted matrix element and the first exponential in (1.32).

It might come as a surprise that such an explicit formula can be derived for $\rho_{\mathsf{SM}}(t)$ in spite of the complicated form of the total Hamiltonian (1.10) of the three-partite system SMR. Moreover, unlike in usual approaches based on master equations, our results (1.32)-(1.33) are not perturbative in the apparatus-reservoir interaction[27]. Their derivation takes advantage of the separation of time scales (1.25), which justifies the use of the following approximations:

(i) the time-evolution operator ${\rm e}^{-{\rm i}tH_{\mathsf{SMR}}}$ at time $t \ll T_{\mathsf{M}}, T_{\mathsf{S}}$ can be replaced by $U(t)\,{\rm e}^{-{\rm i}t(H_{\mathsf{S}}+H_{\mathsf{M}})}$, with

$$U(t) = {\rm e}^{-{\rm i}t(H_{\mathsf{R}}+H_{\rm int}^{\mathsf{SM}}+H_{\rm int}^{\mathsf{MR}})} = {\rm e}^{-{\rm i}tH_{\mathsf{R}}} {\rm e}^{-{\rm i}gtS\otimes P} \mathcal{T} \exp\left\{-{\rm i}\int_0^t {\rm d}\tau (X + g\tau S)\otimes B(\tau)\right\} \tag{1.35}$$

(the last equality is nothing but a Dyson expansion including all orders; in the right-hand side, $\mathcal{T}$ stands for time ordering and $B(t)$ is the reservoir coupling agent in the interaction picture, see (1.23));

[25] In practice, one may prepare M in some state localized near $x = 0$ at time $t = -t_0$ and then let it interact with R between $t = -t_0$ and $t = 0$. If the thermalization time $\tau_{\rm th}$ is small compared with the tunneling escape time $T_{\rm tun}$, one may choose t_0 larger than $\tau_{\rm th}$ but much smaller than $T_{\rm tun}$, so that M is still within the effective potential well when the measurement starts at $t = 0$.

[26] In fact, $V_0 \gg \beta^{-1}$ is equivalent to $\Delta_{\rm th} \ll W$. If $\Delta_{\rm th} \ll W$ then by (1.29) one has also $\Delta_{\rm eff} \leq \sqrt{2}\Delta_{\rm th} \ll W$. Substituting $V(x)$ by $V''(0)x^2/2$ into (1.28), one obtains for any x such that $\Delta_{\rm eff} \ll x \ll W$,

$$2V_{\rm eff}(x) \simeq V''(0)x^2 - 2\gamma_0 x^2 = V''_{\rm eff}(0)x^2 \ \gg \ \beta^{-1}$$

(we have taken here $V(0) = 0$). This shows that $V_{0,\rm eff} \gg \beta^{-1}$.

[27] A simpler derivation of (1.32) and (1.33) can be found in the perturbative regime by using the Redfield master equation, see (1c) in the publication list.

(ii) the pointer-reservoir initial state can be replaced when $\beta \ll T_{\mathsf{M}}$ by its high-temperature approximation

$$\rho_{\mathsf{MR}}(0) \simeq Z_{\mathsf{MR}}^{-1}\, \mathrm{e}^{-\frac{\beta}{2}H_{\mathsf{M}}}\, \mathrm{e}^{-\beta(H_{\mathsf{R}}+H_{\mathrm{int}}^{\mathsf{MR}})} \mathrm{e}^{-\frac{\beta}{2}H_{\mathsf{M}}} \tag{1.36}$$

and its evolution under H_{M} can be neglected, that is, $\mathrm{e}^{-\mathrm{i}\tau H_{\mathsf{M}}}\rho_{\mathsf{MR}}(0)\mathrm{e}^{\mathrm{i}\tau H_{\mathsf{M}}} \simeq \rho_{\mathsf{MR}}(0)$ for $\tau \ll T_{\mathsf{M}}$. The errors incurred are of the order of β^2/T_{M}^2 and $\tau\beta/T_{\mathsf{M}}^2$, as can be seen from the Baker-Campbell-Haussdorf formula.

Let us outline the main steps of the derivation of (1.32) (see [224] for more detail). Using the approximations (i) and (ii), we find after a standard calculation that (1.32) holds if one can prove that

$$\langle x|\rho_{\mathsf{M}}(0)|x'\rangle\, \mathrm{e}^{-D_t(x,x';s_i,s_j)}\mathrm{e}^{-\mathrm{i}\phi_t} = Z_{\mathsf{MR}}^{-1}\int_{-\infty}^{\infty}\mathrm{d}y\, \langle x|\mathrm{e}^{-\frac{\beta}{2}H_{\mathsf{M}}}|y\rangle\langle y|\mathrm{e}^{-\frac{\beta}{2}H_{\mathsf{M}}}|x'\rangle Z_{\mathsf{R},y}\big\langle U[x_j',t]^\dagger U[x_i,t]\big\rangle_{\mathsf{R},y}\,, \tag{1.37}$$

where $U[x_i,t] = \mathcal{T}\exp\{-\mathrm{i}\int_0^t \mathrm{d}\tau\, x_i(-\tau)B(\tau)\}$. The expectation $\langle\cdot\rangle_{\mathsf{R},y} = \mathrm{tr}(\cdot\rho_{\mathsf{R},y})$ is taken with respect to the thermal state $\rho_{\mathsf{R},y} = Z_{\mathsf{R},y}^{-1}\mathrm{e}^{-\beta(H_{\mathsf{R}}+yB)}$ of the reservoir coupled to the pointer localized at position y, and $Z_{\mathsf{R},y}$ denotes the corresponding partition function. The formula (1.37) is derived with the help of the bosonic Wick theorem for the correlation functions $\langle\delta B(t_1)\cdots\delta B(t_n)\rangle_{\mathsf{R},y}$ with $\delta B(t) = B(t) - \langle B(t)\rangle_{\mathsf{R},y}$, as implied by the QCLT for $N \gg 1$ (see Sec. 1.3.2). Actually, it turns out that Wick theorem is equivalent to (see [224], Appendix D)

$$\begin{aligned}\left\langle\left[\mathcal{T}\exp\left\{-\mathrm{i}\int_0^t \mathrm{d}\tau\, k(\tau)\, B(\tau)\right\}\right]^\dagger \mathcal{T}\exp\left\{-\mathrm{i}\int_0^t \mathrm{d}\tau\, l(\tau)\, B(\tau)\right\}\right\rangle_{\mathsf{R},y} &= \exp\left\{\mathrm{i}\int_0^t \mathrm{d}\tau\big(k(\tau)-l(\tau)\big)\langle B(\tau)\rangle_{\mathsf{R},y}\right\}\\ \times\exp\left\{-\int_0^t \mathrm{d}\tau_1\int_0^{\tau_1}\mathrm{d}\tau_2\big(k(\tau_1)-l(\tau_1)\big)\big(k(\tau_2)K_{\mathsf{R}}{}^*(\tau_1-\tau_2) - l(\tau_2)K_{\mathsf{R}}(\tau_1-\tau_2)\big)\right\},\end{aligned} \tag{1.38}$$

where $k(\tau)$ and $l(\tau)$ are two arbitrary smooth real-valued functions. Note that the exponential in the second line is independent of y. Using (1.38) for $\langle\cdot\rangle_{\mathsf{R},y} = \langle\cdot\rangle_{\mathsf{R}}$, $t = -\mathrm{i}\beta$, $k(\tau) = 0$, and $l(\tau) = y$, one gets[28] $Z_{\mathsf{R},y} = Z_{\mathsf{R}}\mathrm{e}^{\beta\gamma_0 y^2}$. By (1.38) again, one finds that the expectation $\langle U[x_j',t]^\dagger U[x_i,t]\rangle_{\mathsf{R},y}$ yields the decoherence factor $\mathrm{e}^{-D_t(x,x';s_i,s_j)}$ given by (1.33). It can be shown that under the stability condition (1.29), the y-dependent integral in the first line of (1.38) entails nothing but a correction of relative order $(\lambda_{\mathrm{th}}/\Delta_{\mathrm{eff}})^2$ to the decoherence exponent D_t. Plugging these results into the right-hand side of (1.37) and using the approximation $\mathrm{e}^{-\beta H_{\mathsf{M}}/2}\mathrm{e}^{\beta\gamma_0 X^2/2} \simeq \mathrm{e}^{\beta\gamma_0 X^2/2}\mathrm{e}^{-\beta H_{\mathsf{M}}/2} \simeq \mathrm{e}^{-\beta H_{\mathrm{eff}}/2}$ with $H_{\mathrm{eff}} = P^2/(2M) + V_{\mathrm{eff}}(X)$, this right-hand side is found to be equal to the left-hand side.

Note that the expression (1.27) for the initial pointer state $\rho_{\mathsf{M}}(0)$ also follows from (1.37), in which one uses the high-temperature approximation[29]

$$\langle x|\mathrm{e}^{-\beta H_{\mathsf{M}}/2}|y\rangle \simeq \mathrm{e}^{-\beta(V(x)+V(y))/4}\mathrm{e}^{-4\pi^2(x-y)^2/\lambda_{\mathrm{th}}^2}\,. \tag{1.39}$$

To get the result, one replaces $V(y)$ by $V''(0)y^2/2$ and neglects terms of the order of $(\lambda_{\mathrm{th}}/\Delta_{\mathrm{th}})^2$.

1.3.4 Entanglement time

Let us first look at the reduced pointer state. In view of (1.32) and (1.33), it is given by $\rho_{\mathsf{M}}(t) = \mathrm{tr}_{\mathsf{S}}[\rho_{\mathsf{SM}}(t)] = \sum_i p_i\rho_{\mathsf{M}|i}(t)$, where $p_i = \mathrm{tr}[\Pi_i\rho_{\mathsf{S}}(0)]$ is the probability of the measurement outcome i and

$$\langle x|\rho_{\mathsf{M}|i}(t)|x'\rangle = \langle x - gs_it|\rho_{\mathsf{M}}(0)|x' - gs_it\rangle\, \exp\left\{-\frac{(x-x')^2}{2}\int_0^t\mathrm{d}\tau_1\int_0^t\mathrm{d}\tau_2\,\mathrm{Re}\,K_{\mathsf{R}}(\tau_1-\tau_2)\right\}\mathrm{e}^{-\mathrm{i}\phi_t}\,. \tag{1.40}$$

The pointer conditional state $\rho_{\mathsf{M}|i}(t)$ has a kernel shifted by gs_it in position and modulated by an i-independent decoherence factor and a phase factor. One infers from (1.30) that the density $\langle x|\rho_{\mathsf{M}}(t)|x\rangle$ of pointer positions exhibits a succession of narrow peaks of width Δ_{eff} centered at $x = gs_it$, associated to each eigenvalue s_i with $p_i > 0$, as represented in Fig. 1.4 in the case of a spin one-half. The peaks associated to distinct eigenvalues s_i and s_j begin to be resolved at the *entanglement time*

$$\tau_{\mathrm{ent}}(s_i,s_j) = \frac{\Delta_{\mathrm{eff}}}{g|s_i - s_j|}\,. \tag{1.41}$$

[28] The fact that the coefficient γ_0 is given by (1.28) can be established by using the analyticity and KMS properties of $K_{\mathsf{R}}(t)$.

[29] One may recognize in this expression the well known short-time behavior of the quantum propagator $\langle x|\mathrm{e}^{-\mathrm{i}tH_{\mathsf{M}}}|x'\rangle$ for $t = -\mathrm{i}\beta/2$ (see e.g. [212]).

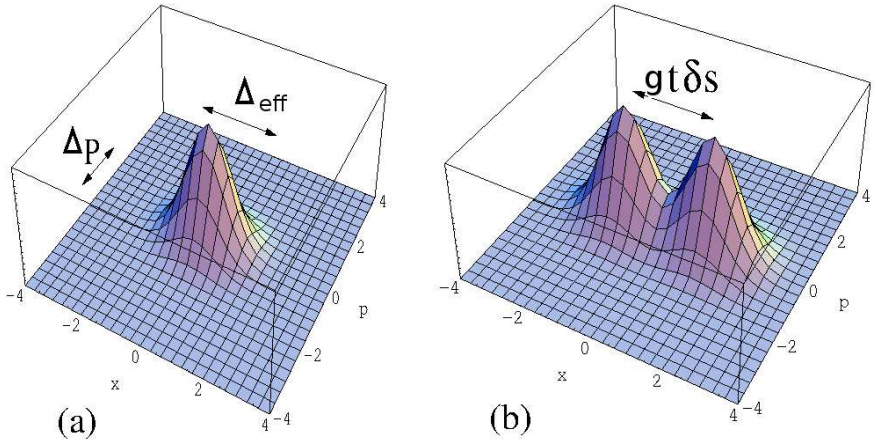

Figure 1.4: Wigner function (in arbitrary units) of the pointer reduced state (1.40) at times (a) $t = 0$ and (b) $t \approx \tau_{\rm ent}$. Here S is a spin one-half and $S = \sigma_z$. We have set the pointer-reservoir coupling to zero, so that $D_t = 0$.

Let δs be the minimum of $|s_i - s_j|$ over all pairs (s_i, s_j) of distinct eigenvalues of S present in the initial state $\rho_{\mathsf{S}}(0)$ (i.e., such that $p_i p_j > 0$)[30]. We denote by $\tau_{\rm ent} = \Delta_{\rm eff}/(g\delta s)$ the maximal entanglement time. At time $t \geq \tau_{\rm ent}$, neighboring peaks of the pointer density can be resolved. Each eigenvalue s_i of the measured observable S is then uniquely tied up with the pointer position gs_it.

Entanglement between S and M is encoded in the off-diagonal ($i \neq j$) contributions in the right-hand side of (1.32). In the absence of pointer-reservoir coupling (i.e., for $K_{\mathsf{R}}(t) = 0$) these off-diagonal contributions do not decay with time for $(x, x') = (gs_it, gs_jt)$: in fact, it follows from (1.30) that $\langle x = gs_it|\rho_{\mathsf{SM}}(t)|x' = gs_jt\rangle$ is approximately equal to $\Pi_i\rho_{\mathsf{S}}(0)\Pi_j\,\langle 0|\rho_{\mathsf{M}}|0\rangle$ at times $t \geq \tau_{\rm ent}$. At large times $t \geq t_{\rm class} = \Delta_{\rm clas}(g\delta s)^{-1}$, the system and pointer are in a superposition of states with macroscopically distinct pointer positions, that is, a Schrödinger cat state similar to (1.13). No classical probabilistic interpretation of the measurement is then possible. This illustrates again the central role played by the pointer-reservoir coupling in the measurement process. The fact that the reduced pointer state $\rho_{\mathsf{M}}(t) = \sum_i p_i\rho_{\mathsf{M}|i}(t)$ is a statistical mixture even in the absence of coupling to the reservoir does by no means solve this problem[31]. Similarly, the system reduced state

$$\rho_{\mathsf{S}}(t) = \sum_{i,j}\Pi_i\rho_{\mathsf{S}}(0)\Pi_j\int_{-\infty}^{\infty}{\rm d}x\,\langle x_i(t)|\rho_{\mathsf{M}}(0)|x_j(t)\rangle\,{\rm e}^{-D_t(x_i(t),x_j(t);s_i,s_j)}{\rm e}^{-{\rm i}\phi_t} \tag{1.42}$$

is almost diagonal at time $t \approx \lambda_{\rm th}/(g\delta s)$, a time much shorter than $\tau_{\rm ent}$ (see (1.31)), because of the form of the initial pointer kernel (1.27). This is, however, not a relevant issue for the measurement since by disregarding the pointer, one dismisses the possibly to get some information on S through the readout of the pointer observable.

1.3.5 Decoherence time

For a nonzero pointer-reservoir coupling, we will argue below that the off-diagonal terms in (1.32) disappear completely after some decoherence time that we now proceed to determine. As $\langle x_i(t)|\rho_{\mathsf{M}}(0)|x'_j(t)\rangle$ almost vanishes when $|x_i(t)| \geq \Delta_{\rm eff}$ or $|x'_j(t)| \geq \Delta_{\rm eff}$, one can appreciate the fate of these off-diagonal contributions by setting $x_i(t) = x'_j(t) = 0$ in (1.32). The decoherence factor then reads

$${\rm e}^{-D_t^{\rm peak}(s_i,s_j)} = \exp\left\{-\frac{g^2(s_i-s_j)^2}{2}\int_0^t{\rm d}\tau_1\int_0^t{\rm d}\tau_2\,\tau_1\tau_2\,{\rm Re}\,K_{\mathsf{R}}(\tau_1-\tau_2)\right\} \tag{1.43}$$

and reveals irreversible decay as soon as t much exceeds the decoherence time $\tau_{\rm dec}(s_i, s_j)$. We may define that time implicitly as $D^{\rm peak}_{\tau_{\rm dec}(s_i,s_j)}(s_i, s_j) = 1$. It is not difficult to prove that $D_t^{\rm peak}(s_i, s_j)$ is a positive increasing

[30] Recall that S has a discrete spectrum. We also suppose here that $p_i = 0$ if s_i belongs to a part of the spectrum containing arbitrarily close eigenvalues (that is, near an accumulation point), so that $\delta s > 0$.

[31] For indeed, one could design a new apparatus interacting with SM to measure the coherences of $\rho_{\mathsf{SM}}(t)$; this would contradict the measurement postulate.

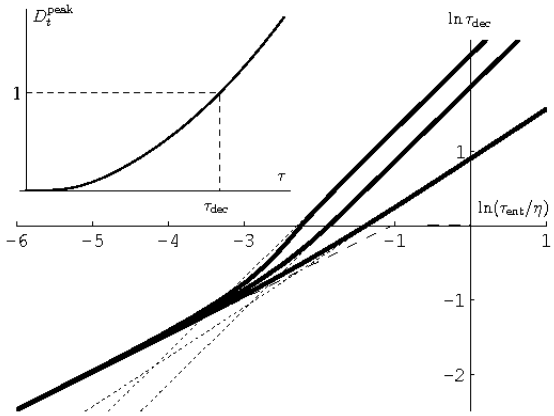

Figure 1.5: Solid curves: decoherence time $\tau_{\rm dec}$ against $\tau_{\rm ent}/\eta$ in a log-log scale. Both times are given in units of $\tau_{\mathsf{R}} = \beta$. The bath correlator is given by (1.46) with $t_{\mathsf{R}} = \omega_D^{-1} = \beta/5$ and $m = 5, 3, 1$ (from top to bottom). Broken curves: approximate expressions (1.47) for $\tau_{\rm dec} \lesssim t_{\mathsf{R}}$ (dashed lines) and $\tau_{\rm dec} \gg \tau_{\mathsf{R}}$ (dotted lines). **Inset:** decoherence exponent $D_t^{\rm peak}$ against $\tau = t/\tau_{\mathsf{R}}$ for $m = 3$.

convex function of time if $s_i \neq s_j$, as is observed in the inset of Fig. 1.5. Because $D_t^{\rm peak}(s_i, s_j) \propto (s_i - s_j)^2$, the maximal decoherence time $\tau_{\rm dec}(s_i, s_j)$ for all pairs (s_i, s_j) of distinct eigenvalues present in the initial state is (not surprisingly) reached when $s_j = s_i + \delta s$. We denote this largest decoherence time by $\tau_{\rm dec}$ and write $D_t^{\rm peak} = D_t^{\rm peak}(s_i, s_i + \delta s)$. The decoherence and entanglement times are related by

$$\left(\frac{\tau_{\rm ent}}{\eta}\right)^2 = \frac{1}{\beta^2} \int_0^{\tau_{\rm dec}} {\rm d}\tau_1 \int_0^{\tau_1} {\rm d}\tau_2 \, \tau_1 \tau_2 \, \frac{{\rm Re}\, K_{\mathsf{R}}(\tau_1 - \tau_2)}{K_{\mathsf{R}}(0)} \, , \tag{1.44}$$

where

$$\eta = \Delta_{\rm eff} K_{\mathsf{R}}(0)^{\frac{1}{2}} \beta < 1 \tag{1.45}$$

is a dimensionless measure of the strength of the apparatus-reservoir coupling[32]. The result (1.44) is valid provided that $\tau_{\rm dec}$ is much smaller than both T_{S} and T_{M}, a condition that must be checked *a posteriori*.

Figure 1.5 displays $\tau_{\rm dec}$ as a function of $\tau_{\rm ent}/\eta$ for a specific choice of bath correlator $K_{\mathsf{R}}(t)$, given by the KMS relation (1.24) and a temperature-independent imaginary part,

$$\widehat{{\rm Im}\, K_{\mathsf{R}}}(\omega) = -{\rm i} J(\omega) \quad , \quad J(\omega) = \widehat{\gamma}\, \omega^m \exp\left(- \frac{\omega^2}{\omega_D^2} \right) \quad , \quad \omega \geq 0 \tag{1.46}$$

(here ω_D is a Debye cut-off frequency). This choice corresponds to a bath of harmonic oscillators linearly coupled to the pointer position X with a power spectrum function $J(\omega)$ [256]. The bath correlation time τ_{R} is the thermal time, $\tau_{\mathsf{R}} = \beta > \omega_D^{-1}$. The three plain curves are obtained by solving numerically the implicit equation (1.44).

Formula (1.44) explicitly yields the decoherence time in two opposite limits.

(i) **Interaction-dominated regime:** when $t \lesssim t_{\mathsf{R}}$, the dynamics is dominated by the system-pointer and pointer-reservoir interactions. One may approximate $K_{\mathsf{R}}(\tau)$ by $K_{\mathsf{R}}(0) = \langle B^2 \rangle_{\mathsf{R}}$ in (1.43) and (1.44).

(ii) **Markov regime:** this regime corresponds to the so-called singular coupling limit[33] $\tau_{\rm dec} \gg \tau_{\mathsf{R}}$. Decoherence is then governed by the small-frequency behavior of $(\widehat{{\rm Re}\, K})(\omega)$.

We present here directly the results in these two limits. The technical justifications can be found in [224]. We first discuss the case of inverse temperatures $\beta \lesssim t_{\mathsf{R}}$. Then the decoherence factor reads

$$ {\rm e}^{-D_t^{\rm peak}} = \exp\left\{ - \left(\frac{t}{\tau_{\rm dec}} \right)^{\gamma} \right\} \quad , \quad \tau_{\rm dec} = \left(c_m^{(\gamma)} \right)^{\frac{1}{\gamma}} \beta \left(\frac{\tau_{\rm ent}}{\beta \eta} \right)^{\frac{2}{\gamma}} \tag{1.47}$$

[32] More precisely, η is the fluctuation of the coupling energy in the uncorrelated state $\rho_{\mathsf{M}}(0) \otimes \rho_{\mathsf{R}}$, in units of β^{-1}. The inequality $\eta < 1$ results from our assumption $\Delta_{\rm th}^2 K_{\mathsf{R}}(0) \beta^2 < 1/2$ ensuring the stability condition (1.29).

[33] Note that this Markovian regime is different from that arising in the van Hove (weak-coupling) limit [73]. The van Hove limit enforces the separation of time scales $\tau_{\mathsf{R}}, T_{\mathsf{S}} \ll \tau_{\rm dec}$ justifying a Born-Markov and a rotating-wave approximations. The latter approximation is inappropriate here due to our restriction (1.25). In contrast, the singular coupling limit $\tau_{\mathsf{R}} \ll \tau_{\rm dec}, T_{\mathsf{S}}$ is the limit of delta correlated reservoirs [121, 103, 189]. This limit, which is the one considered here, also yields Markovian evolutions.

for $t, \tau_{\rm dec} \ll T_{\mathsf{S}}, T_{\mathsf{M}}$, with the exponent γ given by

$$\gamma = \begin{cases} 4 & \text{if} \quad \tau_{\rm dec} \lesssim t_{\mathsf{R}} & \text{(interaction-dominated regime)} \\ 3 & \text{if} \quad \tau_{\rm dec} \gg \tau_{\mathsf{R}} \ , \ m = 1 & \text{(Markov regime, Ohmic reservoir)} \\ 2 & \text{if} \quad \tau_{\rm dec} \gg \tau_{\mathsf{R}} \ , \ m \geq 3 & \text{(Markov regime, super-Ohmic reservoir)} \end{cases} \tag{1.48}$$

and some constants $c_m^{(\gamma)}$ independent of the strengths of the SM- and MR-couplings,

$$c_m^{(4)} = 8 \quad , \quad c_{m=1}^{(3)} = \frac{3\beta K_{\mathsf{R}}(0)}{\int_0^\infty {\rm d}\tau {\rm Re}\, K_{\mathsf{R}}(\tau)} \quad , \quad c_{m\geq 3}^{(2)} = \frac{2\beta^2 K_{\mathsf{R}}(0)}{|\int_0^\infty {\rm d}\tau\, \tau\, {\rm Re}\, K_{\mathsf{R}}(\tau)|} \ . \tag{1.49}$$

By the Cauchy Schwarz inequality one has $|K_{\mathsf{R}}(t)| \leq K_{\mathsf{R}}(0)$, hence $c_{m=1}^{(3)} \geq 3\beta/\tau_{\mathsf{R}}$ and $c_{m\geq 3}^{(2)} \geq (2\beta/\tau_{\mathsf{R}})^2$. For a reservoir at low temperature[34], i.e., such that $\beta \gg t_{\mathsf{R}} = \omega_D^{-1}$, the decoherence factor and decoherence time are still given by (1.47) with $\gamma = 4$ in the interaction-dominated regime $\tau_{\rm dec} \lesssim t_{\mathsf{R}}$. In the opposite limit $t_{\mathsf{R}} \ll \tau_{\rm dec} \ll \beta$ and for correlators with a temperature-independent imaginary part given by (1.46), one has

$$\tau_{\rm ent} \simeq \begin{cases} \eta_D \tau_{\rm dec} \Big(\ln\Big(\frac{\tau_{\rm dec}}{t_{\mathsf{R}}}\Big) - 0.2114) \Big)^{\frac{1}{2}} & \text{if } m = 1 \quad \text{(Ohmic)} \\ \dfrac{\eta_D \tau_{\rm dec}}{\sqrt{m-1}} & \text{if } m \geq 3 \quad \text{(super-Ohmic),} \end{cases} \tag{1.50}$$

where $\eta_D = \Delta_{\rm eff} K_{\mathsf{R}}(0)^{1/2} t_{\mathsf{R}}$ is the pointer-reservoir coupling strength in units of $1/t_{\mathsf{R}}$.

In the interaction-dominated regime, the decoherence time $\tau_{\rm dec}$ presents a universal behavior, namely, it depends on the reservoir through the coupling strength η only. The decoherence factor is a stretched exponential with exponent $\gamma = 4$, instead of the usual Gaussian for systems coupled to their environment in this regime (Sec. 1.2.5). This non-gaussian form can be readily explained by remembering that the reservoir acts on the pointer M, not directly on S, and that it takes time to S to transform the initial state of M into a superposition. In fact, the correct stretched exponential can be obtained, albeit with a wrong numerical factor $c_m^{(4)}$, by replacing the pointer positions x_i and x_j in (1.17) by $x_i(t)$ and $x_j(t)$ as given by (1.34). It follows from the consistency condition $\tau_{\rm dec} \lesssim t_{\mathsf{R}}$ and the stability condition $\eta < 1$ that $\tau_{\rm ent} \lesssim \tau_{\rm dec}(t_{\mathsf{R}}/\beta)$. By the inequality $t_{\mathsf{R}} \leq \beta$, one has $\tau_{\rm ent} \lesssim \tau_{\rm dec}$.

In the Markov regime and at inverse temperatures $\beta \lesssim t_{\mathsf{R}}$, qualitatively different results are obtained for Ohmic and super-Ohmic reservoirs (non-universal regime). Ohmic reservoirs win in efficiency for decoherence over super-Ohmic reservoirs, that is, they lead to much smaller decoherence times[35]. It is interesting to compare our result with the known saturation of the decoherence factor to a positive value in open systems coupled to super-Ohmic baths in the singular coupling limit [188]. In contrast, we see here that, as a consequence of the indirect coupling of S *via the pointer*, the decoherence factor does not saturate but decays to zero, although more slowly than for an Ohmic reservoir.

At low temperature and in the Markov regime, the ratio between the decoherence times for super-Ohmic and Ohmic reservoirs is logarithmic in the large ratio $\tau_{\rm dec}/t_{\mathsf{R}}$. Hence an Ohmic reservoir is not dramatically more efficient than a super-Ohmic one at very low temperature, inversely to the high temperature case.

By invoking (1.43) and $\widehat{({\rm Re}\, h)}(\omega) \geq 0$, it is easy to establish that $D_t^{\rm peak} \leq (t/\tau_{\rm dec}^{(\gamma=4)})^4$ for all times $t \geq 0$, $\tau_{\rm dec}^{(\gamma=4)}$ being the decoherence time in the interaction-dominated regime. As a result, the decoherence time $\tau_{\rm dec}$ is bounded from below by $\tau_{\rm dec}^{(\gamma=4)}$, whatever the value of the ratio $\tau_{\rm dec}/t_{\mathsf{R}}$. This is indeed what is seen in Fig. 1.5. One observes a quite good agreement between the exact and asymptotic behaviors of the decoherence time for $\tau_{\rm dec} \lesssim t_{\mathsf{R}}$ and $\tau_{\rm dec} \gtrsim \tau_{\mathsf{R}}$. The plain curves representing $\tau_{\rm dec}$ split by increasing $\tau_{\rm ent}$ into distinct branches corresponding to distinct m's, as predicted by (1.48). After this splitting, which occurs in the transition region $t_{\mathsf{R}} \lesssim \tau_{\rm dec} \lesssim \tau_{\mathsf{R}}$, the decoherence time $\tau_{\rm dec}$ is larger for larger m's.

By studying carefully the two above limiting regimes, one can prove the following statements[36].

[34] Strictly speaking, extremely low temperatures have to be proscribed because of our hypothesis $\beta \ll T_{\mathsf{M}}$. However, taking e.g. $T_{\mathsf{M}} = 1\,{\rm s}$, $\beta \ll T_{\mathsf{M}}$ holds even for temperatures of the order of 10^{-8} degree Kelvin! Furthermore, the stability condition (1.45) has a better chance to be met at low temperature since $\Delta_{\rm eff}$ is temperature decreasing.

[35] Actually, the ratio $(\tau_{\rm dec}^{(m=1)}/\tau_{\rm dec}^{(m\geq 3)})^3$ is equal to the product of $\beta c_{m=1}^{(3)}/(\tau_{\mathsf{R}} c_{m\geq 3}^{(2)})$ by $\tau_{\mathsf{R}}/\tau_{\rm dec}^{(m\geq 3)}$. Since the last factor must be small compared with unity for consistency and the first one is $\lesssim 1$ by definition of τ_{R}, it follows that $\tau_{\rm dec}^{(m=1)} \ll \tau_{\rm dec}^{(m\geq 3)}$.

[36] The first statement can be justified by estimating in these limits the minimum of the decoherence exponent $D_t(x, x'; s_i, s_j)$ over all positions x and x', with t, s_i, and s_j fixed (let us recall that $\tau_{\rm dec}$ is the time when this exponent is unity for the specific positions $(x, x') = (gs_i t, gs_j t)$). The second statement is established by studying the behavior of the decoherence factor in (1.40) in the asymptotic regimes (for more detail see [225]).

1) For $i \neq j$ one has $\Pi_i \rho_{\mathsf{SM}}(t) \Pi_j \simeq 0$ at time $t \gg \tau_{\rm dec}$, i.e., the off-diagonal contributions in the kernel (1.32) are vanishingly small for all values of (x, x').

2) The decoherence factor for the pointer kernel $\langle x|\rho_{\mathsf{M}|i}(t)|x'\rangle$ (first exponential in (1.40)) remains close to unity at time $t \approx \tau_{\rm dec}$ if $|x_i(t)|$ and $|x_i'(t)|$ are smaller than $\Delta_{\rm eff}$.

We may conclude from the first statement that the dynamical process (1.3) is entirely governed in our model by the two time scales $\tau_{\rm ent}$ and $\tau_{\rm dec}$. The second statement means that decoherence does away with the off-diagonal $(i \neq j)$ terms before the diagonal $(i = j)$ ones change noticeably.

1.3.6 Interaction and measurement times

According to the statement 1) above, the system and pointer are almost in the classical mixed state $\rho_{\mathsf{SM}}(t)$ given by (1.3) at time $t \gtrsim \tau_{\rm dec}$. If moreover $t \gtrsim \tau_{\rm ent}$, the pointer is with probability p_i in a state $\rho_{\mathsf{M}|i}(t)$ with a well-defined position $x = g s_i t$. As emphasized in Sec. 1.2.1, in order to be able to read the outcomes on the pointer without perturbing it significantly, these positions should be separated by macroscopic distances at the end of the measurement (time $t = t_{\rm meas}$). This macroscopic separation can be achieved by using the instability induced by the pointer-reservoir coupling as an amplification mechanism. Actually, let the interaction between S and M be switched off at time

$$t_{\rm int} \approx W_{\rm eff} (g \delta s)^{-1} \; . \tag{1.51}$$

Then all pointer states $\rho_{\mathsf{M}|i}(t)$ are outside the central well of the effective potential (save for possibly one eigenvalue $s_i = 0$). Thanks to (1.31), one has $t_{\rm int} \gg \tau_{\rm ent}$. The "mesoscopic" pointer separation $W_{\rm eff}$ at time $t = t_{\rm int}$ is subsequently amplified by the effective pointer dynamics, till it reaches a macroscopically resolvable magnitude $\Delta_{\rm class}$ at time $t_{\rm meas}$. Then a pointer reading, while still a physical process in principle perturbing M, surely cannot blur the distinction of the peaks in the pointer density. Furthermore, if the measured observable is a spin one-half and the effective pointer potential has the triple-well shape represented in Fig. 1.3, then the pointer ends up at time $t_{\rm meas}$ in one of the two extremal wells separated by the distance $\Delta_{\rm class}$. Assuming that these wells have heights much larger than the thermal energy, M is in an equilibrium state and its state will not subsequently change, thereby ensuring the registration of the result. The time $t_{\rm meas}$ depends on the shape of the effective potential. Since we have assumed $W_{\rm eff} \ll \Delta_{\rm class}$ and $t_{\rm int} \ll T_{\mathsf{M}}$, see (1.25), it is clear that $t_{\rm int} \ll t_{\rm meas}$.

1.3.7 Conclusions

Our apparatus with a single-degree-of-freedom macroscopic pointer coupled to the system and reservoir via the Hamiltonians (1.11) and (1.15) performs a measurement of the system observable S. The precise form of the pointer-reservoir coupling is not important, save for simplifying calculations. Unlike in the description of quantum measurements presented in old textbooks, it has not been necessary to postulate the existence of a new kind of non-unitary dynamics: we have only used the Schrödinger equation and statistical physics considerations. We have still not fully explained the reduction of the wave packet: we have not shown that single runs of the measurement lead to definite outcomes and to conditional states taken from the ensemble $\{\rho_{\mathsf{S}|i} \otimes \rho^{\rm eq}_{\mathsf{M}|i}, p_i\}$. These issues will be discussed in the next section.

However, the main interest of our model is not that it explains (with the limitations mentioned above) the von Neumann projection postulate. Many other models in the literature based on the "system-apparatus-reservoir" approach do so, see e.g. [102, 277, 48, 4, 5, 210]. The interesting point is that, like in Refs. [5, 6], we have been able to determine explicitly the relevant time scales of the measurement process. Indeed, as stressed by the authors of Ref. [6], "a full understanding of quantum mechanics requires knowledge of the time scales involved in measurements" (op. cit. from [6], chapter 1). Four fundamental times characterizing the measurement have been introduced. The *entanglement time* $\tau_{\rm ent}$ is the time after which pointer positions corresponding to distinct eigenvalues s_i of S begin to be resolved. It is given by $\tau_{\rm ent} = \Delta_{\rm eff} (g \delta s)^{-1}$, where g is the system-pointer coupling constant, δs the separation between neighboring eigenvalues, and $\Delta_{\rm eff} \approx \Delta_{\rm th}$ the uncertainty in the initial pointer position. The system and apparatus must interact during a time $t_{\rm int}$ much larger than $\tau_{\rm ent}$. Accordingly, $\tau_{\rm ent}$ provides a good measure of the efficiency of the system-pointer interaction (the smaller $\tau_{\rm ent}$, the more efficient is the coupling). The second fundamental time of the measurement is the *decoherence time* $\tau_{\rm dec}$, that is, the time after which the system-pointer density matrix is almost a statistical mixture of states $\rho_{\mathsf{S}|i} \otimes \rho_{\mathsf{M}|i}(t)$ with $\rho_{\mathsf{S}|i} = \Pi_i \rho_{\mathsf{S}}(0) \Pi_i$ and $\rho_{\mathsf{M}|i}(t)$ the pointer state revealing the eigenvalue s_i. The third time is the system-apparatus interaction time $t_{\rm int}$ mentioned above. It is much smaller than the time

$t_{\rm meas}$ after which the pointer states $\rho_{\mathsf{M}|i}(t)$ relax to the equilibria $\rho^{\rm eq}_{\mathsf{M}|i}$ with positions $x_i = {\rm tr}(\rho^{\rm eq}_{\mathsf{M}|i} X)$ separated by macroscopic distances. In our model, the four time scales $\tau_{\rm ent}$, $\tau_{\rm dec}$, $t_{\rm int}$, and $t_{\rm meas}$ satisfy

$$\tau_{\rm ent} \ll t_{\rm int} \ll t_{\rm meas} \quad , \quad \tau_{\rm ent}, \tau_{\rm dec} \ll T_{\mathsf{S}}, T_{\mathsf{M}} \; , \tag{1.52}$$

where T_{S} and T_{M} are the times associated to the dynamics of the isolated system and pointer. Under this hypothesis, we have shown that $\tau_{\rm dec}$ is given by the implicit equation (1.44) and have derived an explicit relation (1.47) between $\tau_{\rm dec}$ and $\tau_{\rm ent}$ in the interaction-dominated and Markov regimes. These times are ordered as follows:

(i) Interaction-dominated regime: $\tau_{\rm ent} \lesssim \tau_{\rm dec}(t_{\mathsf{R}}/\beta) \leq \tau_{\rm dec} \lesssim t_{\mathsf{R}}$;

(ii) Markov regime ($\tau_{\rm dec} \gg \tau_{\mathsf{R}}$): $\tau_{\rm ent} \leq \tau_{\rm dec}$ for a super-Ohmic reservoir with a pointer-reservoir coupling strength $\eta \lesssim \beta/\tau_{\mathsf{R}} \leq (c^{(2)}_m)^{1/2}$, where η is defined by (1.45); for an Ohmic reservoir this inequality holds for strong enough system-pointer couplings only, more precisely, when $\tau_{\rm ent} \lesssim c^{(3)}_1 \beta/\eta^2$, where $c^{(\gamma)}_m$ is given by (1.49).

Therefore, the only regime with a decoherence faster than resolution of pointer peaks is the Markov regime with $m = 1$ (Ohmic reservoir).

Before closing this section, it is worthwhile to remark that for a reasonably strong pointer-reservoir coupling and a not too strong system-pointer coupling, the decoherence time is small enough to ensure that the whole measurement is performed without producing a Schrödinger cat state as an intermediate step. More precisely, if $\tau_{\rm dec} \leq t_{\rm int}$ then even the mesoscopic superpositions of pointer states with space separation $W_{\rm eff}$ decay to mixtures faster than entanglement can create them. As already pointed out above, this comes from the simultaneous action of the system-pointer and pointer-reservoir interactions. According to (1.47) and (1.51), in the interaction-dominated and Markov regimes the condition $\tau_{\rm dec} \leq t_{\rm int}$ is fulfilled provided that the following inequality holds

$$\eta \geq \sqrt{c^{(\gamma)}_m} \left(\frac{\Delta_{\rm eff}}{W_{\rm eff}} \right)^{\frac{\gamma}{2}} \left(\frac{\beta}{\tau_{\rm ent}} \right)^{\frac{\gamma-2}{2}} . \tag{1.53}$$

1.4 Derivation of the measurement postulate from a statistical interpretation

1.4.1 State reduction for sub-ensembles

We have argued so far that the evolution of the system-apparatus density matrix is given by (1.3) under the influence of the system-apparatus and apparatus-reservoir couplings, thereby justifying from Schrödinger's equation the point 1) in Sec. 1.2.2. In order to establish from first principles the reduction of the wave packet, it remains to clarify the second point 2). Recently, Allahverdyan, Balian and Nieuwenhuizen [6, 7] have derived the state reduction for single runs of a measurement by relying on the usual statistical interpretation of quantum states, taken as a basic postulate. The reduced state of the bipartite system SM, as any quantum state, is interpreted as referring to a thought ensemble of identical systems, in which the real system SM is embedded (see the discussion in Sec. 1.2.2). The system-apparatus density matrix $\rho_{\mathsf{SM}}(t_{\rm meas})$ at the end of the measurement is thus postulated to describe a state ensemble, although one does not know *a priori* which one (as stressed in Sec. 1.2.3, many different ensembles are associated to this density matrix). The main idea of the authors of Refs. [6, 7] is that the system-apparatus dynamics selects the particular ensemble $\{\rho_{\mathsf{SM}|i}, p_i\}$ among all ensembles, with $\rho_{\mathsf{SM}|i} = \rho_{\mathsf{S}|i} \otimes \rho^{\rm eq}_{\mathsf{M}|i}$ the conditional system-apparatus state given outcome i, see (1.5).

More precisely, Allahverdyan *et al.* consider arbitrary sub-ensembles extracted from the density matrix $\rho_{\mathsf{SM}}(t_{\rm meas})$. Let us recall that a sub-ensemble $\{\rho_i, p_{i|I}\}_{i\in I}$ from a given state ensemble $\{\rho_i, p_i\}_{i=1}^n$ is obtained by choosing a subset of indices $I \subset \{1, \ldots, n\}$ and keeping only the states ρ_i with $i \in I$, to which one attaches the probabilities $p_{i|I} = p_i / \sum_{j\in I} p_j$. This operation is a conditioning of the original ensemble subsequent to an updating of information telling us that the identical systems in the ensemble are in fact all in states ρ_i with $i \in I$. This amounts to a filtering (post-selection) among the states produced in the state preparation. The smaller possible sub-ensemble is composed of a single state ρ_i having probability $p_{i|i} = 1$. Given a density matrix ρ, any sub-ensemble of an ensemble forming a convex decomposition of ρ can be obtained as follows: one first decomposes ρ as a sum of two non-negative matrices $\widetilde{\rho}^{\rm (sub)} \geq 0$ and $\widetilde{\rho}^{\rm (comp)} \geq 0$, $\rho = \widetilde{\rho}^{\rm (sub)} + \widetilde{\rho}^{\rm (comp)}$, and then finds a convex decomposition of $\rho^{\rm (sub)} = \widetilde{\rho}^{\rm (sub)} / {\rm tr}(\widetilde{\rho}^{\rm (sub)})$.

The precise statement in Refs. [6, 7] is: for any sub-ensemble of an ensemble forming a convex decomposition of $\rho_{\mathsf{SM}}(t_{\rm meas})$, the irreversible system-apparatus dynamics transforms the density matrix associated to this sub-ensemble into

$$\rho_{\mathsf{SM}}^{(\rm sub)}(t) = \sum_{i=1}^{n} q_i \rho_{\mathsf{S}|i} \otimes \rho_{\mathsf{M}|i}(t) \quad , \quad t \geq t_{\rm meas} + \Delta t \; , \tag{1.54}$$

after some time Δt (which depends on the apparatus-reservoir coupling). Here, $q_i \geq 0$ are some probabilities depending on the sub-ensemble. The formula (1.54) generalizes (1.3), which is recovered by choosing the sub-ensemble equal to the full ensemble.

1.4.2 Emergence of single outcomes in individual measurements

With the aim to show (1.54) and to understand how single runs of a measurement lead to well-defined outcomes i and post-measurement states $\rho_{\mathsf{SM}|i}$, let us consider the following model inspired from the ideas of Ref. [6] applied to our apparatus of Sec. 1.3. Since a single outcome may only arise from the information obtained upon observing the value of the pointer variable X, the readout process by which one acquires this information should be taken into account explicitly (Sec. 1.2.2). Note that this process is purely classical by condition 3 of Sec. 1.2.1. We describe it by imagining an observer measuring the pointer position X at some times $t_{\rm meas} = t_0 < t_1 < t_2 < \cdots < t_m$ separated by a fixed time interval Δt, assumed to be larger than the decoherence time $\tau_{\rm dec}$. Each observation increases the observer knowledge about the pointer. It thus corresponds to a selection of a given sub-ensemble of an ensemble associated to the system-pointer density matrix. Inasmuch as the pointer behaves classically at time t_k, the observation process is completely analogous to that used to find the position of a classical particle distributed according to a probability density $f(x)$. The observation does not disturb significantly the system-pointer state, apart from the updating of information.

To simplify the discussion, we assume that the measured observable S has a non-degenerate spectrum and that the apparatus has pure pointer states $|\mu_{\mathsf{M}|i}^{\rm eq}\rangle$ (the generalization to mixed states $\rho_{\mathsf{M}|i}^{\rm eq}$ is straightforward by taking spectral decompositions). We denote by $\{|\alpha_i\rangle\}$ the orthonormal eigenbasis of S. The system-apparatus state at time $t_0 = t_{\rm meas}$ just before the observation is

$$\rho_{\mathsf{SM}}(t_0) = \sum_{i=1}^{n} |c_i|^2 |\alpha_i\rangle\langle\alpha_i| \otimes |\mu_{\mathsf{M}|i}^{\rm eq}\rangle\langle\mu_{\mathsf{M}|i}^{\rm eq}| \; . \tag{1.55}$$

We denote by $\rho_{\mathsf{SM}|x_0\ldots x_k}(t_k)$ the system-apparatus state after k consecutive observations at times $t_0, t_1, \ldots, t_k$. After the first observation, one has

$$\rho_{\mathsf{SM}|x_0}(t_0) = \Bigl(\sum_{j\in J_0} q_j^{(0)} \Bigr)^{-1} \sum_{j\in J_0} q_j^{(0)} |\Psi_{\mathsf{SM}|j}^{(0)}\rangle\langle\Psi_{\mathsf{SM}|j}^{(0)}| \; , \tag{1.56}$$

where $\rho_{\mathsf{SM}}(t_0) = \sum_{j=1}^{n_0} q_j^{(0)} |\Psi_{\mathsf{SM}|j}^{(0)}\rangle\langle\Psi_{\mathsf{SM}|j}^{(0)}|$ is an arbitrary convex decomposition of $\rho_{\mathsf{SM}}(t_0)$, which may differ from the decomposition (1.55), and J_0 is a subset of $\{1, \ldots, n_0\}$. The vectors $|\Psi_{\mathsf{SM}|j}^{(0)}\rangle$ and $|\alpha_i\rangle|\mu_{\mathsf{M}|i}^{\rm eq}\rangle$ of the two pure state decompositions of $\rho_{\mathsf{SM}}(t_0)$ are related as follows (see Eq. (4.16) in chapter 4)

$$\sqrt{q_j^{(0)}} |\Psi_{\mathsf{SM}|j}^{(0)}\rangle = \sum_{i=1}^{n} |c_i| u_{ji}^{(0)} \, |\alpha_i\rangle|\mu_{\mathsf{M}|i}^{\rm eq}\rangle \; , \tag{1.57}$$

where $(u_{ji}^{(0)})$ is a unitary matrix with size $\max\{n, n_0\}$. Due to the coupling with the reservoir, each of the pure states $|\Psi_{\mathsf{SM}|j}^{(0)}\rangle\langle\Psi_{\mathsf{SM}|j}^{(0)}|$ is transformed by the decoherence process (1.14). By linearity of the evolution, the system-apparatus state reads at time $t_1 = t_0 + \Delta t$, $\Delta t \gg \tau_{\rm dec}$,

$$\rho_{\mathsf{SM}|x_0}(t_1) = \Bigl(\sum_{i=1}^{n} |c_i|^2 \xi_i^{(0)} \Bigr)^{-1} \sum_{i=1}^{n} |c_i|^2 \xi_i^{(0)} |\alpha_i\rangle\langle\alpha_i| \otimes |\mu_{\mathsf{M}|i}^{\rm eq}\rangle\langle\mu_{\mathsf{M}|i}^{\rm eq}| \tag{1.58}$$

with $\xi_i^{(0)} = \sum_{j\in J_0} |u_{ji}^{(0)}|^2$. Repeating this operation m times, we arrive at the system-apparatus state

$$\rho_{\mathsf{SM}|x_0\ldots x_m}(t_m + \Delta t) = \sum_{i=1}^{n} |c_i|^2 \zeta_i^{(m)} |\alpha_i\rangle\langle\alpha_i| \otimes |\mu_{\mathsf{M}|i}^{\rm eq}\rangle\langle\mu_{\mathsf{M}|i}^{\rm eq}| \quad , \quad \zeta_i^{(m)} = \frac{\prod_{k=0}^{m} \xi_i^{(k)}}{\sum_{l=1}^{n} |c_l|^2 \prod_{k=0}^{m} \xi_l^{(k)}} \; , \tag{1.59}$$

in which the coefficients $\xi_i^{(k)}$ are defined as $\xi_i^{(0)}$ but with k-dependent unitary matrices $U^{(k)} = (u_{ji}^{(k)})$ and subsets $J_k \subset \{1, \ldots, n_k\}$. Note that the unitarity of $U^{(k)}$ implies $0 \leq \xi_i^{(k)} \leq 1$. The collapse to the state

$$\rho_{\mathsf{SM}|x_0 \ldots x_m}(t_m + \Delta t) = |\alpha_{i_0}\rangle\langle\alpha_{i_0}| \otimes |\mu^{\mathrm{eq}}_{\mathsf{M}|i_0}\rangle\langle\mu^{\mathrm{eq}}_{\mathsf{M}|i_0}| \tag{1.60}$$

occurs for large m provided that the unitary matrices $U^{(k)}$ are such that $\zeta_i^{(m)}$ converges to δ_{ii_0}.

The latter condition is not always fulfilled, as might be expected since we have not yet made any assumption on the selection of the sub-ensembles. In order to get a unique outcome, this selection must be such that the observer is certain at time t_m that the pointer value is $X = x_{i_0}$. By (1.55), at time t_0 the pointer density is $f_{\mathsf{M}}(x, t_0) = \sum_i |c_i|^2 |\langle x|\mu^{\mathrm{eq}}_{\mathsf{M}|i}\rangle|^2$. This density is modified by the updating of information during each observation and given by conditioning the classical distribution $f_{\mathsf{M}}(x, t_0)$ to the observation results. This conditioning should produce at the end a localized distribution $f_{\mathsf{M}|x_0 \ldots x_m}(x, t_m) = |\langle x|\mu^{\mathrm{eq}}_{\mathsf{M}|i_0}\rangle|^2$, otherwise the observer has failed to identify a specific pointer position. But thanks to (1.59),

$$f_{\mathsf{M}|x_0 \ldots x_m}(x, t_m) = \langle x| \operatorname{tr}_{\mathsf{S}}(\rho_{\mathsf{SM}|x_0 \ldots x_m}(t_m))|x\rangle = \sum_{i=1}^{n} |c_i|^2 \zeta_i^{(m)} \left|\langle x|\mu^{\mathrm{eq}}_{\mathsf{M}|i}\rangle\right|^2 . \tag{1.61}$$

The aforementioned convergence of the coefficients $\zeta_i^{(k)}$ is thus fulfilled if the readout process has provided enough information about the pointer position in order to identify in which of the different equilibria $|\mu^{\mathrm{eq}}_{\mathsf{M}|i}\rangle$ is the pointer.

1.4.3 Discussion

The emergence of the state (1.60) is due to the robustness of the pointer states $|\mu^{\mathrm{eq}}_{\mathsf{M}|i}\rangle$ with respect to the coupling with the reservoir, while linear superpositions of these states rapidly decohere to statistical mixtures. Thus $\{|\alpha_i\rangle \otimes |\mu^{\mathrm{eq}}_{\mathsf{M}|i}\rangle, p_i\}$ is the only ensemble associated to the density matrix $\rho_{\mathsf{SM}}(t_{\mathrm{meas}})$ which is not strongly affected by the dynamics on the time scale Δt. Recalling that at time t_{meas} the pointer states are macroscopically distinct, the decoherence time τ_{dec} giving the evolution of all other ensembles (which should involve linear combinations of the $|\mu^{\mathrm{eq}}_{\mathsf{M}|i}\rangle$) is extremely short, much shorter than any classical time resolution Δt in the readout process. Consequently, the irreversible dynamics of the system and apparatus does not only implement the state transformation (1.3), it also prepares them in the ensemble $\{|\alpha_i\rangle \otimes |\mu^{\mathrm{eq}}_{\mathsf{M}|i}\rangle, p_i\}$.

It is remarkable that this ensemble coincides with the ensemble giving the larger entropy balance (1.8), that is, the largest gain of information on S (equal to $S_{\mathrm{v.N.}}(\rho_{\mathsf{S}}(0))$ if S has a non-degenerate spectrum). All others pure state ensembles $\{|\Psi_{\mathsf{SM}|j}\rangle, q_j\}$ forming a decomposition of $\rho_{\mathsf{SM}}(t_{\mathrm{meas}})$ lead to smaller increases of information, i.e.,

$$S_{\mathrm{v.N.}}\big(\rho_{\mathsf{S}}(0)\big) - \sum_j q_j S_{\mathrm{v.N.}}\big(\operatorname{tr}_{\mathsf{M}}(|\Psi_{\mathsf{SM}|j}\rangle\langle\Psi_{\mathsf{SM}|j}|)\big) \;\leq\; S_{\mathrm{v.N.}}(\rho_{\mathsf{S}}(0)) . \tag{1.62}$$

The minimum over all such decompositions of the sum in the left-hand side is, by definition, the entanglement of formation $E_{\mathrm{EoF}}(\rho_{\mathsf{SM}}(t_{\mathrm{meas}}))$ quantifying the entanglement between S and M (see Sec. 2.3.3 below), which vanishes since $\rho_{\mathsf{SM}}(t_{\mathrm{meas}})$ is separable. It thus appears that the measurement prepares the bipartite system SM in the ensemble of separable states realizing the minimum in the definition of $E_{\mathrm{EoF}}(\rho_{\mathsf{SM}}(t_{\mathrm{meas}}))$.

The above argument is based on a statistical interpretation of the reduced system-apparatus state $\rho_{\mathsf{SM}}(t)$, henceforth it does not explain the origin of the randomness arising in a quantum measurement. This randomness should be (at least partly) linked with our incapacity to have full knowledge on the whole set of quantum correlations and phases of the particles forming the apparatus. Only a few apparatus collective variables (such as the pointer) can be controlled and reset to their initial value between each run of the measurement. Note, however, that it would be incorrect to attribute the occurrence of a specific outcome as being controlled by the initial microscopic state of the apparatus [259, 225].

Chapter 2

Entanglement evolution for quantum trajectories

Pour qu'une chose soit intéressante, il suffit de la regarder longtemps (G. Flaubert).

2.1 Motivations

In this chapter we discuss how the amount of entanglement in an initially entangled open bipartite system evolves with time. A basic property of entanglement is that it cannot be created by means of local operations on each parties. As a result, the entanglement between non-interacting subsystems cannot increase if each subsystem is coupled to its own local reservoir, the reservoirs being independent between themselves. In this case, the most robust states (pointer states) with respect to the interaction with the reservoirs are tensor products of pointer states for each subsystems. In other words, the density operator of the composite system becomes diagonal in a product basis at times much larger than the decoherence time. Thus decoherence transforms an entangled pure state given by a superposition of product states into a statistical mixture of these states, which is by definition separable, i.e., not entangled. Consequently, the entanglement in the composite system is expected to decay with time.

In the last decade, a lot of works have been devoted to the quantitative study of the evolution of entanglement in open bipartite systems AB. Different scenarios have been identified. In the independent reservoir case discussed above, entanglement disappears after a finite time if AB evolves asymptotically to a steady state belonging to the interior of the set of separable states, such as a Gibbs equilibrium at positive temperature [93, 234]. This also occurs for certain initial states if the steady state belongs to the boundary of the set of separable states, for instance, if AB decays to a separable ground state [81, 268, 8]. In the latter situation, one can also find entangled states with entanglement vanishing only asymptotically in the large time limit. Conversely, when the two subsystems are coupled to a common bath, entanglement can be created due to an effective interaction mediated by the bath [44]. Then sudden revivals of entanglement may take place after the state has become separable [92, 166].

In his Ph.D. thesis, S. Vogelsberger has shown that the decay of entanglement is completely different if continuous local measurements are performed on the two independent baths coupled to A and B: then, averaging over the measurement outcomes, the entanglement always decays exponentially (see (2a) in the publication list). He has determined the measurement scheme that better protects the bipartite system from entanglement losses, showing that in some cases entanglement can be completely preserved. Like in the vast majority of papers dealing with entanglement evolution in the absence of measurements[1], his analysis is restricted to two-qubit systems. The main reason is that for higher-dimensional systems it is quite hard to determine whether a state is entangled or separable (and, a fortiori, to estimate the amount of entanglement). Similar less general results have been discovered at the same time by E. Mascarenhas, D. Cavalcanti, V. Vedral, and M. França Santos [165].

The chapter is organized as follows. A general introduction to pure state random evolutions resulting from continuous measurements on the environment (quantum trajectories) is given in Sec. 2.2. We briefly review in Sec. 2.3 the notion of bipartite entanglement for mixed states and the aforementioned scenarios for its time evolution. The results obtained with Vogelsberger are explained in Sec. 2.4. The last section 2.5 contains concluding remarks.

[1] The number of papers on this problem is huge. We will not even try to cite the most relevant ones.

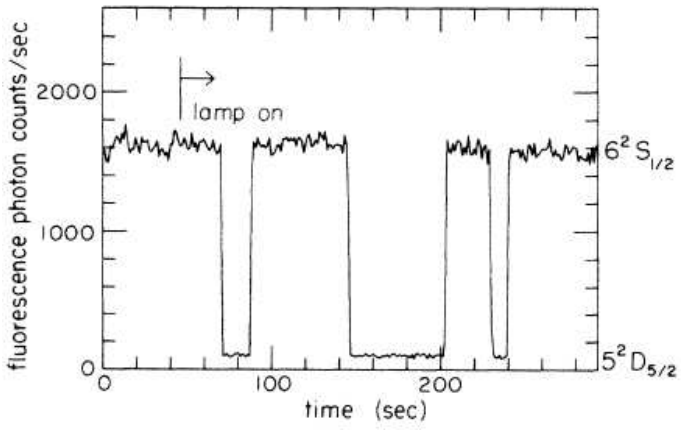

Figure 2.1: Fluorescence signal from a trapped Ba^+ ion at the wavelength corresponding to the $6^2P_{\frac{1}{2}}$-$6^2S_{\frac{1}{2}}$ level transition. This transition is induced by two lasers in resonance with the $6^2P_{\frac{1}{2}}$-$6^2S_{\frac{1}{2}}$ and $6^2P_{\frac{1}{2}}$-$5^2D_{\frac{3}{2}}$ transitions. The signal stops at random times when the ion jumps to the metastable level $5^2D_{\frac{5}{2}}$ (incoherent excitation) and is retrieved when it jumps back from this level to the $6^2S_{\frac{1}{2}}$ level. Taken from [176].

2.2 Quantum trajectories

As stated in the introduction, the dynamics of open quantum systems is often studied with the help of master equations for the reduced density matrix. Another powerful approach has been developed in the nineties. It bears several names, introduced by the various physicists and mathematicians who participated to its (re)discovery: Monte Carlo wave function [69], quantum trajectories [51], quantum state diffusion [101], quantum unraveling [46], or quantum filtering [30, 25], to cite only a few[2]. It consists in describing the system state by a random pure state $|\psi_{\mathsf{S}}(t)\rangle$, the average of which gives the density operator $\rho_{\mathsf{S}}(t) = \overline{|\psi_{\mathsf{S}}(t)\rangle\langle\psi_{\mathsf{S}}(t)|}$ solution of the master equation[3]. The wave function $|\psi_{\mathsf{S}}(t)\rangle$ undergoes a stochastic time evolution caused by continuous measurements on the bath coupled to the system S. This method is now popular in quantum optics, in particular because it provides an efficient way to compute the density matrix $\rho_{\mathsf{S}}(t)$ numerically[4]. However, quantum trajectories are more than simple mathematical or numerical tools: they actually describe the *real* dynamics of S when repeated measurements on the environment with which S interacts are performed. Thanks to improvements in detection techniques in the last decades, experiments at very low temperature have succeeded to observe these trajectories. For instance, the electromagnetic microwave field inside the optical cavity in Haroche's group is coupled to highly excited Rydberg atoms, which cross one-by-one the cavity [113]. The states of these atoms, which constitute the "environment" with which the field interacts, are measured at the exit of the cavity. When the atoms and field are off-resonance, this leads to non-demolition measurements and random jumps between states with different numbers of photons in the cavity can be observed [109]. Another older beautiful experiment, in which the role of atom and field are inverted, is the observation shown in Fig. 2.1 of a random succession of dark and bright periods in the fluorescence signal of a single trapped ion. This telegraphic signal is a direct manifestation of sudden quantum jumps of the ion to a metastable state [176].

[2]This approach has also been introduced in Ref. [99] with the aim to modify the Schrödinger equation in such a way that it includes the collapse of the wave packet. As discussed in the previous chapter, we think that this is meaningless, but Ref. [99] was perhaps the first work in the physics literature giving a detailed account of the quantum trajectory method. The discovery of this method has, however, to be attributed to the mathematician V.P. Belavkin.

[3]In mathematical works, the pure state $|\Psi_{\mathsf{S}}(t)\rangle$ is often replaced with more general random mixed states. This generalization is, however, not so interesting from a numerical viewpoint.

[4]For a system with a n-dimensional Hilbert space, solving the master equation amounts to solve $n \times n$ coupled differential equations, whereas finding the random wave function requires solving n stochastic differential equations. To estimate the average density matrix, the computation must be repeated m times with different noise realizations, but m can be smaller than n for high-dimensional spaces. For a review of the numerical methods in the quantum trajectory approach, see [46], chapter 7.

2.2.1 Quantum jump model

The stochastic dynamics of the system wave function due to the monitoring by measurements on the reservoir can be derived from the system-reservoir dynamics and the reduction of the wave packet. Let us consider a simple model similar to the model of a single atom interacting with the vacuum of the electromagnetic field discussed in the seminal paper of Dalibard, Castin, and Mölmer [69] (for a more general derivation, see the book of H.-P. Breuer and F. Petruccione [46], chapter 6). A system S with Hamiltonian H_S is coupled to a reservoir R with Hamiltonian H_R by the interaction

$$H_{\rm int}^{\mathsf{SR}} = \sum_m g_m \big(M_m \otimes b_m^\dagger + M_m^\dagger \otimes b_m \big) \ , \tag{2.1}$$

where g_m are coupling constants, M_m are operators acting on S, and $b_m^\dagger$, b_m are the creation and annihilation operators for an excitation m in the reservoir. We assume that R is initially in its ground state $|0\rangle$ (zero temperature) and choose the zero of energy such that $H_\mathsf{R}|0\rangle = 0$. At times $t_k = k\delta t$ (with k integer), a detector measures if R is excited or not. We assume that the detection is instantaneous, i.e., the time duration of each measurement is much smaller than δt. Furthermore, we suppose that δt is such that $g_m \delta t$ is very small but $g_m^2 \delta t = \gamma_m$ is of the order of T^{-1}, T being the relevant time scale for the dynamics of S. Mathematically, this corresponds to letting $\delta t \to 0$ with γ_m fixed.

We denote by $|1_m\rangle = b_m^\dagger|0\rangle$ and $|1_m 1_{m'}\rangle = b_m^\dagger b_{m'}^\dagger|0\rangle$ the reservoir states with one excitation m and two excitations m and m', respectively. If the bipartite system SR is in a pure product state[5] $|\psi_\mathsf{S}(t)\rangle|0\rangle$ at time t, then this state is at time $t+\delta t$:

$$\begin{aligned} |\Psi_{\mathsf{SR}}(t+\delta t)\rangle \quad &= \quad \mathrm{e}^{-\mathrm{i}\delta t (H_\mathsf{S} + H_\mathsf{R} + H_{\rm int}^{\mathsf{SR}})} |\psi_\mathsf{S}(t)\rangle|0\rangle = \Big[(1 - \mathrm{i}\delta t H_{\rm eff})|\psi_\mathsf{S}(t)\rangle|0\rangle - \mathrm{i}\delta t \sum_m g_m M_m |\psi_\mathsf{S}(t)\rangle|1_m\rangle \\ &\quad - \frac{\delta t^2}{2} \sum_{m,m'} g_m g_{m'} M_m M_{m'} |\psi_\mathsf{S}(t)\rangle|1_m 1_{m'}\rangle \Big] \Big(1 + \mathcal{O}(g_m \delta t + \delta t/T_\mathsf{S} + \delta t/T_\mathsf{R}) \Big) \ , \end{aligned} \tag{2.2}$$

where T_S and T_R are the typical evolution times under the Hamiltonians H_S and H_R and we have set

$$H_{\rm eff} = H_\mathsf{S} - \frac{\mathrm{i}}{2} \sum_m \gamma_m M_m^\dagger M_m \ . \tag{2.3}$$

The measurement on the reservoir at time $t+\delta t$ causes the wave packet reduction given by the transformation (1.5). The corresponding conditional states are

$$|\Psi_{\mathsf{SR}|0}(t+\delta t)\rangle = \frac{1_\mathsf{S} \otimes |0\rangle\langle 0|\Psi_{\mathsf{SR}}(t+\delta t)\rangle}{\|\langle 0|\Psi_{\mathsf{SR}}(t+\delta t)\rangle\|} \ \propto \ (1 - \mathrm{i}\delta t H_{\rm eff})|\psi_\mathsf{S}(t)\rangle|0\rangle \tag{2.4}$$

if no excitation is detected and

$$|\Psi_{\mathsf{SR}|m}(t+\delta t)\rangle = \frac{1_\mathsf{S} \otimes |1_m\rangle\langle 1_m|\Psi_{\mathsf{SR}}(t+\delta t)\rangle}{\|\langle 1_m|\Psi_{\mathsf{SR}}(t+\delta t)\rangle\|} \ \propto \ M_m|\psi_\mathsf{S}(t\rangle|1_m\rangle \tag{2.5}$$

if an excitation m is detected. In the latter case, we assume that the excitation is removed from the reservoir (e.g. the photon disappears in the photo-detector). Equivalently, one may think that the system interacts in each time interval $[t_k, t_k + \delta t]$ with a different reservoir initially in the state $|0\rangle$ (repeated interaction model). This situation is realized for instance in Haroche's experiments, where the cavity field (system) interacts successively with different atoms (reservoirs) crossing the cavity. Consequently, the final system-reservoir state after a detection of an excitation m at time $t+\delta t$ is $M_m|\psi_\mathsf{S}(t\rangle|0\rangle$, up to normalization. The probability of this detection is very small and equal, according to the Born rules, to

$$\delta p_m(t) = \|\langle 1_m|\Psi_{\mathsf{SR}}(t+\delta t)\rangle\|^2 = \gamma_m \|M_m|\psi_\mathsf{S}(t)\rangle\|^2 \delta t \ . \tag{2.6}$$

The probability to detect more than one excitation is negligible in the limit $g_m \delta t \ll 1$. Each measurement disentangles the system and reservoir, so that SR is in a product state at all times t_k. In the limit $\delta t \to 0$, one can replace the discrete times t_k by a continuous time t and write δt as a differential $\mathrm{d}t$. The system S is then in a pure state $|\psi_\mathsf{S}(t)\rangle$ at all times.

The random dynamics of S can be summarized as follows.

[5] As is common in the physics literature, we omit the symbol $\otimes$ in tensor products of vectors in Hilbert spaces.

(i) A quantum jump

$$|\psi_{\mathsf{S}}(t)\rangle \;\longrightarrow\; |\psi_{\mathsf{S}}^{(m)}(t+\mathrm{d}t)\rangle = \frac{M_m|\psi_{\mathsf{S}}(t)\rangle}{\|M_m|\psi_{\mathsf{S}}(t)\rangle\|} \tag{2.7}$$

occurs with probability $\mathrm{d}p_m(t)$ between times t and $t+\mathrm{d}t$.

(ii) If no jump occurs between t and $t+\mathrm{d}t$ (i.e., no excitation is detected in R), the system wave function evolves as

$$|\psi_{\mathsf{S}}(t+\mathrm{d}t)\rangle = \frac{e^{-\mathrm{i}H_{\mathrm{eff}}\mathrm{d}t}|\psi_{\mathsf{S}}(t)\rangle}{\|e^{-\mathrm{i}H_{\mathrm{eff}}\mathrm{d}t}|\psi_{\mathsf{S}}(t)\rangle\|} \;. \tag{2.8}$$

The probability to have no jump in the finite time interval $[t_0,t]$ is[6]

$$p_{\mathrm{nj}}(t_0,t) = \big\|e^{-\mathrm{i}H_{\mathrm{eff}}(t-t_0)}|\psi_{\mathsf{S}}(t_0)\rangle\big\|^2 \;. \tag{2.9}$$

The above heuristic derivation of the quantum jump model can be made rigorous [18] (see also [16] and [77] for related derivations of evolutions with quantum noises).

2.2.2 Lindblad master equation for the average state

The density matrix of S is obtained by averaging the rank-one projector on $|\psi_{\mathsf{S}}(t)\rangle$ over the random outcomes of the measurements on R,

$$\rho_{\mathsf{S}}(t) = \overline{|\psi_{\mathsf{S}}(t)\rangle\langle\psi_{\mathsf{S}}(t)|} = \int \mathrm{d}p[\psi_{\mathsf{S}}]\,|\psi_{\mathsf{S}}(t)\rangle\langle\psi_{\mathsf{S}}(t)| \;, \tag{2.10}$$

where the overline denotes the mean over all quantum trajectories with distribution $\mathrm{d}p[\psi_{\mathsf{S}}]$. Since a measurement without readout performed on R does not change the reduced state of S, $\rho_{\mathsf{S}}(t)$ describes the state evolution of the system coupled to R in the absence of measurements[7]. A simple calculation (see e.g. [69]) shows that $\rho_{\mathsf{S}}(t)$ satisfies the Lindblad Markovian master equation [159]

$$\frac{\mathrm{d}\rho_{\mathsf{S}}}{\mathrm{d}t} = -\mathrm{i}[H_{\mathsf{S}},\rho_{\mathsf{S}}(t)] + \sum_m \gamma_m\Big(M_m\rho_{\mathsf{S}}(t)M_m^\dagger - \frac{1}{2}\big\{M_m^\dagger M_m,\rho_{\mathsf{S}}(t)\big\}\Big) \;, \tag{2.11}$$

where $\{\cdot,\cdot\}$ denotes the anti-commutator. Let us emphasize here that the derivation of the quantum jump dynamics in the previous subsection also justifies the master equation (2.11) for a zero-temperature reservoir coupled to S via the Hamiltonian (2.1). This derivation involves a re-initialization of the reservoir state at times t_k, which is a "brute force" manner to obtain a Markovian dynamics (equivalently, one may think of a chain of reservoirs interacting one after the other with S, as considered in [16, 17]). Physically, the state re-initialization is due to the relaxation of R to its equilibrium and takes a time larger than the reservoir correlation time τ_{R} defined in Sec. 1.3.1. Hence one must have $\tau_{\mathsf{R}} \ll \delta t \ll T_{\mathsf{S}}$ with $g_m^2\delta t = \gamma_m$ fixed. This limit is the singular coupling limit discussed in Sec. 1.3.5.

It is worth noting that (2.10) defines a particular convex decomposition of the density matrix $\rho_{\mathsf{S}}(t)$. Since there are infinitely many such decompositions, it should not come as a surprise that infinitely many different quantum jump dynamics unravel the same master equation [46]. Each of these unravelings corresponds to a specific measurement scheme on R and can be viewed as a sequence of state preparations associated to $\rho_{\mathsf{S}}(t_k)$ at times t_k (Sec. 1.2.2). These state preparations depend upon the type of information on the system dynamics collected by the measurements. For instance, if one rotates the measurement basis, the jump operators M_m are transformed as

$$M_m \;\longrightarrow\; M_\mu = \sum_m u_{\mu m}\left(\frac{\gamma_m}{\gamma_\mu}\right)^{\frac{1}{2}} M_m \;, \tag{2.12}$$

where $(u_{\mu m})$ is a unitary matrix and γ_μ are the jump rates associated to M_μ. One may check explicitly that this transformation does not change the master equation. Another transformation with this property is

$$M_m \;\longrightarrow\; M_{m,\pm\alpha_m} = M_m \pm \alpha_m \quad,\quad \gamma_m \;\to\; \gamma_{m,\pm\alpha_m} = \frac{\gamma_m}{2} \;. \tag{2.13}$$

[6] This is shown by noting that $p_{\mathrm{nj}}(t_0,t) - p_{\mathrm{nj}}(t_0,t+\mathrm{d}t) = \sum_m \mathrm{d}p_m(t)p_{\mathrm{nj}}(t_0,t)$. Thanks to (2.3) and (2.6), this yields

$$\frac{\partial \ln p_{\mathrm{nj}}(t_0,t)}{\partial t} = -\sum_m \gamma_m\|M_m|\psi_{\mathsf{S}}(t)\rangle\|^2 = \frac{\partial}{\partial t}\ln\|e^{-\mathrm{i}H_{\mathrm{eff}}(t-t_0)}|\psi_{\mathsf{S}}(t_0)\rangle\|^2 \;.$$

[7] This general statement can be established by noting that if the measured observable of R has eigenprojectors Π_m, then, according to (1.3), the post-measurement state $\rho_{\mathsf{SR}}^{\mathrm{p.m.}} = \sum_m 1\otimes\Pi_m\rho_{\mathsf{SR}}\,1\otimes\Pi_m$ has the same partial trace over R than ρ_{SR}.

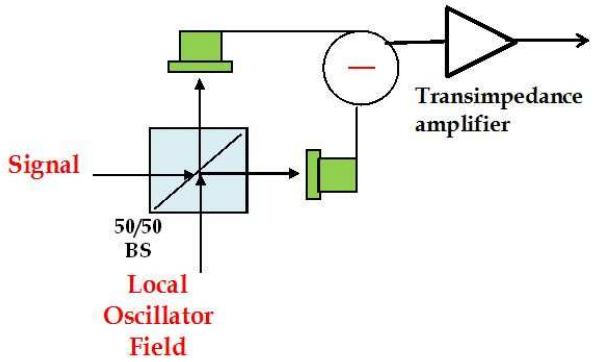

Figure 2.2: In homodyne or heterodyne detections, the signal (e.g. the photons emitted by an atom) is mixed in a 50% beam splitter (BS) with a classical laser field with amplitude α (local oscillator). Two photo-detectors count the number of photons in the two output beams.

If the system is an atom emitting photons and M_m implements jumps provoked by the detection of the emitted photons by a photon-counting device, the new jump operators $M_{m,\pm\alpha_m}$ describe homodyne or heterodyne detection (see Fig. 2.2) used in many experiments in quantum optics [51, 199].

2.2.3 Quantum state diffusion

The homodyne quantum jump dynamics with jump operators $M_{m,\pm\alpha_m}$ converges for large positive time-independent laser amplitudes $\alpha_m \gg 1$ to a diffusive dynamics given by Wiener processes. More precisely, after a coarse graining on a time scale Δt on which many jumps occur but $|\psi_{\mathsf{S}}(t)\rangle$ does not change noticeably, the wave function $|\psi_{\mathsf{S}}(t)\rangle$ converges to the solution of the stochastic Schrödinger equation [261, 46, 222]

$$|\mathrm{d}\psi_{\mathsf{S}}\rangle = \Big(-\mathrm{i}H_{\rm eff}\mathrm{d}t + \sum_m \gamma_m\Big(\mathrm{Re}\,\langle M_m\rangle_t\, M_m - \frac{1}{2}\big(\mathrm{Re}\,\langle M_m\rangle_t\big)^2\Big)\mathrm{d}t + \sum_m \sqrt{\gamma_m}\big(M_m - \mathrm{Re}\,\langle M_m\rangle_t\big)\mathrm{d}w_m\Big)|\psi_{\mathsf{S}}(t)\rangle\ , \tag{2.14}$$

where $\mathrm{d}w_m$ are the Itô differentials for independent real Wiener processes satisfying the Itô rules $\mathrm{d}w_m\mathrm{d}w_n = \delta_{mn}\mathrm{d}t$. We have set $\langle\cdot\rangle_t = \langle\psi_{\mathsf{S}}(t)|\cdot|\psi_{\mathsf{S}}(t)\rangle$. The quantum state diffusion dynamics defined by (2.14) is another unraveling of the master equation (2.10).

In the case of heterodyne detection, the laser amplitudes have oscillating phases and the jump operators $M_{m,\pm\alpha}(t_j) = M_m \pm \alpha_m e^{\mathrm{i}\Omega_m t_j}$ depend on the time t_j of the jth jump [199, 46]. The associated rates are as for homodyne detection. In the limit $\alpha_m^2 \gg \Omega_m/\gamma_m \gg 1$ of large laser intensities and rapidly oscillating amplitudes, with $\alpha_m > 0$, the quantum jump dynamics converges to the solution of [261, 46]

$$\begin{aligned}|\mathrm{d}\psi_{\mathsf{S}}\rangle \;=\; \Big(&-\mathrm{i}H_{\rm eff}\,\mathrm{d}t + \frac{1}{2}\sum_m \gamma_m\Big(\langle M_m\rangle_t^*\, M_m - \frac{1}{2}\big|\langle M_m\rangle_t\big|^2\Big)\mathrm{d}t \\ &+ \sum_m \sqrt{\gamma_m}\Big(\Big(M_m - \frac{1}{2}\langle M_m\rangle_t\Big)\mathrm{d}\xi_m - \frac{1}{2}\langle M_m\rangle_t^*\mathrm{d}\xi_m^*\Big)\Big)|\psi_{\mathsf{S}}(t)\rangle\ ,\end{aligned} \tag{2.15}$$

where $\mathrm{d}\xi_m$ are the Itô differential for independent complex Wiener processes satisfying the Itô rules $\mathrm{d}\xi_m\mathrm{d}\xi_n = 0$ and $\mathrm{d}\xi_m\mathrm{d}\xi_n^* = \delta_{mn}\mathrm{d}t$. The stochastic Schrödinger equation (2.15) describes the coarse-grained evolution of $|\psi_{\mathsf{S}}(t)\rangle$ on a time scale Δt on which $|\psi_{\mathsf{S}}(t)\rangle$ does not change significantly and such that many jumps and many laser amplitude oscillations occur during a lapse of time Δt. These conditions are satisfied when $\alpha_m^2\gamma_m\Delta t \gg \Omega_m\Delta t \gg 1$ and $\gamma_m\Delta t \ll 1$ (see [46] for more detail). Although (2.15) is not exactly the quantum state diffusion equation introduced originally by Gisin and Percival [101], the solutions of the two equations differ by an irrelevant randomly fluctuating phase factor [46]. More general equations involving correlated complex noises satisfying the Itô rules $\mathrm{d}\xi_m\mathrm{d}\xi_n = u_{mn}\mathrm{d}t$ and $\mathrm{d}\xi_m\mathrm{d}\xi_n^* = \delta_{mn}\mathrm{d}t$ have been considered in Ref. [262]. They give back the model of Gisin and Percival when $u_{mn} = 0$.

Linear stochastic evolutions can be obtained from the state diffusion and quantum jump dynamics by relaxing the condition that $|\psi_{\mathsf{S}}(t)\rangle$ must be normalized at all times. The corresponding linear equations involve Wiener

processes for diffusion [25, 99, 222] and Poisson processes for quantum jumps [221, 222]. The average state (2.10) is the same as for the nonlinear dynamics, thus it is normalized at all times.

Let us also mention that the quantum trajectory approach can be extended to non-Markovian dynamics [47, 80, 267] and that models accounting for imperfect measurements on R can be derived.

2.3 Entanglement decay in bipartite systems

We give in this section a brief introduction to the notion of entangled mixed states, which plays an important role in the remaining of this chapter. A more detailed exposition can be found in Sec. 4.4 and chapter 10 below. We then digress on the entanglement evolution when the bipartite system is coupled to its environment.

2.3.1 Bipartite entanglement

Let us recall that a pure state $|\Psi\rangle$ of a bipartite system AB with Hilbert space $\mathcal{H}_{\mathsf{AB}} = \mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}}$ is entangled if it cannot be written as a tensor product of two pure states $|\psi\rangle \in \mathcal{H}_{\mathsf{A}}$ and $|\phi\rangle \in \mathcal{H}_{\mathsf{B}}$. This property is not related to a specific observable. Rather, it reflects the fact that *any* local observables $A \otimes 1$ and $1 \otimes B$ acting respectively on subsystems A and B are uncorrelated, that is, $\langle A \otimes B\rangle_\Psi = \langle A \otimes 1\rangle_\Psi \langle 1 \otimes B\rangle_\Psi$. The notion of entanglement is non trivial when dealing with statistical mixtures because the quantum correlations between $A \otimes 1$ and $1 \otimes B$ must be separated from statistical correlations in the state ensemble, which do not correspond to any entanglement. Consequently, a mixed state ρ_{AB} may be separable even if it is not a product of two density operators ρ_{A} and ρ_{B} on $\mathcal{H}_{\mathsf{A}}$ and $\mathcal{H}_{\mathsf{B}}$.

To see if ρ_{AB} is separable, one must consider *all* possible state preparations associated to ρ_{AB} and look if there exists one such preparation formed by pure separable states only. Actually, we want to consider as separable any state that can be prepared locally, allowing for classical communication between the two parties. More precisely, the state preparation can be achieved by two observers (called Alice and Bob) allowed to exchange classical information and separated by a large distance to eliminate interactions between A and B. Alice prepares identical subsystems A with a fraction p_i of them in state $|\psi_i\rangle \in \mathcal{H}_{\mathsf{A}}$ and, when her state is $|\psi_i\rangle$, tells Bob to prepare the corresponding subsystem B in state $|\phi_i\rangle \in \mathcal{H}_{\mathsf{B}}$. The two observers thereby prepare identical systems AB in product states $|\psi_i\rangle|\phi_i\rangle$ with probability p_i. In mathematical terms, ρ_{AB} is separable if it admits a pure state decomposition $\rho_{\mathsf{AB}} = \sum_i \eta_i |\psi_i\rangle\langle\psi_i| \otimes |\phi_i\rangle\langle\phi_i|$ with $\|\psi_i\| = \|\phi_i\| = 1$, $\eta_i \geq 0$, and $\sum_i \eta_i = 1$. If such a decomposition does not exist, the state is entangled.

The Peres-Horodecki condition gives a simple sufficient criterion for entanglement, which turns out to be necessary and sufficient if the dimensions n_{A} of $\mathcal{H}_{\mathsf{A}}$ and n_{B} of $\mathcal{H}_{\mathsf{B}}$ satisfy $n_{\mathsf{A}} n_{\mathsf{B}} \leq 6$. We recall that the partial transpose $\rho_{\mathsf{AB}}^{T_{\mathsf{B}}}$ of ρ_{AB} in a given product basis $\{|i\rangle|k\rangle\}$ of $\mathcal{H}_{\mathsf{AB}}$ is defined as

$$\langle i|\langle k|\rho_{\mathsf{AB}}^{T_{\mathsf{B}}}|j\rangle|l\rangle = \langle i|\langle l|\rho_{\mathsf{AB}}|j\rangle|k\rangle \; . \tag{2.16}$$

The Peres criterion states that if $\rho_{\mathsf{AB}}^{T_{\mathsf{B}}}$ has negative eigenvalues (i.e., if $\rho_{\mathsf{AB}}^{T_{\mathsf{B}}}$ is not a physical state) then ρ_{AB} is entangled. When $n_{\mathsf{A}} n_{\mathsf{B}} > 6$, there exists entangled states of AB which do not fulfill this criterion. In general, finding out whether a given state is entangled or separable is not an easy task: a simple-to-handle necessary and sufficient condition for bipartite mixed state entanglement is still lacking [130].

In the following, we denote by $\mathcal{E}(\mathcal{H}_{\mathsf{AB}})$ the set of all quantum states of AB, that is, the convex cone formed by all non-negative operators on $\mathcal{H}_{\mathsf{AB}}$ with trace one (we assume here that $\mathcal{H}_{\mathsf{AB}}$ has finite dimension $n_{\mathsf{A}} n_{\mathsf{B}}$). By definition, the separable states of AB form a convex subset $\mathcal{S}_{\mathsf{AB}}$ of $\mathcal{E}(\mathcal{H}_{\mathsf{AB}})$. We say that a state ρ_{AB} lies on the boundary $\partial\mathcal{S}_{\mathsf{AB}}$ of $\mathcal{S}_{\mathsf{AB}}$ when ρ_{AB} is separable and an arbitrarily small perturbation can make it entangled. This is the case e.g. for pure product states[8].

2.3.2 Two-qubit states with maximally mixed marginals

To gain some insight on how the convex set of separable states $\mathcal{S}_{\mathsf{AB}}$ looks like, let us focus on two qubits A and B (i.e., $\mathcal{H}_{\mathsf{A}} \simeq \mathcal{H}_{\mathsf{B}} \simeq \mathbb{C}^2$). An arbitrary trace-one self-adjoint operator on $\mathcal{H}_{\mathsf{AB}} \simeq \mathbb{C}^4$ can be written as

$$\rho = \frac{1}{4}\Big(1_{\mathsf{A}} \otimes 1_{\mathsf{B}} + \mathbf{u}\cdot\boldsymbol{\sigma} \otimes 1_{\mathsf{B}} + 1_{\mathsf{A}} \otimes \mathbf{v}\cdot\boldsymbol{\sigma} + \sum_{a,b=1}^{3} c_{ab}\, \sigma_a \otimes \sigma_b\Big) \; , \tag{2.17}$$

[8] More precisely, $\rho \in \partial\mathcal{S}_{\mathsf{AB}}$ if and only if $\rho \in \mathcal{S}_{\mathsf{AB}}$ and $\forall\, \varepsilon > 0$, $\exists\, \rho_\varepsilon \in \mathcal{E}(\mathcal{H}_{\mathsf{AB}})$ such that $\mathrm{tr}\,|\rho - \rho_\varepsilon| \leq \varepsilon$ and $\rho_\varepsilon \notin \mathcal{S}_{\mathsf{AB}}$. To show that a pure product state $\rho = |\psi\rangle\langle\psi| \otimes |\phi\rangle\langle\phi|$ belongs to $\partial\mathcal{S}_{\mathsf{AB}}$, let us set $|\Psi_\varepsilon\rangle = (1+\varepsilon^2)^{-1/2}(|\psi\rangle|\phi\rangle + \varepsilon|\psi_\perp\rangle|\phi_\perp\rangle)$, where $|\psi_\perp\rangle$ and $|\phi_\perp\rangle$ are orthogonal states to $|\psi\rangle$ and $|\phi\rangle$. Then $\rho_\varepsilon = |\Psi_\varepsilon\rangle\langle\Psi_\varepsilon|$ is entangled and $\mathrm{tr}\,|\rho - \rho_\varepsilon| \leq 2\varepsilon$.

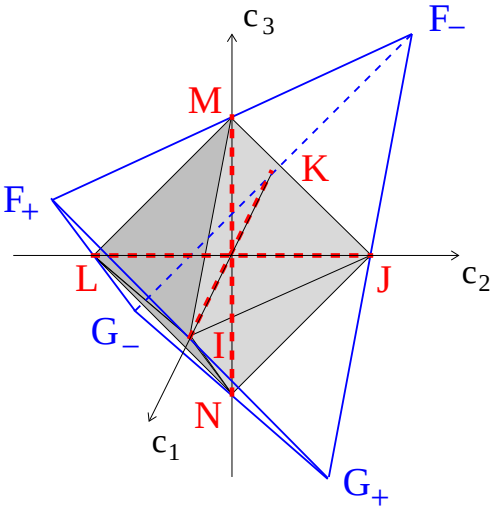

Figure 2.3: Tetrahedron giving the location of the vectors $\mathbf{c}$ of $\mathbb{R}^3$ representing two-qubit states with maximally mixed marginals up to local unitary conjugations. Separable states are located inside the octahedron with vertices I, J, K, L, M, and N. Among separable states, those belonging to the segments $[I, K]$, $[J, L]$, and $[M, N]$ (red dashed lines) are classical states with zero quantum discord (see chapter 11), as shown in Ref. [67].

where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is the vector formed by the three Pauli matrices acting on either $\mathcal{H}_\mathsf{A}$ or $\mathcal{H}_\mathsf{B}$. In order that ρ be a density operator (i.e., that $\rho \geq 0$), the expectation values $\mathbf{u} = \mathrm{tr}[\boldsymbol{\sigma} \otimes 1\, \rho]$ and $\mathbf{v} = \mathrm{tr}[1 \otimes \boldsymbol{\sigma}\ \rho]$ of the spins of A and B and the 3×3 spin-spin correlation matrix $C = (c_{ab}) = (\mathrm{tr}[\sigma_a \otimes \sigma_b\, \rho])$ must fulfill certain conditions (see e.g. Proposition 2.3.4 in [252]). Given two rotations R_A and R_B of $\mathbb{R}^3$, the transformation

$$\mathbf{u} \to R_\mathsf{A}\mathbf{u} \quad , \quad \mathbf{v} \to R_\mathsf{B}\mathbf{v} \quad , \quad C \to R_\mathsf{A} C R_\mathsf{B}^T \tag{2.18}$$

of the spin vectors and correlation matrix amounts to a local unitary conjugation[9] $\rho \to U_A \otimes U_B \rho\, U_A^\dagger \otimes U_B^\dagger$ which does not change the separability property of ρ. By a singular value decomposition of C, one may thus assume that the correlation matrix is diagonal, $C = \mathrm{diag}(c_1, c_2, c_3)$.

We now specialize our discussion to the family of two-qubit states ρ with maximally mixed marginals $\mathrm{tr}_\mathsf{B}(\rho) = \mathrm{tr}_\mathsf{A}(\rho) = 1/2$, which are given by (2.17) with $\mathbf{u} = \mathbf{v} = \mathbf{0}$ [129]. For diagonal C, one has $\rho \geq 0$ if and only if

$$p_0 = \frac{1}{4}(1 - c_1 - c_2 - c_3) \geq 0 \quad , \quad p_a = \frac{1}{4}(1 + c_1 + c_2 + c_3 - 2c_a) \geq 0 \quad , \quad a = 1, 2, 3\,. \tag{2.19}$$

Equivalently, $\mathbf{c} = (c_1, c_2, c_3)$ belongs to the tetrahedron $\mathscr{T}$ with vertices $F_\pm = (\pm 1, \mp 1, 1)$, $G_\pm = (\pm 1, \pm 1, -1)$ associated to the pure states

$$|\Phi^\pm\rangle = \frac{1}{\sqrt{2}}\big(|\uparrow\uparrow\rangle \pm |\downarrow\downarrow\rangle\big) \quad , \quad |\Psi^\pm\rangle = \frac{1}{\sqrt{2}}\big(|\uparrow\downarrow\rangle \pm |\downarrow\uparrow\rangle\big)\,, \tag{2.20}$$

see Fig. 2.3. The states (2.20) are the four Bell states, which are maximally entangled (see below)[10]. They are the eigenvectors of ρ with eigenvalues $p_{2,1}$ and $p_{3,0}$. Thanks to the Peres-Horodecki criterion, ρ is separable if and only if both vectors $\mathbf{c}$ and $\widetilde{\mathbf{c}} = (c_1, -c_2, c_3)$, corresponding respectively to ρ and ρ^{T_B}, belong to $\mathscr{T}$. The convex set of separable Bell diagonal states is thus the octahedron represented in Fig. 2.3.

To illustrate the definition of entangled states given above, let us consider a preparation of two Bell states $|\Psi^\pm\rangle$ with equal probability $p_\pm = 1/2$. Even though $|\Psi^\pm\rangle$ are maximally entangled, the associated density matrix $\rho = (|\Psi^+\rangle\langle\Psi^+| + |\Psi^-\rangle\langle\Psi^-|)/2$ is separable, since it could also have been obtained by preparing two equiprobable separable states $|\downarrow\downarrow\rangle$ and $|\uparrow\uparrow\rangle$.

2.3.3 Entanglement measures

A well known measure quantifying the amount of entanglement in bipartite systems AB is the *entanglement of formation* E_EoF defined as follows [34]. For a pure state, $E_\mathrm{EoF}(|\Psi\rangle) = S(\rho_\mathsf{A})$ is equal to the von Neumann

[9] U_A and U_B are the images of R_A and R_B under the $SU(2)$-representation of the rotation group.

[10] These states have been used e.g. in Ref. [15] to demonstrate experimentally the violation of the Bell inequalities for photon polarization states.

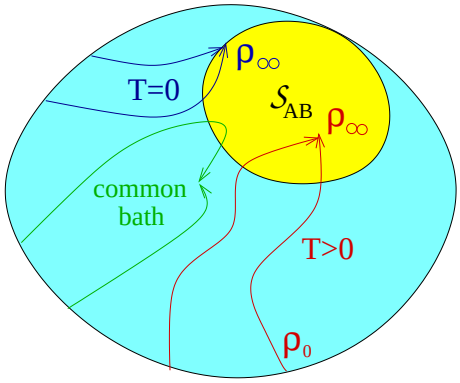

Figure 2.4: Schematic view of the trajectories $t \mapsto \rho(t)$ in the set of quantum states (cyan region) for a bipartite system undergoing an irreversible evolution. An entanglement sudden death or revival occurs when the trajectory crosses the boundary of the set of separable states $\mathcal{S}_{\mathsf{AB}}$ (yellow region).

entropy of the reduced state[11] $\rho_{\mathsf{A}} = \mathrm{tr}_{\mathsf{B}}(|\Psi\rangle\langle\Psi|)$. For a mixed state, $E_{\mathrm{EoF}}(\rho)$ is defined as the minimum of the average entanglement over all ensembles $\{|\Psi_i\rangle, \eta_i\}$ forming a pure state decomposition of ρ, $\rho = \sum_i \eta_i |\Psi_i\rangle\langle\Psi_i|$,

$$E_{\mathrm{EoF}}(\rho) = \min_{\{|\Psi_i\rangle,\eta_i\}} \left\{ \sum_i \eta_i E_{\mathrm{EoF}}(|\Psi_i\rangle) \right\} . \tag{2.21}$$

Note that ρ is separable if and only if $E_{\mathrm{EoF}}(\rho) = 0$. The maximal value of E_{EoF} is $\ln n$ with $n = \min\{n_{\mathsf{A}}, n_{\mathsf{B}}\}$. It is achieved for instance for the two-qubit Bell states (2.20). When $n_{\mathsf{A}} = n_{\mathsf{B}} = 2$, Wootters has shown that $E_{\mathrm{EoF}}(\rho) = h(C(\rho))$, where $h : [0,1] \to [0, \ln n]$ is a convex increasing function (which is explicitly known, see Sec. 10.4.3) and $C(\rho)$ is the *concurrence* given by [265]

$$C(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\} , \tag{2.22}$$

$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$ being the square roots of the eigenvalues of $\rho\, \sigma_2 \otimes \sigma_2\, J\rho J\, \sigma_2 \otimes \sigma_2$. Here, J stands for the anti-unitary operator of complex conjugation in the canonical basis $\{|k\rangle|l\rangle\}_{k,l=\uparrow,\downarrow}$ of $\mathbb{C}^2 \otimes \mathbb{C}^2$. For pure states, the concurrence reads

$$C(\Psi) = |c(\Psi)| \quad , \quad c(\Psi) = \langle\Psi|\sigma_2 \otimes \sigma_2 J|\Psi\rangle = 2(c_{\uparrow\downarrow}c_{\downarrow\uparrow} - c_{\uparrow\uparrow}c_{\downarrow\downarrow})^* \tag{2.23}$$

with $c_{kl} = \langle k|\langle l|\Psi\rangle$ for any $k, l = \uparrow, \downarrow$. The properties of the entanglement of formation and concurrence will be discussed in more detail in chapter 10.

2.3.4 Different scenarios for entanglement evolution

Let us discuss the time evolution of entanglement when two initially entangled subsystems A and B are coupled to independent baths. We assume that A and B do not interact directly, so that the Hamiltonian of AB has the form

$$H_{\mathsf{AB}} = H_{\mathsf{A}} \otimes 1 + 1 \otimes H_{\mathsf{B}} . \tag{2.24}$$

We first suppose that AB converges at large times to a unique stationary separable state ρ_∞. By continuity of the evolution, we conclude that (see Fig. 2.4):

(1) if ρ_∞ belongs to the interior of the set of separable states $\mathcal{S}_{\mathsf{AB}}$, then so does $\rho(t)$ at large enough times. Thus the initial entanglement between A and B disappears completely after a finite time t_{ESD} (entanglement sudden death) [79, 81, 234];

(2) if ρ_∞ belongs to the boundary of $\mathcal{S}_{\mathsf{AB}}$, then either entanglement disappears after a finite time as in case (1), or $\rho(t)$ remains entangled at all times (that is, separability is reached asymptotically). The entanglement behavior depends on the initial state $\rho(0)$ and on the coupling with the bath [268, 8].

Case (1) is realized for instance if AB relaxes to an equilibrium state ρ_{eq} at positive temperature, and case (2) occurs at zero temperature when AB decays to a non-degenerate ground state of the Hamiltonian (2.24)[12]. For

[11] The same result is obtained for the entropy of the reduced state $\rho_{\mathsf{B}} = \mathrm{tr}_{\mathsf{A}}(|\Psi\rangle\langle\Psi|)$ of B, see Sec. 7.1.

[12] Indeed, the Gibbs state $\rho_{\mathrm{eq}} = Z^{-1}\mathrm{e}^{-\beta H_{\mathsf{AB}}} = Z^{-1}\mathrm{e}^{-\beta H_{\mathsf{A}}} \otimes \mathrm{e}^{-\beta H_{\mathsf{B}}}$ is in the interior of $\mathcal{S}_{\mathsf{AB}}$, whereas the ground state of (2.24) is a pure product state and thus belongs to $\partial\mathcal{S}_{\mathsf{AB}}$, see Sec. 2.3.1.

instance, let us consider the Jaynes-Cummings model for a pair of two-level atoms A and B coupled resonantly to independent modes of the electromagnetic field. Entanglement evolution in this model has been considered at zero temperature in the first paper of Yu and Eberly [268]. After a dipole, weak coupling, and rotating wave approximations [65, 185], one finds that the dynamics of the atoms is governed by the Lindblad equation (2.11) with jump operators

$$M_{\pm}^{\mathsf{A}} = \sigma_{\pm} \otimes 1_{\mathsf{B}} \quad , \quad M_{\pm}^{\mathsf{B}} = 1_{\mathsf{A}} \otimes \sigma_{\pm} \tag{2.25}$$

and Hamiltonian (2.24) given by $H_{\mathsf{A}} = H_{\mathsf{B}} = \omega\, \sigma_3$ (here ω is the atomic Bohr frequency, assumed to be the same for both atoms, $\sigma_+ = (\sigma_1 + \mathrm{i}\sigma_2)/2 = |\uparrow\rangle\langle\downarrow|$, and $\sigma_- = \sigma_+^\dagger$). If the jump rates satisfy $0 < \gamma_+^{\mathsf{A}} = \gamma_+^{\mathsf{B}} \leq \gamma_-^{\mathsf{A}} = \gamma_-^{\mathsf{B}}$, an explicit calculation shows that $\rho(t)$ converges to the Gibbs state ρ_{eq} at inverse temperature $\beta = \omega^{-1} \ln(\gamma_-^{\mathsf{A}}/\gamma_+^{\mathsf{A}})$. Then disappearance of entanglement at finite time is the rule. Conversely, if $\gamma_+^{\mathsf{A}} = \gamma_+^{\mathsf{B}} = 0$, the two atoms decay to the ground state $|\downarrow\rangle|\downarrow\rangle$. For the initial state $|\Psi(0)\rangle = c_{\uparrow\uparrow}|\uparrow\rangle|\uparrow\rangle + c_{\downarrow\downarrow}|\downarrow\rangle|\downarrow\rangle$, entanglement disappears at finite time $t_{\mathrm{ESD}} = -\ln(1 - |c_{\downarrow\downarrow}/c_{\uparrow\uparrow}|)/\gamma_-^{\mathsf{A}}$ for inverted populations $|c_{\uparrow\uparrow}| > |c_{\downarrow\downarrow}|$, while separability is asymptotic for $|c_{\uparrow\uparrow}| \leq |c_{\downarrow\downarrow}|$ [8, 186].

Although the above steady state argument is trivial, the fact that in case (2) certain initial states exhibit sudden death of entanglement and other do not has astonished the quantum optics community. The original papers [79, 81, 268] have been followed by a huge amount of works studying entanglement for two qubits with various reservoir couplings in the Markovian and non Markovian regimes.

We now discuss what happens when there are infinitely many stationary states [234]. We consider a simple example of two qubits subject to pure phase dephasing in the interaction picture. The corresponding dynamics is described by the Lindblad equation (2.11) with $H_{\mathsf{S}} = 0$ and the two jump operators

$$M^{\mathsf{A}} = \sigma_3 \otimes 1_{\mathsf{B}} \quad , \quad M^{\mathsf{B}} = 1_{\mathsf{A}} \otimes \sigma_3 \ , \tag{2.26}$$

the associated rates being γ_{A} and γ_{B}. If the initial state $\rho(0)$ is a Bell diagonal state given by (2.17) with $\mathbf{u} = \mathbf{v} = \mathbf{0}$ and $C = \mathrm{diag}(c_1, c_2, c_3)$, then it remains of this form at all times t, with $c_1(t) = \mathrm{e}^{-2\gamma t} c_1$, $c_2(t) = \mathrm{e}^{-2\gamma t} c_2$, and $c_3(t) = c_3$ (we have set $\gamma = \gamma_{\mathsf{A}} + \gamma_{\mathsf{B}}$). In the standard basis $\{|k\rangle|l\rangle\}_{k,l=\uparrow,\downarrow}$, the two-qubit density matrix reads

$$\rho(t) = \frac{1}{4} \begin{pmatrix} 1 + c_3 & 0 & 0 & (c_1 - c_2)\mathrm{e}^{-2\gamma t} \\ 0 & 1 - c_3 & (c_1 + c_2)\mathrm{e}^{-2\gamma t} & 0 \\ 0 & (c_1 + c_2)\mathrm{e}^{-2\gamma t} & 1 - c_3 & 0 \\ (c_1 - c_2)\mathrm{e}^{-2\gamma t} & 0 & 0 & 1 + c_3 \end{pmatrix} . \tag{2.27}$$

This matrix becomes diagonal at large times $t \gg \gamma^{-1}$ (decoherence). Since c_3 is constant in time, there is no relaxation to an equilibrium and the asymptotic state ρ_∞ depends on the initial state. One has $\rho_\infty \in \partial\mathcal{S}_{\mathsf{AB}}$ if and only if $c_3 = \pm 1$. Therefore, entanglement disappears after a finite time for all initial Bell-diagonal states with $c_3 \in]-1, 1[$, whereas for $c_3 = \pm 1$ it disappears asymptotically[13].

We end this brief survey with a few words about non-interacting subsystems coupled to a common bath. Entanglement between these subsystems may increase with time due to the effective interaction mediated by the bath, and the stationary state may be entangled [44]. The same occurs in the reverse situation of two interacting subsystems coupled to independent baths [278]. By continuity of the trajectory, $\rho(t)$ is then entangled at all times larger than a given time t_0. The state $\rho(t)$ can, however, make an excursion to the separable region before that time, so that entanglement vanishes at a finite time $t_{\mathrm{ESD}} < t_0$ and then re-appears at later times $t_{\mathrm{ESR}} \leq t_0$ (entanglement sudden revivals) [92, 166, 186].

[13] By the Peres-Horodecki criterion and our definition of the boundary $\partial\mathcal{S}_{\mathsf{AB}}$ (Sec. 2.3.1), a two-qubit state ρ is in the interior of $\mathcal{S}_{\mathsf{AB}}$ if $\rho^{T_{\mathsf{B}}} > 0$. Since $\rho_\infty^{T_{\mathsf{B}}} = \rho_\infty$ is given by (2.27) with $c_1 = c_2 = 0$, this condition is satisfied when $c_3 \in]-1, 1[$. To show that $\rho_\infty \in \partial\mathcal{S}_{\mathsf{AB}}$ if $c_3 = 1$, one may consider the state ρ_ε given by choosing $c_3 = 0$, $c_1 = -c_2 = \varepsilon$, and $t = 0$ in (2.27). For $0 < \varepsilon \leq 1$, one checks that $\rho_\varepsilon \geq 0$, ρ_ε is entangled, and $\mathrm{tr}\,|\rho_\infty - \rho_\varepsilon| = \varepsilon$ (a similar argument holds for $c_3 = -1$). For $c_3 = \pm 1$ one has $c_1 = \mp c_2$, see Fig. 2.3. If $c_1 \neq 0$, the states (2.27) are entangled at all times $t \geq 0$. In particular, entanglement vanishes asymptotically if the initial state is a Bell state (2.20).

2.4 Entanglement decay when information is extracted from the baths

2.4.1 Main result

We now discuss the time-evolution of entanglement between two non-interacting subsystems A and B coupled to independent baths in the presence of monitoring by continuous measurements on the baths. For fixed realizations of the measurement outcomes, the bipartite system AB remains in a pure state $|\Psi(t)\rangle \in \mathcal{H}_{\mathsf{AB}}$, the evolution of which defines a quantum trajectory $t \mapsto |\Psi(t)\rangle$ (Sec. 2.2). In view of (2.10), the state ensemble $\{|\Psi(t)\rangle, \mathrm{d}p[\Psi]\}$ defines a pure state decomposition of the density matrix $\rho(t)$ of AB, therefore

$$\overline{E_{\mathrm{EoF}}(\Psi(t))} \;\geq\; E_{\mathrm{EoF}}(\rho(t)) \;, \tag{2.28}$$

i.e., the entanglement of formation of $\rho(t)$ is smaller or equal to the average entanglement of formation of the state ensemble selected by the measurements on the baths. The latter depends on the way information on AB is extracted from the reservoirs, that is, on the type of measurements used. As argued in Ref. [178], the quantity $\overline{E_{\mathrm{EoF}}(\Psi(t))}$, albeit it does not fulfill the quantum information theory axioms for an entanglement measure, can be promoted to the rank of relevant measure of entanglement when such information is available. This is the case for instance in Haroche's experiments.

A natural question is whether $\overline{E_{\mathrm{EoF}}(\Psi(t))}$ may vanish at finite times like $E_{\mathrm{EoF}}(\rho(t))$. Two related problems are

(a) find the measurement scheme which better protects AB against entanglement losses (in other words, find the unraveling of a given master equation leading to the maximal value of $\overline{E_{\mathrm{EoF}}(\Psi(t))}$);

(b) find the measurement scheme (if it exists) which selects the state ensemble with the minimal average entanglement of formation $\overline{E_{\mathrm{EoF}}(\Psi(t))}$, that is, such that (2.28) holds with equality.

We assume here that the measurements performed on the two baths are local. The jump operators are then of the form

$$M_m^{\mathsf{A}} \otimes 1 \quad , \quad 1 \otimes M_m^{\mathsf{B}} \;. \tag{2.29}$$

In that case, for the quantum jump dynamics, if the initial state $|\Psi(0)\rangle$ is entangled then $\overline{E_{\mathrm{EoF}}(\Psi(t))} > 0$ at all times. The reason is that the trajectory $t \mapsto |\Psi_{\mathrm{nj}}(t)\rangle$ which does not experience any jump remains entangled at all times. Actually, according to (2.3), (2.24), and (2.29), the non-unitary evolution operator $\mathrm{e}^{-\mathrm{i}tH_{\mathrm{eff}}}$ is a tensor product of two local operators acting on each subsystem. If $|\Psi_{\mathrm{nj}}(t)\rangle \propto \mathrm{e}^{-\mathrm{i}tH_{\mathrm{eff}}}|\Psi(0)\rangle$ would be separable at a given time t then, by reversing the dynamics (i.e., by applying $\mathrm{e}^{\mathrm{i}tH_{\mathrm{eff}}}$ to $|\Psi_{\mathrm{nj}}(t)\rangle$) one would deduce that $|\Psi(0)\rangle$ is separable. Hence $E_{\mathrm{EoF}}(\Psi_{\mathrm{nj}}(t)) > 0$ if $|\Psi(0)\rangle$ is entangled. But the probability to have no jump between times 0 and t, given by (2.9), is nonzero for all times t and thus $\overline{E_{\mathrm{EoF}}(\Psi(t))} > 0$. Note that this argument does not apply if the measurements on the baths are non-local.

If A and B are qubits, our main result shows that the mean entanglement of quantum trajectories can be evaluated exactly if one uses the concurrence instead of the entanglement of formation.

Theorem 2.4.1. *Let A and B be two non-interacting qubits coupled to independent baths. If the dynamics of AB is given either by the quantum jump model of Sec. 2.2.1 with local jump operators (2.29) or by a quantum state diffusion described by Eqs. (2.14) or (2.15) with the same jump operators, then the average concurrence over all quantum trajectories decays exponentially with a rate $\kappa \geq 0$ independent of the initial state $|\Psi(0)\rangle$,*

$$\overline{C(\Psi(t))} = C_0 \mathrm{e}^{-\kappa t} \;, \tag{2.30}$$

where C_0 is the initial concurrence.

The decay rate κ depends on the measurement scheme only. Thanks to the Jensen inequality and the convexity of the function h relating E_{EoF} and C (Sec. 2.3.3), the average entanglement of formation is bounded from below by

$$h(C_0 \mathrm{e}^{-\kappa t}) \;\leq\; \overline{E_{\mathrm{EoF}}(\Psi(t))} \;. \tag{2.31}$$

As a corollary, a negative answer to the problem (b) above can be given in the case where the density matrix $\rho(t) = \overline{|\Psi(t)\rangle\langle\Psi(t)|}$ exhibits a loss of entanglement at a finite time t_{ESD}: then for all *local* measurement schemes, $\overline{E_{\mathrm{EoF}}(\Psi(t))} > E_{\mathrm{EoF}}(\rho(t)) = 0$ at times $t \geq t_{\mathrm{ESD}}$. Therefore, a measurement scheme such that (2.28) holds with

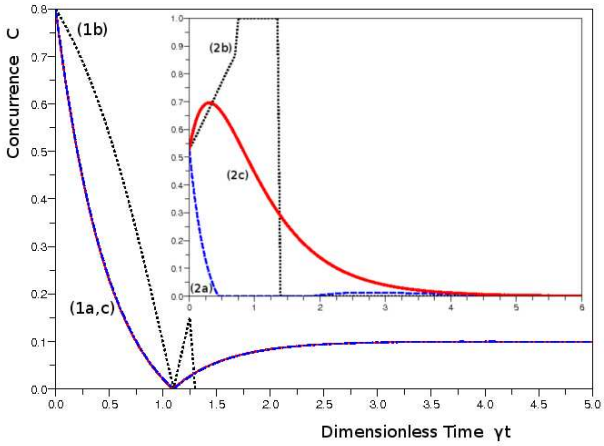

Figure 2.5: Time evolution of the concurrence of a pair of two-level atoms in resonance with the same mode of the electromagnetic field initially in the vacuum: (1a) $C(\rho(t))$ (blue dashed line); (1b) $C(\Psi(t))$ for a single trajectory (black dotted line); (1c) average $\overline{C(\Psi(t))}$ over all trajectories (red line superimposed on the blue line). The initial state is $|\Psi(0)\rangle = (2|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)/\sqrt{5}$. **Inset**: (2a,b,c) same for the initial state $|\Psi(0)\rangle = (7\mathrm{i}|\uparrow\uparrow\rangle + 2\mathrm{i}|\downarrow\downarrow\rangle)/\sqrt{53}$.

equality must necessarily involve measurements of non-local (joint) observables of the two baths. A sort of positive answer to problem (b) has been given in Refs. [52, 249], where a measurement protocol satisfying $\overline{C(\Psi(t))} = C(\rho(t))$ at all times has been found for a specific model. In agreement with the claim above, this protocol involves non-local jump operators, although the associated Lindblad equation has purely local dissipative generators[14]. In Fig. 2.5, we show the concurrences for a model involving a single non-local jump operator $M = \sigma_-^{\mathsf{A}} \otimes 1 + 1 \otimes \sigma_-^{\mathsf{B}}$ and with an Hamiltonian $H_{\mathsf{AB}} = 0$. The jump operator M corresponds to the detection of a photon emitted spontaneously by either atom A or atom B. Here the two atoms are coupled to a common bath. The concurrence $C(\rho(t))$ of the density matrix displays sudden death and revival of entanglement (Sec. 2.3.4). An explicit calculation [251] shows that for any initial state containing at most one excitation (i.e., such that $c_{\uparrow\uparrow} = 0$), the average concurrence $\overline{C(\Psi(t))}$ coincides with $C(\rho(t))$ at all times. In contrast, in the opposite case, $\overline{C(\Psi(t))}$ increases at small times whereas $C(\rho(t))$ decreases, both concurrences converging to the same asymptotic value $C_\infty = |c_{\uparrow\downarrow} - c_{\downarrow\uparrow}|^2/2$ at large times t.

2.4.2 Disentanglement rate

In this subsection we prove Theorem 2.4.1 and determine the decay rate κ. We first focus on the quantum jump dynamics. The result for state diffusion will be deduced by using the fact that the stochastic Schrödinger equations (2.14) and (2.15) arise as limits of quantum jump evolutions (Sec. 2.2.3). To simplify notation, we write $C(t)$ instead of $C(\Psi(t))$. Let us set

$$K_{\mathsf{AB}} = K_{\mathsf{A}} \otimes 1 + 1 \otimes K_{\mathsf{B}} \quad , \quad K_i = \frac{1}{2} \sum_m \gamma_m^i \, M_m^{i\dagger} M_m^i \, , \tag{2.32}$$

where γ_m^i is the rate associated to the jump operator M_m^i, with $i = \mathsf{A}, \mathsf{B}$. We will take advantage of the identities

$$\begin{aligned} &\big\langle O_{\mathsf{A}} \otimes 1 \, \sigma_2 \otimes \sigma_2 \, J \big\rangle_{\Psi(t)} = \big\langle \sigma_2 \otimes \sigma_2 \, J O_{\mathsf{A}}^\dagger \otimes 1 \big\rangle_{\Psi(t)} = \frac{c(t)}{2} \mathrm{tr}_{\mathsf{A}}(O_{\mathsf{A}}) \\ &\langle O_{\mathsf{A}}^\dagger \otimes 1 \, \sigma_2 \otimes \sigma_2 \, J O_{\mathsf{A}} \otimes 1 \rangle_{\Psi(t)} = c(t) \det(O_{\mathsf{A}}^\dagger) \, , \end{aligned} \tag{2.33}$$

where $c(t) = \langle \sigma_2 \otimes \sigma_2 \, J \rangle_{\Psi(t)}$ is given by (2.23) and O_{A} is an arbitrary local operator acting on qubit A. The same identities hold for operators $1 \otimes O_{\mathsf{B}}$ acting on qubit B.

[14] Such non-local operators may be obtained from local jump operators by the transformation (2.12), which does not change the Lindblad equation.

If no jump occurs between t and $t+\mathrm{d}t$, one gets by expanding the exponential in (2.8) and using (2.23)

$$C(t+\mathrm{d}t) = p_{\mathrm{nj}}(t,t+\mathrm{d}t)^{-1}\Big|c(t) + \mathrm{i}\mathrm{d}t\big\langle H_{\mathrm{eff}}^{\dagger}\,\sigma_2\otimes\sigma_2\,J + \sigma_2\otimes\sigma_2\,JH_{\mathrm{eff}}\big\rangle_{\Psi(t)} + \mathcal{O}(\mathrm{d}t^2)\Big|\,, \tag{2.34}$$

where $p_{\mathrm{nj}}(t,t+\mathrm{d}t)$ is the no-jump probability in $[t,t+\mathrm{d}t]$ and $H_{\mathrm{eff}} = \sum_i (H_i - \mathrm{i}K_i)$. With the help of the first identity in (2.33), one finds

$$C(t+\mathrm{d}t)\,p_{\mathrm{nj}}(t,t+\mathrm{d}t) = C(t)\Big(1 - \mathrm{tr}(K_{\mathsf{AB}})\frac{\mathrm{d}t}{2} + \mathcal{O}(\mathrm{d}t^2)\Big)\,. \tag{2.35}$$

If a jump of type (m,i) occurs in the time interval $[t,t+\mathrm{d}t]$, by virtue of (2.6), (2.7), and the second identity in (2.33), the concurrence is

$$C^{(m,i)}_{\mathrm{jump}}(t+\mathrm{d}t) = \frac{\gamma_m^i\,\mathrm{d}t}{\mathrm{d}p_m^i(t)}C(t)\big|\det(M_m^i)\big|\,. \tag{2.36}$$

Collecting (2.35) and (2.36) and using the Markov property of the jump process, one gets $\overline{C(t+\mathrm{d}t)} = \overline{C(t)}(1 - \kappa_{\mathrm{QJ}}\,\mathrm{d}t + \mathcal{O}(\mathrm{d}t^2))$ with

$$\kappa_{\mathrm{QJ}} = \frac{1}{2}\,\mathrm{tr}(K_{\mathsf{AB}}) - \sum_{m,i}\gamma_m^i|\det(M_m^i)|\,. \tag{2.37}$$

We now let $\mathrm{d}t$ go to zero to obtain $\mathrm{d}\overline{C(t)}/\mathrm{d}t = -\kappa_{\mathrm{QJ}}\,\overline{C(t)}$. The solution of this differential equation has the exponential decay claimed in the theorem. One may check that the decay rate κ_{QJ} is non-negative from the equality

$$\kappa_{\mathrm{QJ}} = \sum_{m,i}\frac{\gamma_m^i}{2}\Big(\big|\langle\uparrow|M_m^i|\uparrow\rangle - \mathrm{e}^{2\mathrm{i}\theta_m^i}\langle\downarrow|M_m^{i\dagger}|\downarrow\rangle\big|^2 + \big|\langle\uparrow|(M_m^i + \mathrm{e}^{2\mathrm{i}\theta_m^i}M_m^{i\dagger})|\downarrow\rangle\big|^2\Big)\,, \tag{2.38}$$

where $2\theta_m^i$ is the argument of $\det(M_m^i)$.

As mentioned above, the exponential decay of $\overline{C(t)}$ for quantum trajectories given by the stochastic equation (2.14) (respectively (2.15)) follows from the fact that this equation can be obtained from the quantum jump dynamics for homodyne (heterodyne) detection in the high laser intensity limit. Thus Theorem 2.4.1 is proven. Furthermore, the corresponding decay rate κ_{ho} (respectively κ_{het}) can be derived by making the substitution (2.13) in formula (2.37) and taking $\alpha_m \to \infty$. After some steps of calculation that we do not retrace here (see [251] for more detail), one gets

$$\kappa_{\mathrm{ho}} = \frac{1}{2}\,\mathrm{tr}(K_{\mathsf{AB}}) - \sum_{m,i}\gamma_m^i\Big(\mathrm{Re}\left\{\det(M_m^i)\right\} + \frac{1}{2}\big(\mathrm{Im}\left\{\mathrm{tr}_{\mathbb{C}^2}(M_m^i)\right\}\big)^2\Big) \tag{2.39}$$

$$\kappa_{\mathrm{het}} = \frac{1}{2}\,\mathrm{tr}(K_{\mathsf{AB}}) - \frac{1}{4}\sum_{m,i}\gamma_m^i\big|\mathrm{tr}_{\mathbb{C}^2}(M_m^i)\big|^2\,. \tag{2.40}$$

2.4.3 Optimal measurement schemes for entanglement preservation

We can now tackle the problem of finding the best measurements on the baths to protect entanglement (problem (a) of Sec. 2.4.1). The first results in this direction go back to a paper by Nah and Carmichael [178], who studied a two-level atom A interacting with a cavity field B, the atom and field being coupled to independent baths.

(i) *Complete preservation of entanglement.*

Let us first study the situation where perfect protection of entanglement can be achieved with quantum jumps, i.e., $\kappa_{\mathrm{QJ}} = 0$. In such a case, if the qubits are maximally entangled at time $t=0$ (that is, $C_0 = 1$) then $|\Psi(t)\rangle$ remains maximally entangled at all times $t \geq 0$ for *all* quantum trajectories (that is, $C(t) = \overline{C(t)} = 1$)[15]. This means that one may keep the qubits maximally entangled by extracting locally information from the baths. One easily shows from (2.38) that $\kappa_{\mathrm{QJ}} = 0$ if and only if $M_m^i = \lambda_m^i U_m^i$ with $\lambda_m^i > 0$ and U_m^i a unitary matrix, for any m and i. Then the mean concurrence does not decay, as it can be understood from the following argument [53]. A jump (2.7) transforms $|\Psi(t)\rangle$ into $U_m^i|\Psi(t)\rangle$ and thus does not change the amount of entanglement because U_m^i are local unitaries. The same holds for the evolution between jumps under the effective Hamiltonian $H_{\mathrm{eff}} = H_{\mathsf{AB}} - \mathrm{i}K_{\mathsf{AB}}$. For indeed, the damping part K_{AB} is a c-number and H_{AB} is a sum of

[15] We recall that the concurrence $C(t)$ belongs to $[0,1]$.

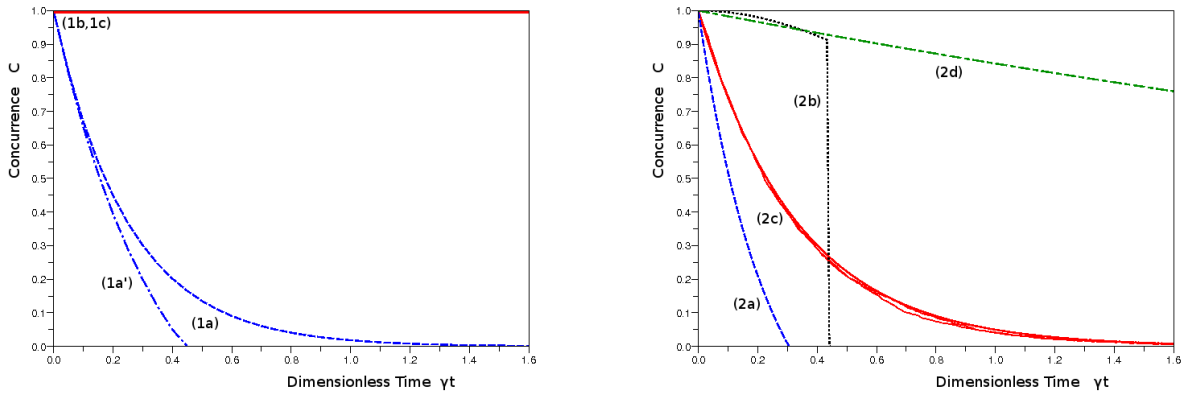

Figure 2.6: **Left panel:** concurrences of 2 qubits coupled to independent baths as a function of γt for a pure dephasing with jump operators (2.26): (1a) density matrix concurrence $C(\rho(t))$ for an initial Bell state $|\Psi(0)\rangle = |\Psi^+\rangle$ (blue dashed line); (1a') same for the initial state $|\Psi(0)\rangle = (|\Psi^+\rangle - \mathrm{i}|\Phi^+\rangle)/\sqrt{2}$ (blue dotted-dashed lines); (1b,1c) $C(\Psi(t)) = \overline{C(\Psi(t))}$ for a single trajectory (red line). **Right panel:** same at positive temperature for the jump operators (2.25) with rates $\gamma^i_+ = \gamma^i_-/2 = \gamma$ and initial state $|\Psi(0)\rangle = \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle - \mathrm{i}|\downarrow\downarrow\rangle)$: (2a) $C(\rho(t))$ (blue dashed line); (2b) $C(\Psi(t))$ for a single trajectory (black dotted line); (2c) $\overline{C(\Psi(t))}$ averaged over 1500 trajectories and from (2.30) (red lines) ; (2d) $\overline{C(\Psi(t))}$ for the best measurement scheme with jump operators (2.42).

two local Hamiltonians, see (2.24) and (2.32). A particular case for which $\kappa_{\rm QJ} = 0$ is when the jump operators $M^i_m \propto \mathbf{u}^i_m \cdot \boldsymbol{\sigma}^i$ are proportional to local spin operators. The corresponding evolution under the master equation (2.11) is a pure phase dephasing. Taking for instance the jump operators given by (2.26), the entanglement in the density matrix $\rho(t)$ disappears after a finite time for all initial entangled pure states[16] $|\Psi(0)\rangle$ such that $c_{kl} = \langle kl|\Psi(0)\rangle \neq 0$ for all $k, l = \uparrow, \downarrow$, whereas it never vanishes e.g. if $c_{\uparrow\downarrow} = c_{\downarrow\uparrow} = 0$ or if $c_{\uparrow\uparrow} = c_{\downarrow\downarrow} = 0$. This is illustrated by the decay of the concurrence $C(\rho(t))$ displayed in the left panel of Fig. 2.6. In contrast, the concurrence $C(\Psi(t))$ of all quantum trajectories do not decay.

(ii) *Two-level atoms coupled to the electromagnetic field with a Jaynes-Cummings interaction.*

Another physically relevant example is a pair of two-level atoms resonantly coupled to independent modes of the electromagnetic field. If the jumps are caused by a detection of a photon emitted by one of the atoms, which can be done in principle by surrounding each atom by photo-detectors, the jump operators are as in (2.25). Thanks to (2.30) and (2.37), the average concurrence is[17] $\overline{C(t)} = C_0 e^{-(\gamma^{\mathsf{A}}_+ + \gamma^{\mathsf{B}}_+ + \gamma^{\mathsf{A}}_- + \gamma^{\mathsf{B}}_-)t/2}$. It is compared in the right panel of Fig. 2.6 with the concurrence of the density matrix, obtained by solving the master equation (2.11). At positive temperatures, the latter concurrence vanishes after a finite time for all initial states, as a consequence of the entanglement sudden death (Sec. 2.3.4). We would like to determine the optimal measurement scheme to protect the entanglement between the two atoms. With this goal, we replace the photon-counting jump operators $M^i_\pm = \sigma^i_\pm$ by M^i_μ using the transformation (2.12) with a $N \times 2$ unitary matrix $(u^i_{\mu m})^{m=\pm}_{\mu=1,\ldots,N}$ for each $i = \mathsf{A}, \mathsf{B}$. Note that we do not mix jump operators for different i's, as this would lead to non-local measurements. From (2.37), the new entanglement decay rate is found to be

$$\kappa_{\rm QJ} = \frac{1}{2} \sum_{\mu,i} \left(\sqrt{\gamma^i_-} |u^i_{\mu -}| - \sqrt{\gamma^i_+} |u^i_{\mu +}| \right)^2 . \tag{2.41}$$

By optimizing over all unitaries and making use of the Cauchy-Schwarz inequality and $\sum_\mu |u^i_{\mu\,\pm}|^2 = 1$, the smallest disentanglement rate is reached e.g. for $N = 2$ and $u^i_{1\,\pm} = \pm u^i_{2\,\pm} = 1/\sqrt{2}$, i.e.,

$$\sqrt{\gamma^i_1} M^i_1 = 2^{-\frac{1}{2}} \left(\sqrt{\gamma^i_+} \sigma^i_+ + \sqrt{\gamma^i_-} \sigma^i_- \right) \quad , \quad \sqrt{\gamma^i_2} M^i_2 = 2^{-\frac{1}{2}} \left(\sqrt{\gamma^i_+} \sigma^i_+ - \sqrt{\gamma^i_-} \sigma^i_- \right) \tag{2.42}$$

[16] This follows from similar arguments as in Sec. 2.3.4. In fact, if $c_{kl} \neq 0$ then the asymptotic state $\rho_\infty = \sum_{k,l} |c_{kl}|^2 |kl\rangle\langle kl|$ is in the interior of $\mathcal{S}_{\mathsf{AB}}$.

[17] It is instructive to derive this result directly, without relying on Theorem 2.4.1. One notes that a jump with jump operator σ^{A}_- or σ^{B}_-, which corresponds to the detection of a photon emitted by atom A or B, disentangles the atoms completely. Actually, if e.g. a detector surrounding A makes a click, then one is certain that A is in its ground state $|\downarrow\rangle$, so that AB is in a product state (see [251]).

with $i = \mathsf{A}, \mathsf{B}$. It is given by

$$\kappa_{\rm QJ}^{\rm opt} = \frac{1}{2} \sum_{i=\mathsf{A},\mathsf{B}} \left(\sqrt{\gamma_-^i} - \sqrt{\gamma_+^i} \right)^2 . \tag{2.43}$$

Note that $\kappa_{\rm QJ}^{\rm opt} = \kappa_{\rm QJ}$ at zero temperature ($\gamma_+^i = 0$) and $\kappa_{\rm QJ}^{\rm opt} = 0$ (perfect protection) at infinite temperature ($\gamma_+^i = \gamma_-^i$). The decay of $\overline{C(\Psi(t))}$ for the optimal quantum jump dynamics is shown in the right panel of Fig. 2.6. Based on our result, the authors of Ref. [53] have proposed to engineer an artificial infinite temperature reservoir for a two-level atom via an adiabatic elimination of a third level populated with a laser pump, and to use quantum jumps or state diffusion to preserve entanglement.

(iii) *Optimal homodyne detection*

We now turn to homodyne detection. The corresponding disentanglement rate $\kappa_{\rm ho}$ is given by (2.39). Unlike $\kappa_{\rm QJ}$, this rate changes when the operators M_m^i acquire a phase factor, $M_m^i \to {\rm e}^{-{\rm i}\theta_m^i} M_m^i$. This arises for homodyne detection with complex laser amplitudes $\alpha_m^i = |\alpha_m^i| {\rm e}^{{\rm i}\theta_m^i}$, $|\alpha_m^i| \gg 1$. Minimizing over the laser phases θ_m^i yields

$$\kappa_{\rm ho}^{\rm opt} = \frac{1}{2} {\rm tr}(K_{\mathsf{AB}}) - \sum_{m,i} \gamma_m^i \Big(\Big| \det(M_m^i) - \frac{1}{4} \big({\rm tr}_{\mathbb{C}^2}(M_m^i) \big)^2 \Big| + \frac{1}{4} \Big| {\rm tr}_{\mathbb{C}^2}(M_m^i) \Big|^2 \Big) . \tag{2.44}$$

It is easy to convince oneself that

$$\kappa_{\rm ho}^{\rm opt} \leq \kappa_{\rm QJ} \quad , \quad \kappa_{\rm ho}^{\rm opt} \leq \kappa_{\rm het} . \tag{2.45}$$

The first (respectively second) inequality is strict excepted if the two eigenvalues of M_m^i have the same modulus (are equal) for all m and i. Thus optimal homodyne detection protects entanglement better than - or, if the aforementioned conditions are fulfilled, as well as - photon counting and heterodyne detection. For the jump operators given by (2.42), homodyne detection does not do better than quantum jumps, $\kappa_{\rm QJ}^{\rm opt} = \kappa_{\rm ho}^{\rm opt}$. Let us stress that the optimal measurements (in particular, the laser phases θ_m^i minimizing the rate $\kappa_{\rm ho}$) only depend on the Lindblad operators M_m^i in the master equation and are the same for all initial states of the qubits.

2.5 Conclusions

We have found explicit formulas for the mean concurrence of quantum trajectories and have shown that the monitoring of the two qubits obtained thanks to continuous measurements on the baths may be used to protect entanglement. These results shed new light on the phenomenon of entanglement sudden death. If the measurements are performed locally on the two independent baths, the mean concurrence $\overline{C(\Psi(t))}$ is either constant in time or vanishes exponentially with a rate depending on the measurement scheme only. A constant value implies a perfect protection of maximally entangled states for all trajectories. In the case of pure dephasing and for Jaynes-Cumming couplings at infinite temperature, we have found measurement schemes independent of the initial state leading to such a perfect entanglement protection. Despite obvious analogies, this way to keep entanglement differs from the strategies based on the quantum Zeno effect[18]. A corollary of our result is that, if the density matrix $\rho(t)$ suffers from a disappearance of entanglement at finite times, a measurement scheme which could prepare the two qubits at all times in a state ensemble with a minimal mean entanglement of formation $\overline{E_{\rm EoF}(\Psi(t))} = E_{\rm EoF}(\rho(t))$ must necessarily involve measurements of non-local observables of the two baths. In contrast, we have found that if the two qubits are coupled to a common bath, $\overline{C(\Psi(t))}$ may coincide with the concurrence $C(\rho(t))$ of the density matrix for certain initial states.

Let us point out that generalizing our result to higher-dimensional systems is not straightforward. One could use the generalized concurrence defined by Eq.(10.4) in chapter 10, but unlike what happens for two qubits, $C(t + {\rm d}t)$ does not appear to be a function of $C(t)$ and the behavior of $\overline{C(t)}$ is more complicated.

[18] Such a strategy has been proposed in Ref. [166]. In the quantum jump and quantum state diffusion models considered here, the time interval between consecutive measurements is small but not exactly zero, whereas in [166] a perfect entanglement protection is reached in the idealized limit $\gamma\, {\rm d}t \to 0$, i.e., when the measurements completely prevent the decay of the superradiant state.

Chapter 3

Decoherence in Bose-Josephson junctions

Nous irons droit à l'essentiel, mais pour comprendre l'essentiel il est nécessaire de parler des détails
(journaliste de France Inter).

3.1 Motivations

In this chapter, we investigate a physical model which is more directly related to experiments than in the other chapters. We are interested by ultracold gases of interacting bosonic atoms in two modes, realized by coupling two trapped Bose-Einstein condensates (BECs) which may exchange particles. Since the first experimental observations in 1995 of Bose-Einstein condensation with Rubidium atoms in the group of E. Cornell and C. Wiemann and with Sodium atoms in the group of W. Ketterle, many laboratories have produced BECs from metastable vapors of various Alkali atoms at very low temperature (of the order of 10-100 nK) and low densities (of the order of 10^{13}-$10^{14}\,\mathrm{cm}^{-3}$). The trapping of the atoms is realized either by applying an inhomogeneous magnetic field or by means of optical potentials created by counter-propagating off-resonant laser fields forming standing waves. A so-called *external Bose-Josephson junction* is obtained when the condensed atoms are trapped in a double-well potential. The two spatial wave functions localized within each wells constitute the two modes of the junction. Tunneling between the wells leads to an inter-mode coupling. In *internal Bose-Josephson junctions*, the atoms are on the contrary trapped in a single well. The two modes correspond to two distinct hyperfine internal states, which are coupled by a resonant microwave or radio-frequency field. The two systems have the same dynamics. They have been realized experimentally with optically trapped ^{87}Rb atoms [87, 106] in the group of M. Oberthaler in Heidelberg and with magnetically trapped ^{87}Rb atoms on a chip [39, 202] in the group of P. Treutlein in Basel.

An interesting feature of ultracold trapped atomic gases is the experimental tunability of the physical parameters governing their dynamics. The shape of the optical potential is controlled by the laser fields creating it, enabling in particular to design quasi one-dimensional or two-dimensional confinements (anisotropic traps) and disordered potentials (speckle). Artificial gauge fields can be created by applying additional laser fields [70]. The strength and even the sign of the inter-atomic interactions can be controlled by using Feshbach resonances in the presence of a uniform magnetic field [89, 91]. For these reasons, BECs have attracted a lot of interest in recent years, in view of their applications to quantum information technologies and their ability to simulate the ground state properties and dynamics of many-body solid state systems. In particular, the superfluid Mott-insulator phase transition has been observed experimentally in a BEC trapped in an optical lattice potential [104].

One of our motivations to study Bose-Josephson junctions (BJJs) is that they generate dynamically macroscopic superpositions of coherent states. The first proposal to generate such superpositions with light consisted of sending photons through a medium presenting a strong Kerr non-linearity [272]. In such media, the dynamical phases of Fock states are nonlinear in the photon number, thus the phase of an initial coherent state is split. An alternative way is to use a cavity field coupled to atoms crossing the cavity [113]. In the experiments of the Haroche group, superpositions of coherent states of the cavity field have been produced, and their progressive transformation into statistical mixtures as time evolves has been observed [76] (see also the circuit QED experiments in Yale [250]). Superpositions of coherent states for the motional degree of freedom of a single Be^+ trapped ion have been also produced in the group of D. Wineland [172]. In gases of ultracold bosonic atoms, interactions between atoms lead to non-linearities similar to the Kerr non-linearity. Starting from a spin

coherent state (that is, a product state in which all atoms are in the same superposition of the two modes), the unitary dynamics builds up superpositions of spin coherent states after a sudden quench to zero of the inter-mode coupling. The main question is to know whether such superpositions of coherent states can be formed in a BJJ even in the presence of experimental noises and interaction of the trapped atoms with their environment. The impact of decoherence on the superpositions has been investigated in Ref. [131] by considering the coupling of the atoms with the electromagnetic vacuum. However, these decoherence effects seem to be negligible in the Heidelberg and Basel experiments. In what follows, we study the impact on the superpositions of (i) a phase noise produced by magnetic fluctuations in internal BJJs and by relative fluctuations of the bottoms of the two wells in external BJJs (ii) atom losses in the BEC provoked by recombination and collision processes. These two effects are presumably the main sources of decoherence in the aforementioned experiments. Their impact on spin squeezed states, which are produced by the unitary dynamics of the BJJ at earlier times than the macroscopic superpositions, has been analyzed in detail in Refs. [218, 270, 271]. Unlike macroscopic superpositions, spin squeezed states have been already observed experimentally [87, 106, 202].

Multipartite-entangled states such as squeezed states and macroscopic superpositions of coherent states are particularly interesting for high-precision interferometry. Indeed, they can be used to estimate phase shifts with a resolution below the shot noise limit, that is, better than what can be done with independent atoms. The highest possible phase sensitivity is smaller than the shot noise value $(\Delta\phi)_{\rm SN} \sim 1/\sqrt{N}$ by a factor $1/\sqrt{N}$, N being the number of atoms. It is achieved when the input state of the interferometer is a macroscopic superposition of two coherent states with opposite phases on the Bloch sphere [197]. In order to quantify the amount of quantum correlations useful for interferometry, we study the time evolution of the quantum Fisher information. This quantity is related to the Cramér-Rao lower bound on the precision with which an unknown parameter - the phase shift in interferometry - can be determined via measurements on the output state and statistical estimators. Important applications of interferometry with entangled states is the enhancement of the precision of atomic clocks and magnetic field sensors.

The chapter is organized as follows. In Sec. 3.2 we review the dissipation-free dynamics of BJJs after a sudden quench to zero of the inter-mode coupling. In Sec. 3.3 we study decoherence and phase diffusion due to a phase noise and to atom losses. The time evolution of the quantum Fisher information in the presence of atom losses is discussed in Sec. 3.4. Sec. 3.5 addresses the problem of the impact of a single atom loss event on a macroscopic superposition, leading to qualitative explanations of the results of Sec. 3.4. The last section presents some conclusive remarks.

3.2 Unitary dynamics of Bose-Josephson junctions

3.2.1 Bose-Hubbard Hamiltonian

We will describe the condensed atoms in external or internal BJJs by the two-mode Bose-Hubbard Hamiltonian [168]

$$H_0 = \sum_{i=1,2} \left(E_i n_i + \frac{U_i}{2} n_i(n_i-1) \right) + U_{12} n_1 n_2 + K(a_1^\dagger a_2 + a_2^\dagger a_1) \,, \tag{3.1}$$

where a_i, $a_i^\dagger$, and $n_i = a_i^\dagger a_i$ are the bosonic annihilation, creation, and number operators in mode $i = 1, 2$.

For an *external BJJ*, the Hamiltonian (3.1) is obtained formally by replacing the field operator $\widehat{\psi}(x)$ in the second-quantized many-body Hamiltonian of the N atoms by $\psi_1(x)a_1 + \psi_2(x)a_2$, where $\psi_i(x)$ is the (real) wave function of an atom localized in the ith well in the lowest energy level $E_i = \int {\rm d}x (\nabla\psi_i)^2(x)/(2m_{\rm at}) + \int {\rm d}x\, V_i(x)\psi_i(x)^2$ (here $V_i(x)$ is the single-well potential and $m_{\rm at}$ the atomic mass). The contact interactions between the atoms lead to the quadratic terms in n_i, with interaction energies $U_i = (4\pi\ell_i/m_{\rm at}) \int {\rm d}x\, \psi_i(x)^4$, with ℓ_i the scattering length in well i. Hereafter, we assume repulsive interactions, i.e., $\ell_i > 0$. Quadratic terms involving overlap integrals of $\psi_1(x)$ and $\psi_2(x)$ are neglected, so that in particular $U_{12} = 0$ (no interactions between atoms localized in distinct wells). The last contribution in (3.1) originates from the overlap integrals in the kinetic and external potential parts of the many-body Hamiltonian. It describes the tunneling between the two wells, with tunnel energy K. The rigorous derivation of the Bose-Hubbard Hamiltonian (3.1) from the many-body problem is an interesting open issue (see Part III below). In particular, it would be important to identify precisely the weak interaction and large inter-well distance limits in which the two-mode approximation is justified (see [168] for a physical discussion in this respect).

For an *internal BJJ*, the energy E_i in (3.1) is the energy of the hyperfine level i. The interaction energy U_i for atoms in the same mode i is given by the same formula as above albeit with $\psi_i(x)$ the solution of the Gross-Pitaevskii equation in the single-well potential. The cross-interaction energy U_{12} cannot anymore

be neglected. The tunneling term originates from a resonant microwave or radio-frequency field coupling the two hyperfine levels, so that K is an experimentally controllable parameter. Even though the validity of the two-mode approximation is questionable in the experiments of Refs. [106, 202], we will assume here that the Bose-Hubbard Hamiltonian (3.1) captures qualitatively the important physics.

It is convenient to rewrite this Hamiltonian in terms of the angular momentum operators[1]

$$J_x = \frac{1}{2}\big(a_1^\dagger a_2 + a_2^\dagger a_1\big) \quad , \quad J_y = -\frac{\mathrm{i}}{2}\big(a_1^\dagger a_2 - a_2^\dagger a_1\big) \quad , \quad J_z = \frac{1}{2}\big(a_1^\dagger a_1 - a_2^\dagger a_2\big) \tag{3.2}$$

(Schwinger representation). If the BJJ has a well-defined fixed total atom number $N = n_1 + n_2$, the Bose-Hubbard Hamiltonian takes the form

$$H_0 = \chi J_z^2 - \lambda J_z + 2K J_x \ , \tag{3.3}$$

where terms depending only on N have been dropped out and

$$\chi = \frac{U_1 + U_2 - 2U_{12}}{2} \quad , \quad \lambda = E_2 - E_1 + (N-1)\frac{U_2 - U_1}{2} \ . \tag{3.4}$$

It is worth noting that the Hamiltonian (3.3) also appears in the physics of superconductor-insulator-superconductor tunnel junctions (Josephson junctions).

3.2.2 Spin coherent states

When tunneling dominates inter-atomic interactions, the atoms in the BJJ are independent and delocalized in the two modes, more precisely, they are in a Hartree state with single atom wave function $(\psi_1(x) + \psi_2(x))/\sqrt{2}$. This state is an example of $SU(2)$-coherent state. Let $|n_1, n_2\rangle$ be the Fock states, that is, joint eigenstates of the number operators in modes 1 and 2 with eigenvalues n_1 and n_2. The $SU(2)$-coherent states are defined as [274]

$$|N;\theta,\phi\rangle = \sum_{n_1=0}^{N} \left(\begin{array}{c} N \\ n_1 \end{array} \right)^{\frac{1}{2}} \frac{(\tan(\theta/2))^{n_1}}{[1+\tan^2(\theta/2)]^{\frac{N}{2}}}\, \mathrm{e}^{-\mathrm{i}n_1\phi} |n_1, n_2 = N - n_1\rangle \ . \tag{3.5}$$

It is easy to show that $|N;\theta,\phi\rangle \propto (\mathrm{e}^{-i\phi}\sin(\theta/2)a_1^\dagger + \cos(\theta/2)a_2^\dagger)^N|0\rangle$ (here, $|0\rangle$ is the vacuum state). Hence $|N;\theta,\phi\rangle$ is a product state in which all atoms are in the same superposition of the two modes.

An arbitrary (pure or mixed) state ρ can be represented by its Husimi distribution on the Bloch sphere of radius $N/2$ (the classical phase space in our problem),

$$Q_N(\theta,\phi) = \frac{1}{\pi}\langle N;\theta,\phi|\rho|N;\theta,\phi\rangle \ . \tag{3.6}$$

This distribution provides a useful information on the phase content of ρ. The coherent state (3.5) has a Husimi distribution with a single peak at (θ,ϕ) of width $\approx 1/\sqrt{N}$, as illustrated in the panel (a) of Fig. 3.1. In analogy with quantum optics, we represent this state on the Bloch sphere as a disc of diameter $\sqrt{N}/2$ centered at $N(\sin\theta\cos\phi, \sin\theta\sin\phi, -\cos\theta)/2$ (see Fig. 3.2). The coordinates of the center of the disc are the expectation values of the angular momentum operators J_x, J_y, and J_z, whereas its diameter gives the quantum fluctuations of $J_{\mathbf{n}} = \mathbf{J}\cdot\mathbf{n}$ in directions $\mathbf{n}$ tangential to the sphere. The Fock state $|n_1 = N, n_2 = 0\rangle$ is a coherent state with $\theta = \pi$, located at the north pole of the Bloch sphere. The unitary operator $\mathrm{e}^{-\mathrm{i}\phi J_{\mathbf{n}}}$ performs a rotation by an angle ϕ around the axis specified by the unit vector $\mathbf{n}$.

3.2.3 Dynamics after a sudden quench to zero of the tunnel amplitude

Let us consider a BJJ with N_0 atoms initially in the ground state of the Hamiltonian (3.3) for $E_1 = E_2$ and in the limit of small interactions with respect to tunneling, i.e., $H_0 = 2KJ_x$ This state is the coherent state

$$|\psi(0)\rangle = |N_0;\phi=0\rangle \equiv |N_0;\theta=\frac{\pi}{2},\phi=0\rangle \ . \tag{3.7}$$

We are interested in the dynamics of the BJJ driven by interactions after a sudden quench to zero of the inter-mode coupling K. Going to the rotating frame, we may suppose that $\lambda = 0$. It is immediate from (3.5) that the atomic state $|\psi^{(0)}(t)\rangle = \mathrm{e}^{-\mathrm{i}t\chi J_z^2}|\psi(0)\rangle$ displays a periodic evolution with period $T = 2\pi/\chi$ if N_0 is even and $T/2$ if N_0 is odd. It can be shown that the evolution between $t = 0$ and $t = T$ first builds up spin

[1] It is easy to check that J_x, J_y, and J_z satisfy the commutation relations of angular momenta.

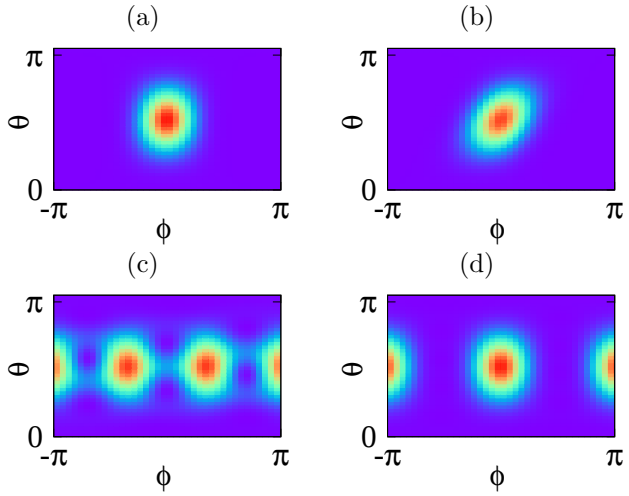

Figure 3.1: Husimi distributions in a dissipation-free BJJ at some specific times: (a) $t = 0$ (coherent state); (b) $t = T/40$ (spin squeezed state) , (c) $t = T/6$ (3-component superposition of coherent states), (d) $t = T/4$ (2-component superposition). Here $U_1 = U_2 = \frac{2\pi}{T}$, $U_{12} = E_1 = E_2 = 0$, and $N_0 = 10$.

squeezed states (states with a reduced fluctuation of $J_{\mathbf{n}}$ in one direction $\mathbf{n}$ and an enhanced fluctuation in the perpendicular direction)[2], the highest squeezing being produced at time $t \approx TN_0^{-2/3}$ [146]. Later on, when the Husimi distribution covers the whole equator of the Bloch sphere, interferences and quantum superpositions should be expected to come into play. Indeed, at the times

$$t_q = \frac{\pi}{\chi q} = \frac{T}{2q} \quad , \quad q = 2, 3, \ldots , \tag{3.8}$$

one finds that the atoms are in superpositions of coherent states[3] [230, 272]

$$|\psi^{(0)}(t_q)\rangle = \sum_{k=0}^{q-1} c_{k,q} \big| N_0; \phi_{k,q} \big\rangle \tag{3.9}$$

with coefficients $c_{k,q}$ of equal moduli $q^{-1/2}$ and phases $\theta = \pi/2$ and $\phi_{k,q} = \phi_{0,q} + 2\pi k/q$. In particular, the BJJ is at time $t = t_2$ in the superposition $(|N_0; \phi_{0,2}\rangle - |N_0; \phi_{0,2} + \pi\rangle)/\sqrt{2}$ of two coherent states located on the equator of the Bloch sphere at diametrically opposite points. Panels (c) and (d) of Fig. 3.1 show the Husimi distributions of the states (3.9) for $q = 2$ and $q = 3$.

As stressed above, the observation of Schrödinger cat states such as (3.9) with large atom numbers N_0 is an exciting challenge. Due to the tunability of the interactions and trapping potential and since the sources of decoherence in BECs are relatively well understood, cold atoms and more specifically BJJs seem to be good candidates for observing such states[4]. Note that spin squeezing has been already obtained experimentally in BJJs [87].

3.2.4 Structure of the atomic state in the Fock basis

The density matrix

$$\rho^{(0)}(t) = |\psi^{(0)}(t)\rangle\langle\psi^{(0)}(t)| = \mathrm{e}^{-\mathrm{i}t\chi J_z^2} |\psi(0)\rangle\langle\psi(0)| \mathrm{e}^{\mathrm{i}t\chi J_z^2} \tag{3.10}$$

[2]This can be understood intuitively as follows. The initial state is represented on the Bloch sphere by a disc centered at $(N/2, 0, 0)$ with radius $\sqrt{N}/4$ (see above). One may think of the unitary $\mathrm{e}^{-\mathrm{i}t\chi J_z^2}$ as a rotation around the z-axis with an angle $t\chi J_z$. As the upper (lower) half of the disc rotates in the anticlockwise (clockwise) direction, an elongated ellipse is formed, as illustrated in panel (b) of Fig. 3.1. This ellipse corresponds to a spin squeezed state.

[3]Equation (3.9) can be derived by using (3.5) and the Fourier expansions

$$\mathrm{e}^{-\mathrm{i}\pi n_1^2/q} = \begin{cases} u \sum_{k=0}^{q-1} \mathrm{e}^{\mathrm{i}\pi k^2/q} \mathrm{e}^{-2\mathrm{i}\pi n_1 k/q} & \text{if } q \text{ is even} \\ u' \sum_{k=0}^{q-1} \mathrm{e}^{\mathrm{i}\pi k(k+1)/q} \mathrm{e}^{-\mathrm{i}\pi n_1 (2k+1)/q} & \text{if } q \text{ is odd} \end{cases}$$

with $|u|^2 = |u'|^2 = 1/q$ by the Parseval identity. The phase $\phi_{0,q}$ is equal to $-N_0\pi/q$ if q is even and $(-N_0 + 1)\pi/q$ if q is odd.

[4]Due to unavoidable atom losses, N_0 should, however, be smaller than the hundreds of atoms used so far in the Heidelberg and Basel experiments [202, 106].

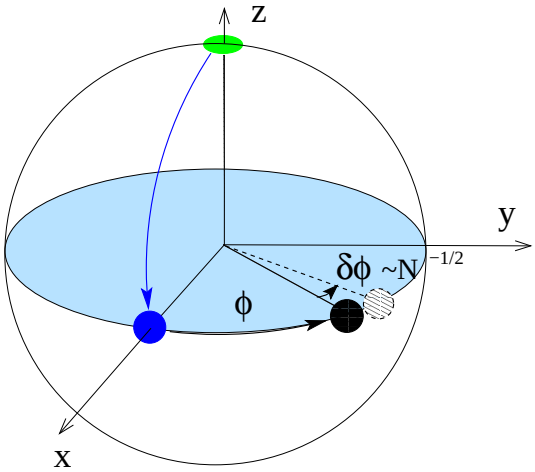

Figure 3.2: Rotations on the Bloch sphere in the interferometric scheme (see Sec. 3.4.1): the input coherent state at the north pole (green disk) is rotated around the y-axis by an angle $\pi/2$ (blue disk) and afterward around the z-axis by the unkown phase ϕ (black disk). The precision $\Delta\phi$ on the estimation of ϕ is larger than the size $\sqrt{N}/2$ of the disk, representing the angular momentum fluctuations, divided by the radius $N/2$ of the sphere. The last rotation around the y-axis is not represented.

has a simple form in the Fock basis. In all what follows, we set $n_2 = N_0 - n_1$ and $n_2' = N_0 - n_1'$. The moduli of the matrix elements of $\rho^{(0)}(t)$ are time-independent and behave in the limit $N_0 \gg 1$ like

$$\big|\langle n_1, n_2|\rho^{(0)}(t)|n_1', n_2'\rangle\big| = \frac{1}{2^{N_0}} \begin{pmatrix} N_0 \\ n_1 \end{pmatrix}^{\frac{1}{2}} \begin{pmatrix} N_0 \\ n_1' \end{pmatrix}^{\frac{1}{2}} \sim \sqrt{\frac{2}{\pi N_0}} \exp\Big\{-\frac{(n_1 - N_0/2)^2 + \big(n_1' - N_0/2\big)^2}{N_0}\Big\} . \quad (3.11)$$

At the time t_q of formation of the superposition (3.9), it is convenient to decompose the density matrix as a sum of a "diagonal part" $[\rho^{(0)}(t_q)]_{\rm d}$, corresponding to the statistical mixture of the coherent states, and an "off-diagonal part" $[\rho^{(0)}(t_q)]_{\rm od}$ describing the coherences between these states, i.e.,

$$[\rho^{(0)}(t_q)]_{\rm d} = \sum_{k=0}^{q-1} \rho^{(0)}_{kk,q} \quad , \quad [\rho^{(0)}(t_q)]_{\rm od} = \sum_{k\neq k'}^{q-1} \rho^{(0)}_{kk',q} \quad , \quad \rho^{(0)}_{kk',q} = c_{k,q} c^*_{k',q} |N_0; \phi_{k,q}\rangle\langle N_0; \phi_{k',q}| \, . \quad (3.12)$$

These diagonal and off-diagonal parts exhibit remarkable structures in the Fock basis, which enable to read them easily from the total density matrix[5]:

$$\begin{aligned} \langle n_1, n_2|[\rho^{(0)}(t_q)]_{\rm d}|n_1', n_2'\rangle &= 0 \quad \text{if } n_1' \neq n_1 \text{ modulo } q \\ \langle n_1, n_2|[\rho^{(0)}(t_q)]_{\rm od}|n_1', n_2'\rangle &= 0 \quad \text{if } n_1' = n_1 \text{ modulo } q. \end{aligned} \quad (3.13)$$

The off-diagonal part does almost not contribute to the Husimi distribution. The Husimi plots in panels (c) and (d) of Fig. 3.1 thus essentially show the diagonal parts only. On the other hand, the quantum correlations are contained in the off-diagonal part.

3.3 Dissipative dynamics of Bose-Josephson junctions

In this section, we consider the same evolution with initial state (3.7) under the Hamiltonian (3.3) with $K = \lambda = 0$ as in the preceding section, but we add noise or atom losses on top of the unitary dynamics. We study the effects of decoherence and phase relaxation on the atomic state, in particular the superpositions at time t_q.

[5] The first line follows from (3.5) and the identity $\sum_{k=0}^{q-1} \mathrm{e}^{2\mathrm{i}k(n'-n)\pi/q} = q$ if $n = n'$ modulo q and 0 otherwise. The second line is obtained by using $[\rho^{(0)}(t_q)]_{\rm od} = \mathrm{e}^{-\mathrm{i}\pi J_z^2/q}|N_0; \phi = 0\rangle\langle N_0; \phi = 0|\mathrm{e}^{\mathrm{i}\pi J_z^2/q} - [\rho^{(0)}(t_q)]_{\rm d}$.

3.3.1 Decoherence and phase relaxation due to a phase noise

Random time fluctuations of the energies E_i.

In the external BJJ of Ref. [87], random changes of the shape of the two lowest potential wells are caused by fluctuations in the direction of the laser beam producing a sinusoidal optical potential on top of the harmonic trap (the other wells are not occupied by atoms at the temperature of the experiment). This induces time fluctuations of the lowest energies E_1 and E_2 of these two wells. Similarly, in the internal BJJs of Refs. [106, 202], the presence of a randomly fluctuating inhomogeneous magnetic field leads to fluctuations of the hyperfine energy levels E_i. We account for this noise by considering the Hamiltonian

$$H(t) = \chi {J_z}^2 - \lambda(t) J_z \ , \tag{3.14}$$

where $\lambda(t)$ is a classical stochastic process. We assume here a stationary process with a correlation function

$$k(t-t') = \overline{\lambda(t)\lambda(t')} - \overline{\lambda(0)}^2 \tag{3.15}$$

independent of the number of atoms N_0 (this is justified if one can neglect the fluctuations of the interaction energies U_i, see (3.4)). The overline in (3.15) refers to the average over the noise realizations. We set $\overline{\lambda} = \overline{\lambda(0)} = \overline{\lambda(t)}$. Since $[H(t), J_z] = 0$ at all times, J_z is conserved as in the noiseless case. As a consequence, the model can be solved exactly, without having to rely on a Markov approximation. For a given realization $\lambda(t)$, the Schrödinger-evolved state is obtained from the state $|\psi^{(0)}(t)\rangle$ in the absence of noise through a rotation around the z-axis by a random angle $\phi(t) = -\int_0^t \mathrm{d}\tau \lambda(\tau)$, i.e., $|\psi(t)\rangle = \mathrm{e}^{-\mathrm{i}\phi(t)J_z}|\psi^{(0)}(t)\rangle$. We denote by $f(\phi,t) = \overline{\delta(\phi(t)-\phi)}$ the distribution of the phase $\phi(t)$. Performing the average over $\lambda(t)$ of the pure state $|\psi(t)\rangle$ yields the density matrix[6] $\rho(t) = \overline{|\psi(t)\rangle\langle\psi(t)|}$.

We obtain

$$\rho(t) = \int_{-\infty}^{\infty} \mathrm{d}\phi \, f(\phi,t) \, \mathrm{e}^{-\mathrm{i}\phi J_z} \rho^{(0)}(t) \mathrm{e}^{\mathrm{i}\phi J_z} \ , \tag{3.16}$$

where $\rho^{(0)}(t)$ is the lossless density matrix (3.10). By projecting (3.16) over the Fock basis we get

$$\langle n_1, n_2|\rho(t)|n_1', n_2'\rangle = \widehat{f}(n_1'-n_1,t)\langle n_1, n_2|\rho^{(0)}(t)|n_1', n_2'\rangle \ , \tag{3.17}$$

where $\widehat{f}(m,t) = \overline{\mathrm{e}^{\mathrm{i}m\phi(t)}}$ is the Fourier transform of $f(\phi,t)$ (characteristic function) and as before $n_2 = N_0 - n_1$, $n_2' = N_0 - n_1'$.

To be specific, let us consider a Gaussian noise. Then $\widehat{f}(m,t) = \mathrm{e}^{-a^2(t)m^2/2}\mathrm{e}^{-\mathrm{i}\overline{\lambda}tm}$ with a N-independent variance $a^2(t)$ given by

$$a^2(t) = \int_0^t \mathrm{d}\tau \int_0^t \mathrm{d}\tau' k(\tau-\tau') = \begin{cases} k(0)\, t^2 & \text{if } t \le t_c \ \text{(small time regime)} \\ 2t \displaystyle\int_0^\infty \mathrm{d}\tau \, k(\tau) & \text{if } t \gg \tau_c \ \text{(Markov regime).} \end{cases} \tag{3.18}$$

The noise correlation times τ_c and t_c are defined in analogy with τ_R and t_R in Sec. 1.3.5. We thus find

$$\langle n_1, n_2|\rho(t)|n_1', n_2'\rangle = \exp\Big\{-\frac{1}{2}a^2(t)(n_1-n_1')^2\Big\}\mathrm{e}^{\mathrm{i}\overline{\lambda}t(n_1-n_1')}\langle n_1, n_2|\rho^{(0)}(t)|n_1', n_2'\rangle \ . \tag{3.19}$$

The effect of the noise is to suppress the off-diagonal elements of $\rho(t)$ in the Fock basis (decoherence). At long times $t \gg (\int_0^\infty \mathrm{d}\tau k(\tau))^{-1}$, $\rho(t)$ converges to a statistical mixture of Fock states with the same probabilities as for the initial state (see (3.11)),

$$\rho(\infty) = \sum_{n_1=0}^{N_0} \frac{1}{2^{N_0}} \begin{pmatrix} N_0 \\ n_1 \end{pmatrix} |n_1, N_0-n_1\rangle\langle n_1, N_0-n_1| = \int_0^{2\pi} \frac{\mathrm{d}\phi}{2\pi} |N_0;\phi\rangle\langle N_0;\phi| \ . \tag{3.20}$$

The last equality means that the phase ϕ is uniformly distributed on $[0, 2\pi]$ at large times t or large noise intensities $\int_0^\infty \mathrm{d}\tau k(\tau)$. The spreading of the phase along the equator (phase relaxation) can be seen on the Husimi distributions shown in Fig. 3.3. In the absence of noise, the distribution is peaked at $\phi = 0$ and π, which correspond to the two coherent states of the superposition. The peaks are smeared for $a_2 \simeq 1$ and the distribution reaches a flat profile for $a_2 \gg 1$. Equation (3.19) predicts a decay of the visibility $\nu(t) = 2\,\mathrm{tr}(J_x\rho(t))/N_0$ like

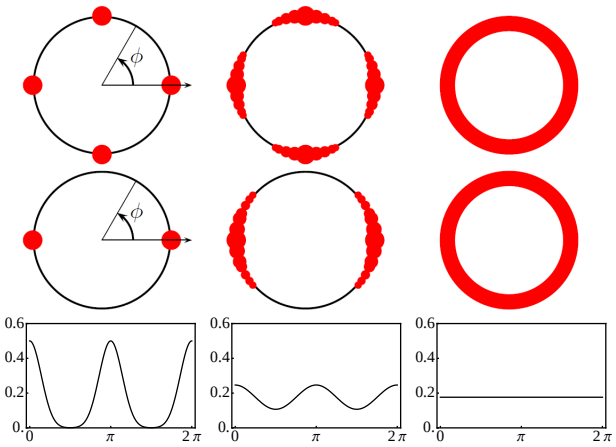

Figure 3.3: Phase relaxation along the equator of the Bloch sphere. Top panels: superposition with $q = 4$ (time $t_4 = T/8$) and $a_4 = 0,\ 0.64,\ 2.05$ (from left to right). Middle panels: superposition with $q = 2$ (time $t_2 = T/4$) for the same noise intensities $\int_0^\infty \mathrm{d}\tau\, k(\tau)$ in the Markov regime (see (3.18)), i.e., $a_2 = 0,\ 0.9,\ 2.9$. The circle sizes illustrate qualitatively the phase distribution $f(\phi, t_{2,4})$. For intermediate noise (middle column), the superposition is closer to the steady state (last column) for $q = 4$ than for $q = 2$. Bottom panels: Husimi distribution (3.6) of $[\rho(t_2)]_\mathrm{d}$ as a function of ϕ for $\theta = \pi/2$, for the same values of a_2 and $\overline{\lambda} = 0$, $N_0 = 10$.

$\mathrm{e}^{-a^2(t)/2}$ with respect to the noiseless case. A Gaussian decay is observed experimentally [107], indicating that the correlation time τ_c is much larger than the time of formation of squeezed states (non Markovian regime).

Effect of the phase noise on macroscopic superpositions.

Let us now study the impact of decoherence and phase relaxation on the superposition (3.9) at time $t = t_q$ as a function of $a_q = a(t_q)$. In view of (3.19), both the diagonal and off-diagonal parts of the noiseless density matrix are multiplied in the Fock basis by the decoherence factor $\mathrm{e}^{-D_q(n_1,n_1')} = e^{-a_q^2(n_1-n_1')^2/2}$. Using also the structure (3.13) of this matrix, we infer that

1) $[\rho(t_q)]_\mathrm{d} \to \rho(\infty)$ when $a_q q \gg 1$ (phase relaxation);

2) $[\rho(t_q)]_\mathrm{od} \to 0$ when $a_q \gg 1$ (decoherence).

As corroborated by Fig. 3.4, in the strong noise limit the diagonal part of $\rho(t)$ relaxes to the steady state (3.20) and the off-diagonal part is washed away. Remarkably, the decoherence factor e^{-D_q} does not depend on the atom number N_0. This number gives the separation on the Bloch sphere between the coherent states of the two-component superposition. Also note the different noise scales relevant for decoherence, a_q, and for phase relaxation, $a_q q$. Hence, when increasing the noise intensity, $[\rho(t_q)]_\mathrm{d}$ approaches $\rho(\infty)$ before $[\rho(t_q)]_\mathrm{od}$ vanishes. The higher the number of components q in the superposition, the more pronounced is this effect. In fact, superpositions with higher q are less affected by decoherence since they are formed at shorter times and $a(t)$ increases with time. Conversely, phase relaxation has a stronger effect on superpositions with higher q in the Markov regime[7], as illustrated in Fig. 3.3. As a consequence, by increasing the noise intensity the superposition (3.9) is not transformed into a statistical mixture of coherent states but relaxes directly to the mixture of Fock states (3.20).

The surprising fact that decoherence is not enhanced by increasing the distance on the Bloch sphere between the coherent states in the superposition is specific to the noise considered. Indeed, the noise in the Hamiltonian (3.14) is applied *perpendicularly* to the equator of the Bloch sphere where the coherent states lay. As a result, it is *insensitive to the separation* between these states. The superpositions would be much more affected by a J_x or J_y noise parallel to the equatorial plane, since such noises would "see" the separation between the components. We also stress that, even though decoherence is not due here to a coupling with a reservoir, exactly the same dynamics would have been obtained for the reduced density matrix of a BJJ interacting with a bath via a Hamiltonian of the form (1.15) with X replaced by J_z.

[6]This is the analog of tracing out the bath degrees of freedom for systems coupled to quantum baths (chapter 1).

[7]Indeed, in view of (3.13) and (3.19), the $n_1 \neq n_1'$ matrix elements of $[\rho(t_q)]_\mathrm{d}$ are damped in the Markov regime by a factor equal to or smaller than $\exp(-a_q^2 q^2/2) \simeq \exp[-(\pi q/\chi)\int_0^\infty \mathrm{d}\tau k(\tau)]$. In contrast, in the small-time regime all the q-component superpositions relax to $\rho(\infty)$ at the same noise intensity $k(0)$, since $a_q q$ is independent of q.

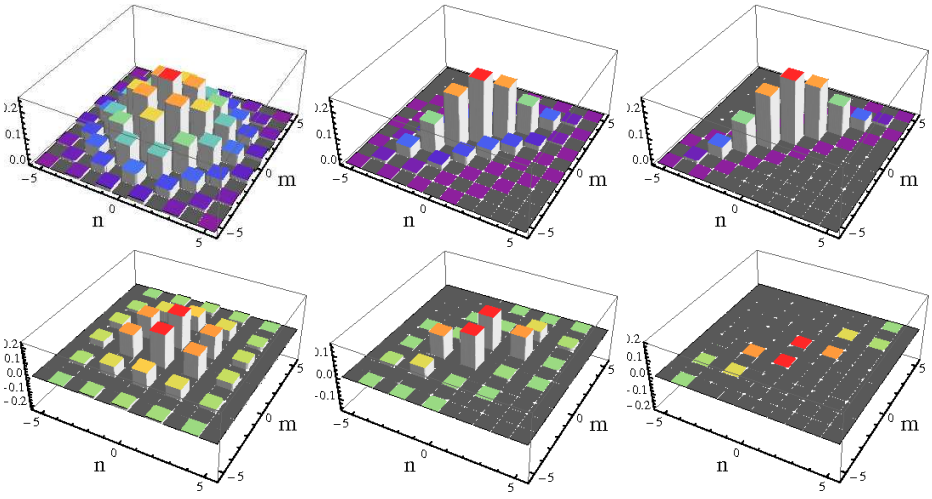

Figure 3.4: **Top panels:** relaxation of the diagonal part of the density matrix $\langle n, N_0-n|[\rho(t_2)]_{\rm d}|m, N_0-m\rangle$ in the Fock basis to the diagonal matrix (3.20) as the noise is increased from $a_2=0$ (left) to $a_2=0.9$ (middle) and $a_2=2.9$ (right). The time $t=t_2$ is the formation time of the superposition with $q=2$ and one has $N_0=10$. **Bottom panels:** off-diagonal part of the density matrix, whose decay to zero indicates decoherence among the coherent states of the superposition, for the same noise strengths a_2 and time $t=t_2$.

3.3.2 Dynamics in the presence of atom losses

Master equation and quantum jumps.

Scattering events lead to atom losses in the trap and constitute the main source of decoherence in the Heidelberg and Basel experiments. In fact, the phase noise considered in the previous subsection can be decreased by using a spin-echo technique [106]. Three kinds of loss processes may play a role: one-body losses, due to inelastic collisions between trapped atoms and the background gas; two-body losses, resulting from scattering of two atoms in the magnetic trap, which changes their spin and gives them enough kinetic energy to be ejected from the trap; and three-body losses, where a three-body collision event produces a molecule and ejects a third atom out of the trap.

Let $m_{\nu,1}$ and $m_{\nu,2}$ be the number of atoms lost in mode 1 and 2 during the νth loss process, with $|m_\nu| = m_{\nu,1}+m_{\nu,2}=r$ for a r-body loss. We introduce the jump operators

$$M_{m_\nu} = a_1^{m_{\nu,1}} a_2^{m_{\nu,2}} \tag{3.21}$$

and loss rates Γ_{m_ν}. For external BJJs one has $\Gamma_{m_\nu}=0$ for $m_{\nu,1}m_{\nu,2}>0$ (no inter-mode losses). We account for the three aforementioned loss processes by considering the Markovian Lindblad master equation [10, 137, 138]

$$\frac{{\rm d}\rho}{{\rm d}t} = -{\rm i}\big[H_0,\rho(t)\big] + \sum_{m\in\{1,2,3\}^2,|m|\leq 3} \Gamma_m\Big(M_m\,\rho\,M_m^\dagger - \frac{1}{2}\big\{M_m^\dagger M_m,\rho\big\}\Big)\,, \tag{3.22}$$

where $\rho(t)$ is the atomic density matrix[8].

Equation (3.22) does not couple sectors with different numbers of atoms N. If the matrix elements of $\rho(t)$ between states with different N vanish initially, this will be the case at all times $t\geq 0$. Then

$$\rho(t) = \sum_{N=0}^{N_0} \widetilde{\rho}_N(t) \quad , \quad \widetilde{\rho}_N(t) = w_N(t)\rho_N(t)\,, \tag{3.23}$$

where $\widetilde{\rho}_N(t)$ (respectively $\rho_N(t)$) is the unnormalized (normalized) density matrix with a well-defined[9] number of atoms N and $w_N(t)\geq 0$ is the probability of finding N atoms at time t (thus $\sum_N w_N(t)=1$). The conditional

[8]The loss rates Γ_m actually depend on the macroscopic wave function of the condensate [137, 138] and thus on the number of atoms N in the BJJ. As far as the number of lost atoms at the revival time T remains small with respect to the initial atom number N_0, one may nevertheless assume that Γ_m are time-independent in the time interval $[0,T]$.

[9]This means that $\langle n_1,n_2|\rho_N(t)|n_1',n_2'\rangle = 0$ for $n_1+n_2\neq N$ or $n_1'+n_2'\neq N$.

state $\rho_N(t)$ describes the state of the BJJ when one post-selects among many single-run measurements those yielding a total number of atoms at time t equal to N. In this sense, $\rho_N(t)$ contains a more precise physical information[10] than $\rho(t)$.

When inter-mode losses are absent (like for external BJJs), the master equation (3.22) can be solved in the Fock basis by exact diagonalization; however, for two- and three-body losses the solution has a quite complicated form and the computation of expectation values or other quantities depending on $\rho(t)$ must be done numerically[11]. Alternatively, quantum trajectories provide a natural and efficient tool to study the conditional states $\rho_N(t)$. We use here the quantum jump approach (Sec. 2.2.1). Let J be the number of loss events (jumps) in the time interval $[0,t]$. We write $\mathbf{m}=(m_1,\ldots,m_J)\in\{1,2,3\}^{2J}$ the random sequence of loss types $m_\nu=(m_{\nu,1},m_{\nu,2})$ and $\mathbf{s}=(s_1,\ldots,s_J)$ the random sequence of loss times $0\le s_1\le\cdots\le s_J\le t$. The total number of atoms ejected from the condensate between times 0 and t is equal to $|\mathbf{m}|=\sum_{\nu=1}^J(m_{\nu,1}+m_{\nu,2})$. The random wave function at time t reads

$$\begin{aligned}|\psi_J(t)\rangle &= \frac{|\widetilde{\psi}_J(t)\rangle}{\|\widetilde{\psi}_J(t)\|}\\ |\widetilde{\psi}_J(t)\rangle &= \mathrm{e}^{-\mathrm{i}(t-s_J)H_{\rm eff}}M_{m_J}\mathrm{e}^{-\mathrm{i}(s_J-s_{J-1})H_{\rm eff}}M_{m_{J-1}}\cdots\mathrm{e}^{-\mathrm{i}(s_2-s_1)H_{\rm eff}}M_{m_1}\mathrm{e}^{-\mathrm{i}s_1H_{\rm eff}}|\psi(0)\rangle\ ,\end{aligned}\tag{3.24}$$

where $H_{\rm eff}=H_0-\mathrm{i}\sum_m\Gamma_mM_m^\dagger M_m/2$ is defined by (2.3). Using (2.6) and (2.9), one finds that the probability to have J loss events between times 0 and t, with the νth event of type m_ν occurring in the time interval $[s_\nu,s_\nu+\mathrm{d}s_\nu]$, $\nu=1,\ldots,J$, reads

$$\mathrm{d}p_{\mathbf{m}}^{(t)}(\mathbf{s};J)=\Gamma_{m_1}\cdots\Gamma_{m_J}\big\|\widetilde{\psi}_J(t)\big\|^2\mathrm{d}s_1\cdots\mathrm{d}s_J\ .\tag{3.25}$$

As pointed out in Sec. 2.2.2, the density matrix solution of the master equation (3.22) is obtained by averaging the projector on $|\psi_J(t)\rangle$ over all quantum trajectories (that is, over the number of jumps J and the jump times s_ν and types m_ν). We thus recover the block structure (3.23) of $\rho(t)$ with

$$\widetilde{\rho}_N(t)=\sum_{J=1}^N\sum_{\mathbf{m}}\delta_{N_0,N+|\mathbf{m}|}\,\Gamma_{m_1}\cdots\Gamma_{m_J}\int_{0\le s_1\le\cdots\le s_J\le t}\mathrm{d}s_1\cdots\mathrm{d}s_J\,|\widetilde{\psi}_J(t)\rangle\langle\widetilde{\psi}_J(t)|\ .\tag{3.26}$$

Conditional state in the subspace with N_0 atoms.

We start by determining the unormalized conditional state with the initial number of atoms N_0,

$$\widetilde{\rho}_{N_0}^{(\text{no loss})}(t)=|\widetilde{\psi}_0(t)\rangle\langle\widetilde{\psi}_0(t)|\quad,\quad|\widetilde{\psi}_0(t)\rangle=\mathrm{e}^{-\mathrm{i}tH_{\rm eff}}|N_0;\phi=0\rangle\,.\tag{3.27}$$

This corresponds to the contribution of quantum trajectories with no jump in the time interval $[0,t]$. In the Fock basis diagonalizing both the Hamiltonian (3.1) (recall that we set the tunneling amplitude K to zero) and the damping terms $-\mathrm{i}\Gamma_mM_m^\dagger M_m/2$, the conditional state takes the form

$$\langle n_1,n_2|\widetilde{\rho}_{N_0}^{(\text{no loss})}(t)|n_1',n_2'\rangle=\mathrm{e}^{-t[d_{N_0}(n_1)+d_{N_0}(n_1')]}\langle n_1,n_2|\rho_{N_0}^{(0)}(t)|n_1',n_2'\rangle\ ,\tag{3.28}$$

where $\rho_{N_0}^{(0)}(t)$ is the state (3.10) in the absence of losses, $d_{N_0}(n_1)=\sum_m\Gamma_m\langle n_1,n_2|M_m^\dagger M_m|n_1,n_2\rangle/2$, and $n_2=N_0-n_1$, $n_2'=N_0-n_1'$. We restrict here our attention to symmetric three-body losses[12], i.e., $\Gamma_{3,0}=\Gamma_{0,3}$ and $\Gamma_{1,2}=\Gamma_{2,1}$. Let us set

$$a=\frac{1}{2}\big(\Gamma_{2,0}+\Gamma_{0,2}-\Gamma_{1,1}\big)+(N_0-2)\kappa\quad,\quad\kappa=\frac{3}{2}\Gamma_{3,0}-\frac{1}{2}\Gamma_{2,1}\,.\tag{3.29}$$

If $a>0$, the damping factor in (3.28) (exponential factor in the right-hand side) is Gaussian: in fact, then

$$d_{N_0}(n_1)=a(n_1-\overline{n}_1)^2\tag{3.30}$$

up to an irrelevant n_1-independent constant that can be absorbed in the normalization of the state, with

$$\overline{n}_1=\frac{1}{4a}\Big(\Delta\Gamma_1-\Delta\Gamma_2+N_0(2\Gamma_{0,2}-\Gamma_{1,1})+2N_0(N_0-2)\kappa\Big)\quad,\quad\Delta\Gamma_r=\Gamma_{0,r}-\Gamma_{r,0}\,.\tag{3.31}$$

[10] To have access to this information, one must be able to extract samples with a well-defined atom number initially (since we assumed an initial state with N_0 atoms) and after the evolution time t. Even though the precise measurement of N is still an experimental challenge, the precision has increased by orders of magnitude during the last years.

[11] See the Appendix A of [228] for more detail.

[12] See [228] for the study of the asymmetric three-body loss case, for which $d_{N_0}(n_1)$ is cubic in n_1.

In order to estimate the typical loss rates at which the state (3.28) is affected by the Gaussian damping, we focus on two particular cases.

(i) Symmetric loss rates, $\Delta\Gamma_r = 0$, $r = 1, 2, 3$. In this case $\overline{n}_1 = N_0/2$ and the damping factor in (3.28) is a Gaussian centered at $(n_1, n_1') = (N_0/2, N_0/2)$. This center coincides with the peak of the matrix elements in the absence of losses, which have a width $\sqrt{N_0}$, see (3.11). Thus the effect of the Gaussian damping begins to set in for times t such that $at \approx 1/N_0$. In particular, the macroscopic superposition at time t_q is noticeably affected by damping for $a \gtrsim \chi q/N_0$.

(ii) Completely asymmetric two-body losses and no three-body losses, $\Gamma_{0,2} = \Gamma_{1,1} = \kappa = 0$. Then $\overline{n}_1 = \Delta\Gamma_1/(2\Gamma_{2,0}) + 1/2$. The onset of the damping on the q-component superposition is at the loss rate $\Gamma_{2,0} \approx \chi q/N_0^{3/2}$, which is smaller by a factor of $\sqrt{N_0}$ compared with case (i), except for strongly asymmetric one-body loss rates satisfying $\Delta\Gamma_1 \approx \Gamma_{2,0}N_0$ (in which case this onset occurs when $\Gamma_{2,0} \approx \chi q/N_0$ as in (i)). Therefore, when $\Delta\Gamma_1$ is not of the order of $\Gamma_{2,0}N_0$, the Gaussian damping affects more strongly the superpositions than in the symmetric case.

Conditional state in the subspace with $N < N_0$ atoms.

We now proceed to study the contribution to the total density matrix $\rho(t)$ of quantum trajectories having $J \geq 1$ jumps in the time interval $[0, t]$. It is easy to see that each jump (2.7) with jump operator (3.21) transforms a coherent state $|N_0; \theta, \phi\rangle$ into a coherent state $|N_0 - |m|; \theta, \phi\rangle$. This coherent state is rotated on the Bloch sphere by the evolution between jumps driven by the effective Hamiltonian $H_{\rm eff}$. This is due to the non-linearity of $H_{\rm eff}$ in n_i and to the different numbers of atoms in the BJJ in the time intervals $[0, s_1]$, $[s_1, s_2], \ldots, [s_J, t]$, leading to different interaction energies. More precisely, for three-body loss rates satisfying $\Gamma_m \ll (N_0 t)^{-1}$, $|m| = 3$, the wave function (3.24) is proportional to

$$|\psi_J(t)\rangle \;\propto\; \mathrm{e}^{-\mathrm{i}tH_{\rm eff}}\big|N_0 - |\mathbf{m}|; \theta_{\mathbf{m}}(\mathbf{s}), \phi_{\mathbf{m}}(\mathbf{s})\big\rangle \;, \tag{3.32}$$

where $\theta_{\mathbf{m}}(\mathbf{s})$ and $\phi_{\mathbf{m}}(\mathbf{s})$ are random angles depending on the loss types m_ν and loss times s_ν. These angles are given by

$$\theta_{\mathbf{m}}(\mathbf{s}) = 2\arctan\Big(\exp\Big\{-\sum_{\nu=1}^{J}\frac{s_\nu}{2}\big(\delta_1 m_{\nu,1} + \delta_2 m_{\nu,2}\big)\Big\}\Big) \quad , \quad \phi_{\mathbf{m}}(\mathbf{s}) = \sum_{\nu=1}^{J} s_\nu(\chi_1 m_{\nu,1} + \chi_2 m_{\nu,2}) \;, \tag{3.33}$$

where we have introduced the interaction energies

$$\chi_1 = U_1 - U_{12} \quad , \quad \chi_2 = -(U_2 - U_{12}) \;, \tag{3.34}$$

and the loss rate differences

$$\delta_1 = 2\Gamma_{2,0} - \Gamma_{1,1} + (3\Gamma_{3,0} - \Gamma_{1,2})N_0 \quad , \quad \delta_2 = -\big(2\Gamma_{0,2} - \Gamma_{1,1} + (3\Gamma_{0,3} - \Gamma_{2,1})N_0\big) \;. \tag{3.35}$$

To give an insight on the derivation of these formulas, let us consider the special case of a trajectory $t \mapsto |\psi_1(t)\rangle$ having a single loss event of type $m = (2, 0)$ at time $s \in [0, t]$, for a BJJ subject to two-body losses only[13]. We first determine how an initial Fock state $|n_1, n_2\rangle$ is transformed when two atoms are lost in mode 1 at time s. Using (3.24), this state becomes $\sqrt{n_1(n_1-1)}e^{-\mathrm{i}\Phi_{t,s}(n_1,n_2)}|n_1 - 2, n_2\rangle$ at time t, where $\Phi_{t,s}(n_1, n_2)$ is a complex dynamical phase. Setting $n_2 = N_0 - n_1$ and denoting by $H_{\rm eff}(n_1, n_2)$ the quadratic eigenvalues of $H_{\rm eff}$, this phase reads

$$\begin{aligned}\Phi_{t,s}(n_1, n_2) &= (t-s)H_{\rm eff}(n_1 - 2, n_2) + sH_{\rm eff}(n_1, n_2) \\ &= tH_{\rm eff}(n_1 - 2, n_2) + n_1\phi_1(s) + \mathrm{i}\, n_1 \ln\Big(\tan\Big(\frac{\theta_1(s)}{2}\Big)\Big) + c_{t,s} \;,\end{aligned} \tag{3.36}$$

where $c_{t,s}$ is independent of n_1 and $\theta_1(s)$, $\phi_1(s)$ are given by (3.33) with $J = 1$, $m_{\nu,1} = 2$, $m_{\nu,2} = 0$, and $\Gamma_m = 0$ for $|m| \neq 2$. The second term in the last line of (3.36) corresponds to the dynamical phase associated to the change in the atomic interaction energy because of the reduction of particles at time s, and the third term describes a corresponding change in the damping. Replacing $|n_1, n_2\rangle$ in the Fock-state expansion (3.5) of the initial coherent state by the above transformed state, we obtain the formula (3.32).

[13] See [228] for a more general derivation of (3.32)-(3.35) including also one- and three-body losses and an arbitrary number of loss events.

Equation (3.32) means that, apart from damping effects due to the non self-adjoint part in the effective Hamiltonian $H_{\rm eff}$, atom losses can be accounted for by external noises rotating the state around the Bloch sphere. More precisely, the conditional state $\rho_N(t)$ is the same state as if there were no atom loss, one had initially N atoms, and the BJJ was subject to phase noises in the angles ϕ and θ described by the random Hamiltonian (3.14) but with a complex stochastic process $\lambda(t)$ satisfying $\int_0^t \mathrm{d}\tau\lambda(\tau) = -\phi_{\mathbf{m}}(\mathbf{s}) - \mathrm{i}\ln(\tan(\theta_{\mathbf{m}}(\mathbf{s}))/2$ and with an additional damping term. This analogy between atom losses and phase noise has been discovered independently by Sinatra, Dornstetter, and Castin [219]. For a single loss event ($J = 1$), the noise fluctuations have magnitude

$$\delta\theta_{m_1} \simeq \frac{1}{2}\delta s_{m_1}\left|\sum_{i=1,2}\delta_i m_{1,i}\right| \quad , \quad \delta\phi_{m_1} = \delta s_{m_1}\left|\sum_{i=1,2}\chi_i m_{1,i}\right| \tag{3.37}$$

(we assume here $\delta\theta_{m_1} \ll 1$), where δs_{m_1} is the fluctuation of the loss time s.

The conditional density matrix $\rho_N(t)$ with $N < N_0$ atoms can be determined by summing over all trajectories having J loss events such that $|\mathbf{m}| = N_0 - N$, see (3.26). This requires the evaluation of the norm $\|\widetilde{\psi}_j(t)\|$. The calculation does not present difficulties, but it is maybe the longest one I have made in the last years, because of the large number of jump operators and rates in the problem. The result is that $\rho_N(t)$ is given in the Fock basis by

$$\langle n_1, n_2|\rho_N(t)|n_1', n_2'\rangle = \mathcal{E}_N(t; n_1, n_1')\langle n_1, n_2|\widetilde{\rho}_N^{(\text{no loss})}(t)|n_1', n_2'\rangle \ , \tag{3.38}$$

where $\widetilde{\rho}_N^{(\text{no loss})}(t)$ is the density matrix conditioned to no loss event for an initial coherent state with N atoms, see (3.28), and $\mathcal{E}_N(t; n_1, n_1')$ is an envelope depending on time and on the matrix entries n_1 and n_1'. For $N = N_0 - r$ and if only r-body losses occur, this envelope is

$$\mathcal{E}_N^{(\text{1-jump})}(t; n, n') \ \propto \sum_{m,|m|=r} \Gamma_m C_{N,m}(t; n, n') \tag{3.39}$$

with

$$C_{N,m_1}(t; n, n') = \frac{1 - \mathrm{e}^{-t[G_{m_1} + (\delta_1 m_{1,1} + \delta_2 m_{1,2})(n+n'-N)/2 + \mathrm{i}(\chi_1 m_{1,1} + \chi_2 m_{1,2})(n-n')]}}{G_{m_1} + (\delta_1 m_{1,1} + \delta_2 m_{1,2})(n + n' - N)/2 + \mathrm{i}(\chi_1 m_{1,1} + \chi_2 m_{1,2})(n - n')} \ , \tag{3.40}$$

where G_{m_1} is a function of N_0, r, m_1, and Γ_{m_1} that we shall not give explicitly here (see [228] for more detail). If, in addition to the above condition on three-body losses, one assumes that[14] the two-body loss rates satisfy $\Gamma_m \ll t^{-1}$, $|m| = 2$, and that the total number $N_0 - N$ of atoms lost between times 0 and t is much smaller than N_0, the envelope in (3.38) takes a simple form in terms of the single jump envelopes,

$$\mathcal{E}_N(t; n, n') = \sum_{J_1,J_2,J_3\geq 0, J_1+2J_2+3J_3=N_0-N} \frac{1}{J_1!J_2!J_3!}\prod_{r=1}^{3}\left[\mathcal{E}_{N_0-r}^{(\text{1-jump})}(t; n, n')\right]^{J_r} . \tag{3.41}$$

3.4 Time evolution of the quantum Fisher information

We introduce in this section the notion of quantum Fisher information and its interpretation in terms of the best achievable phase sensitivity in interferometry. We then use this Fisher information to quantify the amount of quantum correlations in the atomic states of a BJJ in the presence of atom losses as a function of time. A more detailed presentation on the Fisher information and phase estimation will be given in chapter 9 below.

3.4.1 Interferometric sequence

The goal of interferometry is to estimate an unknown phase shift ϕ with the highest possible precision. We assume here that the reader is familiar with optical Mach Zehnder interferometers (otherwise a quick look to Sec. 9.2.1 and Fig. 9.1 might be helpful!). In atom interferometry, an input state is first transformed into a superposition of two modes, analogous to the two arms of an optical interferometer. These modes acquire distinct phases ϕ_1 and ϕ_2 during the subsequent quantum evolution. They are finally recombined to read out interference fringes, from which the phase difference $\phi = \phi_1 - \phi_2$ is inferred. The interferometric sequence can

[14] These are not strong restrictions since for large N_0 the mean number $\langle N\rangle_t$ of atoms at time t when the BJJ is subject to two- or three-body losses is a function of $N_0^{|m|-1}\Gamma_m t$ with $|m| = 2$ or 3, respectively (this follows e.g. from the phenomenological rate equations, which yield $\langle N\rangle_t \simeq N_0(\Gamma_m N_0 t+1)^{-1}$ and $N_0(2\Gamma_m N_0^2 t+1)^{-1/2}$ for symmetric two- and three-body losses, respectively). Hence our conditions are still fulfilled if a large fraction (e.g. 50%) of the initial atoms are lost between times 0 and t.

be described by means of rotations generated by the momentum operators (3.2). Let us consider the case where the two modes correspond to two internal states of the atoms in a trapped BEC (internal BJJ). The N atoms are initially in the lower energy state (mode $i = 1$) and the input state is the Fock state $|n_1 = N, n_2 = 0\rangle$. The application of a $\pi/2$ pulse with frequency in resonance with the two internal levels plays the role of a beam splitter in optical interferometers. It brings the input state onto the coherent state $|N; \theta = \pi/2, \phi = 0\rangle$. Then the state is rotated around the z-axis by the free evolution. This rotation implements the different phases accumulated in the two arms of an optical interferometer [273]. The consecutive rotations on the Bloch sphere are represented in Fig. 3.2. Finally, the state is rotated again around the y-axis by an angle of $-\pi/2$ radians to recombine the two paths. Since the composition of rotations $\mathrm{e}^{\mathrm{i}\frac{\pi}{2}J_y}\mathrm{e}^{-\mathrm{i}\phi J_z}\mathrm{e}^{-\mathrm{i}\frac{\pi}{2}J_y}$ is equivalent to a rotation by an angle $-\phi$ around (Ox), the output state of the interferometer reads

$$\rho_{\mathrm{out}}(\phi) = \mathrm{e}^{-\mathrm{i}\phi J_{\mathbf{n}}}\rho_{\mathrm{in}}\mathrm{e}^{\mathrm{i}\phi J_{\mathbf{n}}} \, , \tag{3.42}$$

where ρ_{in} is the input state and the unit vector $\mathbf{n} = -\mathbf{e}_x$ defines the interferometer direction. The rotations are realized fast enough so that the non-linearity induced by the interactions and atom losses can be neglected during the whole interferometric sequence [106].

3.4.2 Phase estimation and quantum Fisher information

The phase shift ϕ is determined by means of a statistical estimator depending on the results of measurements on the output state $\rho_{\mathrm{out}}(\phi)$. As argued in Sec. 9.2 below, the best precision that can be achieved (that is, optimizing over all possible estimators and measurements) is given by

$$(\Delta\phi)_{\mathrm{best}} = \frac{1}{\sqrt{\mathcal{N}\,\mathcal{F}_Q(\rho_{\mathrm{in}}, J_{\mathbf{n}})}} \, , \tag{3.43}$$

where $\mathcal{N}$ is the number of measurements and $\mathcal{F}_Q(\rho, J_{\mathbf{n}})$ is the quantum Fisher information defined by Eq. (8.42) in chapter 8. The latter quantity thus measures the amount of quantum correlations in ρ useful for interferometry. If $\mathcal{F}_Q(\rho_{\mathrm{in}}, J_{\mathbf{n}}) > \langle N\rangle$, a better phase sensitivity than the shot noise value $(\Delta\phi)_{\mathrm{SN}} = 1/\sqrt{\mathcal{N}\,\langle N\rangle}$, which corresponds to using as input $\langle N\rangle$ independent atoms (coherent state), is obtained. The highest value of the Fisher information for N_0 atoms is $\mathcal{F}_Q(\rho_{\mathrm{in}}, J_{\mathbf{n}}) = N_0^2$. It can be shown that $\mathcal{F}_Q(\rho_{\mathrm{in}}, J_{\mathbf{n}}) \geq \langle N\rangle$ implies that the atoms are entangled [134] and that high values of $\mathcal{F}_Q(\rho_{\mathrm{in}}, J_{\mathbf{n}})$ imply multipartite entanglement between a large number of atoms [135, 236].

Since $J_{\mathbf{n}}$ does not couple subspaces with different N's, it follows from (8.42) and from the block structure (3.23) of ρ that

$$\mathcal{F}_Q(\rho, J_{\mathbf{n}}) = \sum_{N=0}^{N_0} w_N \mathcal{F}_Q(\rho_N, J_{\mathbf{n}}) \, , \tag{3.44}$$

where $\mathcal{F}_Q(\rho_N, J_{\mathbf{n}})$ is the Fisher information of the conditional state ρ_N with N atoms and w_N is the corresponding probability.

In order to obtain a measure of quantum correlations independent of the direction $\mathbf{n}$ of the interferometer, we optimize the Fisher information over all unit vectors $\mathbf{n}$ and define [133],

$$\mathcal{F}_Q(\rho) = \max_{\|\mathbf{n}\|=1} \mathcal{F}_Q(\rho, J_{\mathbf{n}}) = 4C_{\max} \, . \tag{3.45}$$

From the expression (8.42) of $\mathcal{F}_Q(\rho, J_{\mathbf{n}})$, one easily sees that $C_{\max}$ is the largest eigenvalue of the 3×3 real symmetric covariance matrix

$$C_{ab} = \frac{1}{2} \sum_{k,l,p_k+p_l>0} \frac{(p_k - p_l)^2}{p_k + p_l} \mathrm{Re}\left\{\langle k|J_a|l\rangle\langle l|J_b|k\rangle\right\} \quad , \quad a, b = 1, 2, 3 \, , \tag{3.46}$$

where $|k\rangle$ and p_k stand for the eigenvectors and eigenvalues of ρ. For simplicity, we write in the following $\mathcal{F}_{\mathrm{tot}}(t) \equiv \mathcal{F}_Q(\rho(t))$ for the optimized Fisher information of the total atomic density matrix $\rho(t)$ of the BJJ at time t. When studying the quantum correlations of the conditional states, we optimize over $\mathbf{n}$ independently in each N-atom subspace and define[15] $\mathcal{F}_N(t) \equiv \mathcal{F}_Q(\rho_N(t))$.

[15] Note that $\mathcal{F}_{\mathrm{tot}}(t)$ is not equal to $\sum_N w_N(t)\mathcal{F}_N(t)$, because the optimal directions may be different in each subspace.

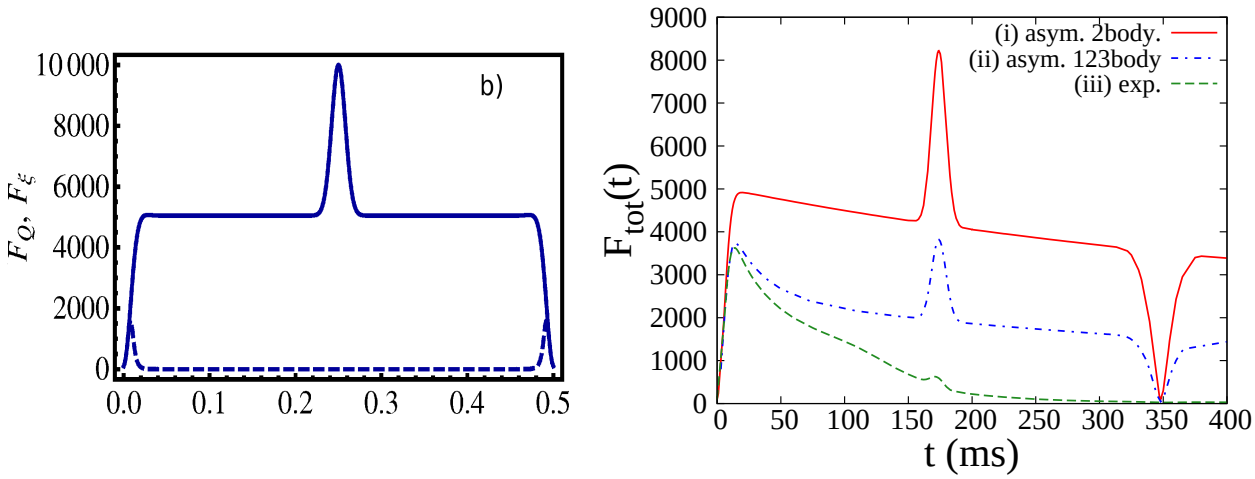

Figure 3.5: **Left panel:** Fisher information $F_{\rm tot}(t)$ during the quenched dynamics of a BJJ with $N_0 = 100$ atoms as a function of time (in units of the revival time T) in the absence of losses. The dashed line represents the amount of squeezing $F_\xi(t) = N_0/\xi(t)^2$ [197]. **Right panel:** same in the presence of atom losses (time in ms) for $U_2 = U_{12}$, $U_1 - U_{12} = 18.056\,$Hz, and (i) asymmetric two-body losses $\Gamma_{0,2} = 0.0127\,$Hz and $\Gamma_{2,0} = 0$ without one- and three-body losses; (ii) one-, two-, and three-body losses in the second mode with rates $\Gamma_{0,1} = 0.4\,$Hz, $\Gamma_{0,2} = 0.0127\,$Hz, $\Gamma_{0,3} = 1.08 \times 10^{-6}\,$Hz, and no losses in the first mode; (iii) symmetric one- and three-body losses and asymmetric two-body losses, $\Gamma_{1,0} = \Gamma_{0,1} = 0.2\,$Hz, $\Gamma_{0,2} = 0.0127\,$Hz, $\Gamma_{2,0} = 0$, and $\Gamma_{3,0} = \Gamma_{0,3} = 0.54 \times 10^{-6}\,$Hz. The case (iii) roughly corresponds to the experimental conditions in Refs. [106, 202].

3.4.3 Time evolution of the Fisher information in Bose-Josephson junctions

Let us study the time evolution of the optimized quantum Fisher information (3.45) in the BJJ after a sudden quench to zero of the inter-mode coupling K. In the absence of losses, the two-component superposition has the highest possible Fisher information $\mathcal{F}_Q[\rho^{(0)}(t_2)] = N_0^2$, which is for $N_0 \gg 1$ approximately twice larger than that of the superpositions with q components, $3 \leq q \lesssim N_0^{1/2}$ [197]. The time evolution of $\mathcal{F}_{\rm tot}(t) = \mathcal{F}_{N_0}(t)$ for an atomic sample with $N_0 = 100$ atoms is represented in the left panel of Fig. 3.5. At $t = 0$, the BJJ is in the coherent state (3.7) and $\mathcal{F}_{N_0}(0)$ has the shot noise value N_0. The Fisher information first increases, reaches a plateau with value $N_0(N_0+1)/2$ at time $t \approx T/\sqrt{N_0}$, and then displays a maximum N_0^2 at the time of formation $t = t_2 = T/4$ of the two-component superposition. It evolves afterward symmetrically back to the initial value N_0 at time $T/2$, when the atomic state $|\psi^{(0)}(t)\rangle$ coincides with the coherent state $|N_0; \phi = \pi\rangle$ (see Ref. (3c) in the publication list).

We now compare this evolution with that obtained for the same sample in the presence of atom losses for experimentally relevant loss rates extracted from Refs. [218, 202, 271]. The Fisher information is obtained in this case from an exact diagonalization of the master equation (3.22) and a numerical diagonalization of the density matrix $\rho(t)$ [228]. It is displayed in the right panel of Fig. 3.5. When the BJJ is subject to two-body losses in the mode $i = 2$ only (upper curve), we observe that the Fisher information exhibits a well pronounced peak at time t_2 as in the lossless case, showing that the quantum correlations of the superpositions are well preserved. This occurs provided that the atomic interaction energies U_i are chosen such that $U_2 = U_{12}$ (i.e., $\chi_2 = 0$), in such a way as to suppress the ϕ-noise given by (3.33) in the mode losing atoms. A much lower Fisher information would be obtained if one takes symmetric energies $U_1 = U_2$, keeping the same value for the revival time $T = 4\pi/(U_1 + U_2 - 2U_{12})$ (see below). In the Heidelberg and Basel experiments, two-body losses indeed occur mainly in the upper hyperfine level. When one- and three-body losses, which are also present in these experiments, are added, the coherences of the superpositions are still preserved provided that all losses occur in the second mode and one takes as before $U_2 = U_{12}$, as it can be inferred from the still relatively high values of the Fisher information in the middle curve of Fig. 3.5. However, when symmetric one- or three-body losses are added, the quantum correlations are destroyed much more rapidly and the peak in the Fisher information at time t_2 disappears. In the aforementioned experiments, one-body losses are symmetric since they are due to collisions with atoms from the background gas, which are equally likely for the two internal states. We conclude from Fig. 3.5 that for $U_2 = U_{12}$, these symmetric one-body losses are more detrimental to the macroscopic superpositions than the two-body losses in the second mode.

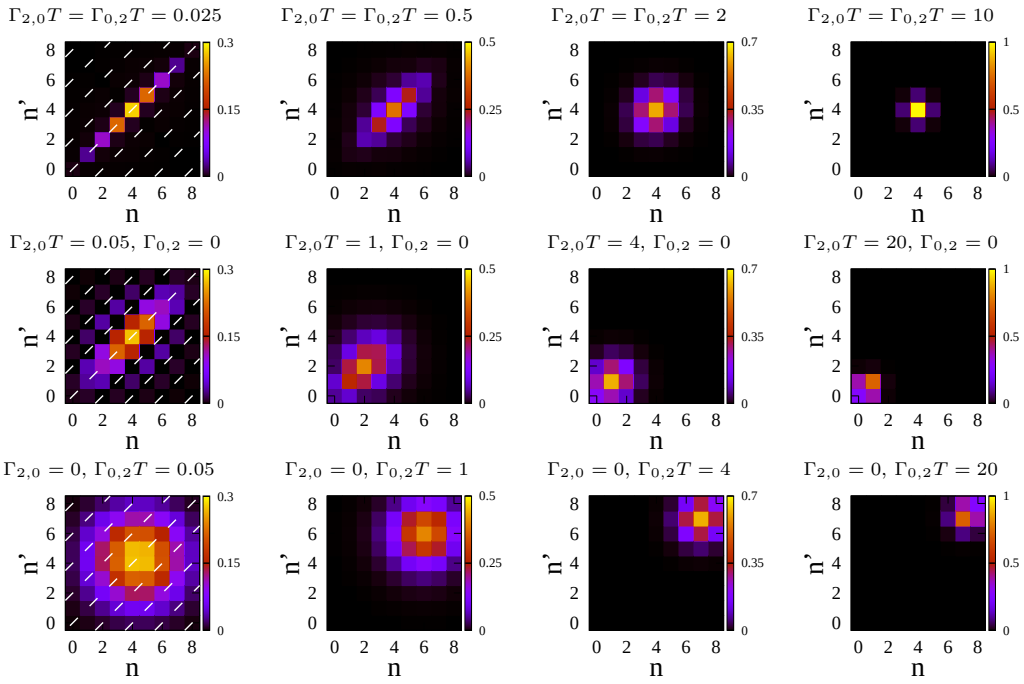

Figure 3.6: Moduli $|\langle n, N_1 - n|\rho_{N_1}(t_2)|n', N_1 - n'\rangle|$ of the density matrix in the Fock basis at time $t_2 = T/4$ in the subspace with $N_1 = N_0 - 2$ atoms for increasing two-body loss rates (from left to right). The upper, middle, and bottom panels correspond to symmetric losses and energies ($\Gamma_{2,0} = \Gamma_{0,2}$ and $U_1 = U_2$), asymmetric losses and symmetric energies ($\Gamma_{0,2} = 0$ and $U_1 = U_2$), and asymmetric losses and energies ($\Gamma_{2,0} = 0$ and $U_2 = U_{12}$), respectively. The revival time $T = 2\pi/\chi$ is the same in all cases. Panels in the same column have the same total loss rate $\Gamma_{2,0} + \Gamma_{0,2}$. White dashed lines are marking the values of (n, n') for which the matrix elements of the diagonal part $[\rho^{(0)}(t_2)]_{\rm d}$ of the two-component superposition do not vanish. Only two-body losses in the same mode are considered, i.e., $\Gamma_m = 0$ for $|m| = 1, 3$ and $m = (1, 1)$. The initial number of atoms is $N_0 = 10$.

3.5 Impact of a single loss event on macroscopic superpositions

The quantum Fisher information in Fig. 3.5 gives the amount of quantum correlations in the density matrix $\rho(t)$ averaged over all loss events. In this section, we investigate separately the quantum correlations in the conditional states $\rho_N(t)$ with given atom numbers N (Sec. 3.3.2). In particular, we would like to answer the following question: how does a *single* loss event, occurring at an unknown random time between $t = 0$ and the formation time t_q of a macroscopic superposition, affects the coherences in the superposition? In the case of photons, it is believed that macroscopic superpositions of coherent states loose their coherence completely after a single photon loss [113]. This means that the conditional state given that one photon has been lost is close to a separable mixture of coherent states. In what follows, we study this problem for a BJJ. With this aim, we determine the Fisher information $\mathcal{F}_{N_0-r}(t_q)$ in the conditional density matrix $\rho_{N_0-r}(t_q)$, where $r = 1, 2, 3$ for one-, two-, and three-body losses. We find quite different values depending on the degree of asymmetry between the loss rates and interaction energies in the two modes of the BJJ. This provides a qualitative explanation of the different behaviors of the total Fisher information in the presence of losses in Fig. 3.5.

3.5.1 Numerical results

Let us first describe the results obtained by my collaborator K. Pawlowski via an exact diagonalization of the master equation (3.22) and a subsequent numerical evaluation of the optimized Fisher information (3.45). For concreteness, we restrict ourselves to two-body losses, assuming no one-body, three-body, and inter-mode losses (i.e., all loss rates vanish save for $\Gamma_{2,0}$ and $\Gamma_{0,2}$). The density matrix $\rho_{N_1}(t_2)$ with $N_1 = N_0 - 2$ atoms is shown in Fig. 3.6 in the Fock basis. If the interaction energies in the two modes are equal, $U_1 = U_2$, we observe that $\rho_{N_1}(t_2)$ is almost diagonal in this basis for weak symmetric loss rates $\Gamma_{2,0} = \Gamma_{0,2} \lesssim \chi/N_0$ (upper left panel). In contrast, for completely asymmetric rates with $\Gamma_{0,2} = 0$, $\rho_{N_1}(t_2)$ has non-vanishing off-diagonal elements for odd values of $n_1' - n_1$ (middle left panel). Moreover, if there is no loss in the first mode and one tunes the

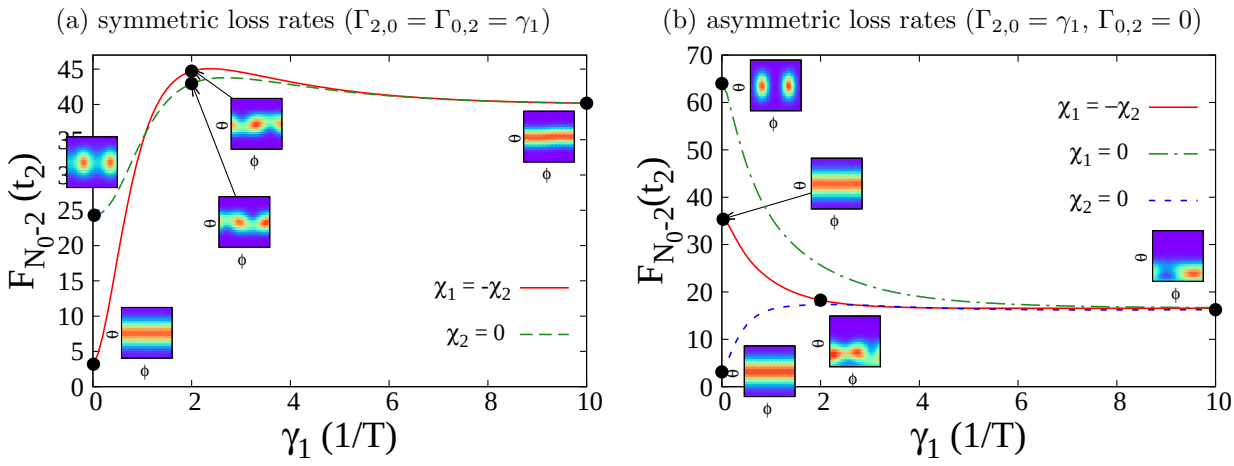

Figure 3.7: Fisher information optimized in the subspace with $(N_0 - 2)$ atoms at time t_2 as a function of the two-body loss rate (in units of T^{-1}) for (a) symmetric losses $\Gamma_{2,0} = \Gamma_{0,2} = \gamma_1$ with $U_1 = U_2$ (red solid line) and $U_2 = U_{12}$ (green dashed line); (b) completely asymmetric losses with $\Gamma_{0,2} = 0$ and $\Gamma_{2,0} = \gamma_1$, for $U_1 = U_2$ (red solid line), $U_1 = U_{12}$ (green dot dashed line), and $U_2 = U_{12}$ (blue dashed line). The energies U_i are chosen in such a way that the revival time $T = 2\pi/\chi$ does not change. **Insets:** plots of the Husimi distributions for some specific choices of loss rates (indicated by circles and arrows). Other parameters as in Fig. 3.6.

energies such that $U_2 = U_{12}$, keeping the effective interaction energy $\chi = (U_1 + U_2 - 2U_{12})/2$ fixed, the density matrix has the same structure as that of a two-component superposition with N_1 atoms (lower left panel). This is confirmed by looking at the Fisher information $\mathcal{F}_{N_1}(t_2)$ displayed in Fig. 3.7. If one of the two modes does not lose atoms and $U_i = U_{12}$ in the other mode, $\mathcal{F}_{N_1}(t_2)$ is approximately equal for small loss rates $\Gamma_m \ll \chi/N_0$ to the Fisher information N_1^2 of a two-component superposition (upper curve in the right panel). At stronger loss rates $\Gamma_m \approx \chi$, $\mathcal{F}_{N_1}(t_2)$ decreases to much lower values. Conversely, for symmetric losses and energies, $\mathcal{F}_{N_1}(t_2)$ starts below the shot-noise limit at weak losses and increases with Γ_m to reach a maximum when $\Gamma_m \simeq \chi/\pi$ (left panel).

3.5.2 Analytical results for small loss rates

The above numerical observations can be explained by relying on the quantum jump approach. The general results of Sec. 3.3.2 take a simpler form for small loss rates satisfying

$$\Gamma_m \ll N_0^{1-|m|} t_q^{-1} \quad , \quad |m| = 1, 2, 3 \, . \tag{3.47}$$

Then the θ-noise can be neglected since $\delta\theta_m$ is much smaller than the quantum fluctuations in the coherent states forming the components of the superposition (3.9) (the latter are of the order of $1/\sqrt{N_0}$). In contrast, due to large fluctuations $\delta s_m \approx t_q$ of the loss time, which has an almost flat distribution on $[0, t_q]$, the fluctuations of $\phi_m(s)$ are quite large. For instance, for symmetric interaction energies $U_1 = U_2$ (i.e., $\chi_1 = -\chi_2 = \chi$), (3.37) yields a fluctuation $\delta\phi_{m_1} \approx |m_{1,1} - m_{1,2}|\pi/q$ of the order of the phase separation $\phi_{k+1,q} - \phi_{k,q} = 2\pi/q$ between the coherent states, save for two-body inter-mode losses ($m_{1,1} = m_{1,2} = 1$) for which the phase noise vanishes.

The effect of this phase noise on the conditional density matrix $\rho_{N_1}(t_q)$ with $N_1 = N_0 - r$ atoms at time t_q is obtained in the Fock basis by multiplying the matrix elements of the superposition $\rho^{(0)}_{N_1}(t_q)$ in the absence of losses by a damping factor $\mathcal{D}_q(n_1, n_1')$ and an envelope $\mathcal{E}_{q,r}(n_1, n_1')$ given by[16]

$$\mathcal{D}_q(n, n') = \exp\Big\{ -\frac{\pi}{\chi q}\big(d_{N_1}(n) + d_{N_1}(n')\big)\Big\} \quad , \quad \mathcal{E}_{q,r}(n, n') = \frac{q\chi}{\pi} \sum_{|m|=r} \Gamma_m C_{N_1,m}(t_q; n, n') \, , \tag{3.48}$$

see (3.28), (3.38), and (3.39). In the small loss limit (3.47), the coefficient G_m in (3.40) can be neglected and $C_{N_1,m}(t_q; n, n')$ can be approximated for symmetric energies $U_1 = U_2$ by

$$C_{m_1}(t_q; n, n') = \frac{1 - \exp\big\{-\mathrm{i}\frac{\pi}{q}(m_{1,1} - m_{1,2})(n - n')\big\}}{\mathrm{i}\chi(m_{1,1} - m_{1,2})(n - n')} \, . \tag{3.49}$$

[16] Recall that we assume here the occurrence of a single r-body loss process in $[0, t_q]$. The factor $q\chi/\pi$ in front of the last sum is put for convenience and disappears in the state normalization.

We now argue that the behavior of the Fisher information $\mathcal{F}_{N_1}(t_q)$ shown in Fig. 3.7 comes from the combination of two effects: a channel effect for $U_1 = U_2$ and $q = 2, 3$, and the suppression of phase noise in the ith mode when $U_i = U_{12}$. Before discussing these two effects, we show that the phase noise always induces a complete phase relaxation when $U_1 = U_2$.

(i) *Complete phase relaxation for $U_1 = U_2$.*

Let us first study the impact of the phase noise on $\rho_{N_1}(t_q)$ for *symmetric interaction energies* $U_1 = U_2$ and small loss rates satisfying (3.47). We recall that phase noise flattens the Husimi distribution of the superposition in the ϕ direction (phase relaxation); for strong phase noises the diagonal part of the density matrix relaxes to a statistical mixture of Fock states with completely undefined phases (Sec. 3.3.1). We find that the loss of $r = 2$ atoms in the same mode leads to complete phase relaxation. Actually, from (3.13), (3.28), and (3.38), the matrix elements of $[\rho_{N_1}(t_q)]_{\rm d}$ in the Fock basis vanish for $n_1' \neq n_1$ modulo q. We may thus restrict our attention to $n_1' = n_1 + pq$ for integer p's. If $p \neq 0$, the envelope (3.48) reads

$$\mathcal{E}_{q,r}(n_1, n_1 + pq) \simeq \begin{cases} -\mathrm{i}\Delta\Gamma_1 \frac{1-(-1)^p}{\pi p} & \text{for } r = 1 \\ \Gamma_{1,1} & \text{for } r = 2 \\ -\mathrm{i}(\Delta\Gamma_3 + 3\Delta\Gamma_{3,\rm int}) \frac{1-(-1)^p}{3\pi p} & \text{for } r = 3 \end{cases} \tag{3.50}$$

with $\Delta\Gamma_{3,\rm int} = \Gamma_{1,2} - \Gamma_{2,1}$. Therefore, for weak two-body losses and $\Gamma_{1,1} = 0$, the diagonal part of $\rho_{N_1}(t_q)$ is equal to a statistical mixture of Fock states, that is, $\langle n_1, N_1 - n_1 | \, [\rho_{N_1}(t_q)]_{\rm d} | n_1', N_1 - n_1' \rangle = 0$ for $n_1 \neq n_1'$. This is confirmed in the upper and middle left panels of Fig. 3.6, where one observes vanishing matrix elements along the dashed lines $n_1' = n_1 \pm 2$, $n_1' = n_1 \pm 4$. This also explains the ϕ-independent profile of the Husimi distributions in Fig. 3.7 for $\chi_1 = -\chi_2$ and small $\Gamma_{2,0}$. For one- and three-body losses, complete phase relaxation occurs for symmetric losses ($\Delta\Gamma_1 = \Delta\Gamma_3 = \Delta\Gamma_{3,\rm int} = 0$) only[17]. Let us also stress that no phase relaxation occurs in the inter-mode channel $m = (1, 1)$.

(ii) *Loss of quantum correlations when $U_1 = U_2$ and $q = 2$ or 3: channels effects.*

Phase relaxation does not tell us anything about quantum correlations, which are contained in the off-diagonal part of the density matrix. We now determine this part, which corresponds to the matrix elements of $\rho_{N_1}(t_q)$ in the Fock basis such that $n_1' \neq n_1$ modulo q (see (3.13)). We still assume symmetric energies $U_1 = U_2$ and small losses satisfying (3.47). In view of (3.49), the main effect of phase noise is to multiply the matrix elements in the absence of noise by a factor of $(n_1 - n_1')^{-1}$ (decoherence). This factor decays to zero as one moves away from the diagonal but does not modify substantially the elements close to the diagonal. Indeed, non vanishing off-diagonal matrix elements for $n_1' = n_1 \pm 1$ and $n_1' = n_1 \pm 3$ are visible in the left middle panel of Fig. 3.6. This also explains the relatively high value of the Fisher information $\mathcal{F}_{N_0-2}(t_2)$ for $\Gamma_{2,0} \ll \chi/N_0$, $\Gamma_{0,2} = 0$, and $\chi_1 = -\chi_2$ in Fig. 3.7(b). For such loss rates and energies we are in the noise regime of weak decoherence pointed out in Sec. 3.3.1: phase noise is more efficient in washing out the phase content of each component of the superposition than in destroying the coherences between them.

However, we observe in Figs. 3.6 and 3.7 that the situation is quite different for symmetric two-body losses ($\Gamma_{2,0} = \Gamma_{0,2}$ and $\Gamma_{1,1} = 0$): then $[\rho_{N_1}(t_2)]_{\rm od}$ vanishes completely and the Fisher information at weak losses is smaller than N_0. This comes from a cancellation when adding the contributions of the $m = (2, 0)$ and $m = (0, 2)$ loss channels, which occurs only at time t_2 and in the absence of inter-mode losses[18]. In fact, for such loss rates $\mathcal{E}_{r,r}(n, n') \simeq 2\Gamma_{r,0}\delta_{n,n'}$ for $r = 2, 3$, thus $[\rho_{N_1}(t_2)]_{\rm od} = 0$ and the whole density matrix $\rho_{N_1}(t_2)$ is diagonal in the Fock basis. As a consequence of this channel effect, the two-component (three-component) superposition suffers in the absence of inter-mode losses from a *complete decoherence* in the N_1-atom subspace, for arbitrary small symmetric two-body (three-body) loss rates[19]. Such a channel effect does not occur for completely asymmetric losses involving only one channel.

We emphasize that complete decoherence does also not occur for symmetric one-body and inter-mode three-body losses[20]. As a result, the channel effect is suppressed when one-body losses are also present and one can

[17] This can be understood intuitively as follows. For weak losses the random phase $\phi_{1,0}(s) = s\chi$ ($\phi_{0,1}(s) = -s\chi$) produced by the loss of one atom in the mode $i = 1$ ($i = 2$) is uniformly distributed in $[0, \pi/q]$ ($[-\pi/q, 0]$). Since the components of the superposition have a phase separation of $2\pi/q$, one needs equal loss probabilities in the two modes to wash out its phase content completely. A similar argument applies to three-body losses.

[18] A similar cancellation occurs at time t_3 when the two modes are subject to three-body losses with symmetric rates $\Gamma_{3,0} = \Gamma_{0,3}$ and $\Gamma_{1,2} = \Gamma_{2,1} = 0$.

[19] Note that this is not in contradiction with the fact that the total state $\rho(t_2)$ converges to $\rho^{(0)}(t_2)$ when $\Gamma_m \to 0$, since the probability $w_{N_1}(t_2)$ converges to zero in this limit and thus $\rho_{N_1}(t_2)$ does not contribute to the total state.

[20] For instance, one has $\mathcal{E}_{q,1}(n, n') = 2\Gamma_{1,0}\,\mathrm{sinc}[\pi(n - n')/q]$ for $\Delta\Gamma_1 = 0$.

have more than one loss event between $t = 0$ and $t = t_q$. A surprising consequence is that by adding one-body losses to the two-body (three-body) losses, decoherence on the conditional state with $N_0 - 2$ ($N_0 - 3$) atoms at time t_2 (t_3) is reduced.

(iii) *Protecting macroscopic superpositions by tuning the interaction energies.*

Let us now proceed to the case of *asymmetric interaction energies* $U_1 \neq U_2$. In order to keep the formation time $t_q = \pi/(\chi q)$ of the superposition constant, we vary U_1 and U_2 while fixing $2\chi = \chi_1 - \chi_2$. We still consider weak losses satisfying (3.47). Then the phase relaxation described above is incomplete, as well as decoherence at times t_2 or t_3 for symmetric losses. An interesting situation is $U_2 = U_{12} < U_1$, i.e., $\chi_2 = 0$ and $\chi_1 = 2\chi$. Then $\phi_{0,r}(s) = 0$ by (3.33), thus the second mode is protected against phase noise, whereas the first mode is subject to a strong noise with fluctuations $\delta\phi_{r,0} \approx 2\pi r/q$. Taking for simplicity vanishing inter-mode rates, one gets from (3.40) and (3.48)

$$\mathcal{E}_{q,r}(n,n') = \Gamma_{0,r} + q\Gamma_{r,0}\frac{1-\exp\{-\mathrm{i}\frac{2\pi r}{q}(n-n')\}}{2\mathrm{i}\pi r(n-n')} \ . \tag{3.51}$$

For symmetric rates $\Gamma_{0,r} = \Gamma_{r,0}$, the off-diagonal matrix elements of $\rho_{N_1}(t_2)$ in the Fock basis coincide with those of a two-component superposition up to a factor of the order of one half[21]. Loosely speaking, $\rho_{N_1}(t_2)$ is a "half macroscopic superposition". Such a state has a relatively large Fisher information, as shown in Fig. 3.7(a). An even larger Fisher information is obtained for completely asymmetric losses, namely, if atoms are lost in the protected mode $i = 2$ only ($\Gamma_{r,0} = 0$). Then $\mathcal{E}_{q,r}(n,n') = \Gamma_{0,r}$ for all (n,n') and the conditional state $\rho_{N_1}(t_q)$ coincides with a superposition of q coherent states with N_1 atoms, slightly modified by the damping factor $\mathcal{D}_q(n_1,n_1')$. This is in agreement with Fig 3.7(b), where one observes the convergence at weak losses and asymmetric energies of $\mathcal{F}_{N_0-2}(t_2)$ to the highest possible value $(N_0-2)^2$. Two well-pronounced peaks at $\phi = \pm\pi/2$ are visible in the corresponding Husimi distributions.

In summary, by tuning the atomic interaction energies U_i such that $\chi_1 = 0$ or $\chi_2 = 0$ one can protect one mode against phase noise, to the expense of enlarging noise in the other mode, thereby limiting decoherence effects on the conditional state with N_1 atoms. One can similarly switch the phase noise off in the two-body inter-mode loss channel $m = (1,1)$ by tuning the energies such that $\chi_2 = -\chi_1$ (symmetric energies $U_1 = U_2$) and in the three-body inter-mode loss channels $m = (2,1)$ or $(1,2)$ by taking $\chi_2 = -2\chi_1$ or $\chi_1 = -2\chi_2$ ($U_1 = U_2 \mp 2\chi/3$). We emphasize that it is impossible to suppress the noise in two different loss channels at the same time. Therefore, the optimal energy tuning for protecting the macroscopic superpositions is to switch phase noise off in the channel losing more atoms.

Let us remark that, although the result above applies in principle to both internal and external BJJs, for the latter $U_{12} = 0$ and thus one must tune the interaction energy U_i to zero to switch phase noise off in the well i. But the two- and three-body loss rates Γ_m depend on U_i and this tuning actually decreases these rates (recall that two- and three-body losses are due to the scattering of atoms in the trap). Therefore, excepted for one-body losses caused by scattering with the background gas, the protection of the superposition is a trivial effect in external BJJs.

We now claim that when a large number $N_0 - N$ of atoms leave the BJJ (i.e., many loss events occur between $t = 0$ and $t = t_q$), this number being much smaller than N_0, one may still protect the superpositions from decoherence by tuning the energies U_i. However, in order to protect the superpositions efficiently in this way, the loss rates must be strongly asymmetric. This can be understood from the product form (3.41) of the envelope in (3.38): if the single jump envelopes (3.51) are constant in (n,n') then so is $\mathcal{E}_N(t_q;n,n')$, and thus the conditional state $\rho_N(t_q)$ is close to a macroscopic superposition with N atoms, apart from small damping effects. In order that the single jump envelope be constant, one must tune the interaction energies such that $U_i = U_{12}$ in the mode i losing the atoms. As soon as the number of lost atoms becomes large, this tuning is inefficient since the probability that all losses occur in the same mode decreases exponentially with the number of loss events.

Thus, the suppression of the effective phase noise in the mode losing atoms by the energy tuning for strongly asymmetric loss rates provides a good explanation of the time evolution of the quantum Fisher information seen in Fig. 3.5.

[21] More precisely, for all $n \neq n'$, $\mathcal{E}_{q,r}(n,n) = 2\Gamma_{r,0}$ and $\mathcal{E}_{2,r}(n,n')/\mathcal{E}_{2,r}(n,n) = 1/2$ for $r = 2$ and $1/2 + (1-(-1)^{n-n'})/(2\mathrm{i}\pi r(n-n'))$ for $r = 1,3$.

3.5.3 Intermediate and strong loss rates

To complete the interpretation of Fig. 3.7, we briefly discuss three effects of atom losses occurring at intermediate and strong loss rates.

(iv) *Increasing the loss rates reduces phase noise.*

Surprisingly, in the regime of intermediate loss rates the ϕ-noise *decreases* if one increases Γ_m. This results from the decrease of the loss time fluctuations δs_m at increasing Γ_m, leading to a decrease of the phase fluctuations $\delta\phi_m$ in (3.37). In fact, while for small rates the loss time is uniformly distributed on the interval $[0, t_q]$, for larger rates the loss has more chance to occur at small times and δs_m gets smaller. It can be shown [228] that this decreasing of the phase noise sets in for $\Gamma_m \approx \chi q N_0^{1-|m|}$. Therefore, by increasing the loss rates one may protect the conditional state with N_1 atoms against decoherence[22].

(v) *Effect of the θ-noise.*

The fact that the peaks of the Husimi distributions in Fig. 3.7(b) at intermediate and strong losses are centered at values of θ smaller than $\pi/2$ is due to the θ-noise. For larger initial atom numbers N_0, the θ-noise is always small, its fluctuations being of the order of $1/N_0$ when $N_0 \gg 1$ (see [228]).

(vi) *Damping effects.*

Increasing further the rates Γ_m, the damping due to the effective Hamiltonian begins to play the major role. The combination of this damping with the reduced phase noise effect described above leads again to different behaviors of $\mathcal{F}_{N_1}(t_q)$ as a function of the loss rates for symmetric and asymmetric losses. Let us recall from Sec. 3.3.2 that the onset of damping is for $\Gamma_m \approx \chi q N_0^{1-|m|}$ in the symmetric case and $\Gamma_m \approx \chi q N_0^{1/2-|m|}$ in the asymmetric case. Since phase noise reduction begins when $\Gamma_m \approx \chi q N_0^{1-|m|}$, there exists for symmetric losses a small range of rates Γ_m for which phase noise is reduced by increasing Γ_m while the damping is still relatively small. This explains the increase of $\mathcal{F}_{N_1}(t_2)$ with $\Gamma_{2,0}$ seen in Fig. 3.7(a). At the point where $\mathcal{F}_{N_1}(t_2)$ reaches a maximum, two peaks are clearly visible in the Husimi distribution, as opposed to the flat distribution observed at weak losses for $\chi_1 = -\chi_2$. This nicely illustrates phase noise reduction. In contrast, in the asymmetric loss case, damping effects counter-balance phase noise reduction and $\mathcal{F}_{N_1}(t_2)$ decreases with $\Gamma_{2,0}$ (even though some peaks show up in the Husimi plots). For $\Gamma_{2,0} = \Gamma_{0,2} \gg \chi$ and even initial atom numbers N_0, it can be argued [228] that $\rho_{N_0-2}(t_q)$ converges to a Fock state with $(N_0-2)/2$ atoms in each mode, which has a high Fisher information $(N_0-2)N_0/2$. For asymmetric losses, instead, $\rho_{N_0-2}(t_q)$ converges to a superposition of Fock states with $n_1 = 0$ or 1 atom in the first mode, and $\mathcal{F}_{N_0-2}(t_2) \approx N_0$, as seen in Fig. 3.7(b)[23]. Let us stress that these effects on the conditional state with N_1 atoms at strong losses are absent in the total density matrix $\rho(t)$, because of the small probability to have a single loss event between $t = 0$ and $t = t_q$.

3.6 Conclusions

We have shown that the impact of decoherence on superpositions of coherent states in Bose-Josephson junctions subject to one-, two- and three-body atom loss processes is well described by a strong effective phase noise and a channel effect. The last effect gives rise to enhanced decoherence on the two-component (three-component) superposition after summing over the two loss channels when the interaction energies and the two-body (three-body) loss rates are the same in the two modes and there are no inter-mode losses. Conversely, if all losses occur mostly in one mode, it is possible to protect the superpositions from decoherence by adjusting the interaction energies U_i of each mode, keeping their sum fixed. For instance, in the absence of inter-mode losses, the effective phase noise can be suppressed in the mode loosing more atoms by choosing an energy U_i in this mode equal to the inter-mode interaction U_{12}. For the internal BJJs with Rubidium atoms of the Heidelberg experiment [106], this can be done by applying a uniform magnetic field and reducing the scattering length ℓ_1 in the mode $i = 1$ loosing less atoms thanks to Feshbach resonances. Then, because the scattering lengths ℓ_2 and ℓ_{12} are almost equal, one has $U_2 \simeq U_{12}$, whereas $|U_1 - U_{12}|$ can be large. For experimentally relevant loss rates and

[22] This counter-intuitive effect does not manifest itself in the Fisher information $\mathcal{F}_{\rm tot}(t)$ of the total state $\rho(t)$ because when increasing $\Gamma_{2,0}$, the subspaces contributing to $\mathcal{F}_{\rm tot}(t)$ in (3.44) have less atoms and hence are less quantum correlated, and the increase of $\mathcal{F}_{N_0-2}(t)$ is counter-balanced by the decrease of the probability $w_{N_0-2}(t)$. As a consequence, $\mathcal{F}_{\rm tot}(t)$ is getting smaller.

[23] Similarly, the comparison of the two first rows in Fig. 3.6 shows that an increase of $\Gamma_{2,0}$ makes non-vanishing off-diagonal matrix elements to appear when $\Gamma_{2,0} = \Gamma_{0,2}$, as a consequence of phase noise reduction, while for $\Gamma_{0,2} = 0$ the same operation moves the peak in the density matrix towards $n_1 = n_1' = 0$, as a consequence of damping.

initial atom numbers, we have found that the amount of coherence left at the time of formation of the two-component superposition can be made in this way substantially higher, provided that the losses are strongly asymmetric (see Fig. 3.5). This condition is met experimentally for two-body losses but not for one-body losses, which are symmetric in the two modes. As a consequence, in the range of parameters corresponding to the experimental situation studied above, we predict that one-body losses lead to much stronger decoherence effects on macroscopic superpositions than the asymmetric two-body processes.

In the Basel experiment [202], the atoms are trapped by using an inhomogeneous magnetic field and thus a tuning of the scattering lengths with Feshbach resonances is impossible. One must then imagine other ways to reduce phase noise in the upper hyperfine level $i = 2$ suffering from two-body losses. In this experiment, the two BECs of atoms in hyperfine states $i = 1, 2$ can be split by trapping them in different level-dependent potentials. This splitting provokes a decrease of the inter-mode interaction energy U_{12}. An interesting open problem is to find the best experimentally realizable strategy to protect in this case the macroscopic superpositions from decoherence due to losses of atoms.

Part II

Quantum correlations and Distinguishability of quantum states

Introduction

The fundamental role played by the theory of information in physics has been demonstrated in the last century along with the development of statistical physics [22]. More recently, it has been recognized that information is also at the heart of quantum physics, leading to the emergence of a new field called quantum information. In few words, quantum information theory is concerned with the use of quantum systems to accomplish information-processing tasks which are either not feasible classically or are done classically much less efficiently [180]. These tasks can be related to a computational problem or to communication, for instance, sending encrypted information in a secure way. Computational tasks are performed on a quantum computer made of qubits. Such qubits are two-level quantum systems in arbitrary superpositions of $|0\rangle$ and $|1\rangle$ instead of being either in state 0 or 1 as with classical bits. A quantum algorithm is a unitary quantum evolution on a set of qubits followed by a measurement, the outcomes of which should provide the solution of the problem. For example, the celebrated Shor algorithm factorizes an integer with N digits into prime numbers in a time $\mathcal{O}(N^2 \ln N \ln(\ln N))$ [217], instead of the exponential time required by all known classical algorithms. Quantum computers with a few qubits have been implemented in physics laboratories. There is still a lot of debate about whether we will be able in the future to manipulate coherently many qubits and address them locally during a sufficiently long computational time, and which quantum systems are the most promising [180, 41].

The fact that quantum algorithms and communication protocols can outperform their classical analogs is usually attributed to quantum correlations. Such correlations in composite quantum systems are at the origin of the violation of the Bell inequalities, which has been confirmed experimentally [194]. These quantum correlations are quite different in nature from classical correlations in stochastic processes. For a long time they have been identified with entanglement. However, there is now increasing evidence that other types of quantum correlations in mixed states, which may be present even in unentangled states and are captured notably by the quantum discord [184, 120], might be of relevance in certain quantum algorithms and communication protocols [72, 152, 190, 163, 54, 108, 68].

In this second part, we review the basic properties of the entanglement measures and quantum discord and present a geometrical description of these notions based on the Bures distance on the set of quantum states. In this approach, the quantum discord turns out to be related to the problem of discriminating non-orthogonal quantum states. Two central questions guide the foregoing discussions and can be formulated as follows. How well can one distinguish unknown quantum states pertaining to a given ensemble by performing a measurement on a system? If this system consists of several particles, does the amount of information one gets from measurements on a single particle tell us something about the way the particles are correlated? Quantum measurements and entropies obviously come into the game in these two questions. They constitute the subjects of chapters 5 and 7. Some answers to the first question are given in chapters 6 and 9, devoted respectively to state discrimination and to related topics called hypothesis testing and parameter estimation. The Bures distance and Uhlmann fidelity are introduced in chapter 8. A detailed account of their properties is given there. The remaining chapters (chapters 10, 11, and 12) address the problem of quantifying quantum correlations and provide answers to the second question. It is neither our purpose to discuss thoroughly the (huge amount of) quantum correlation measures found in the literature nor to study how these correlations could explain the quantum efficiencies. Well-documented surveys on quantum entanglement already exist, see e.g. [130, 110], as well as on the quantum discord and related measures [170]. The precise role of entanglement as a resource in quantum computing and quantum communication is still not completely understood, in spite of recent progresses (such as the proof that, in order to offer an exponential speedup over classical algorithms, a quantum algorithm using pure states must produce entanglement which is not restricted to blocks of qubits of fixed size as the system size increases [145]). The role played by the discord as a quantum resource is, in turn, still poorly understood and constitutes a challenging issue (see [170]).

We concentrate in our exposition on the mathematical and fundamental aspects of the theory. In particular, we will not investigate here the physical implementations and the system-dependent irreversible dynamical processes destroying (or sometimes producing) quantum correlations, which have been treated in the first part. We present the detailed proofs of some selected fundamental results, instead of relating all important achievements obtained so far. Most of these results have been published in physics journals, and are sometimes explained in the original papers without full mathematical rigor in their derivation. Others have been published in mathematical journals with full proofs, which are nevertheless given here for completeness. We try to emphasize how the results are connected between themselves and to stress the similarities in the arguments used to derive them. This sometimes leads to new proofs.

Quantum information is a rapidly growing field of research and the amount of articles and surveys devoted to it is already considerable. Researchers who got interested by this subject recently (including the author) may

fear to have difficulties to form a clear opinion about the most pertinent open questions. Significant contributions have been made by physicists, mathematicians, and computer scientists, who constitute a broad community with different viewpoints. I hope that this review-like part of this thesis may be useful to mathematicians, by providing examples of interesting problems and explaining the mathematical tools used to tackle them. It may also be of help to physicists wishing to get acquainted with such tools, which could be useful to derive new results. This work is intended to be complementary to other surveys containing collections of results without explicit derivations and to more introductory monographs like [180], which do not include the most recent advances.

Some comments on the structure of the following chapters might be helpful. The contents of chapter 6, chapter 8, and chapters 10-11 are largely independent. On the other hand, chapter 6 is partly related to Sec. 5.5, and chapter 9 makes use of the results of chapter 6 and Sec. 8.5. The material of Secs. 7.1 and 7.2 is relevant for chapters 10 and 11. Chapter 12 needs more or less the knowledge of all previous chapters. The main definitions and theorems presented in chapters 4 and 5 are used in the whole part. Two appendices contain textbook issues about operator convex functions and some less standard trace inequalities.

Before closing this introduction, let us warn the reader that we will be exclusively concerned by quantum systems with *finite-dimensional Hilbert spaces*. This is motivated for two reasons. Firstly, this is the case of most systems in quantum information theory. Secondly, in this way one avoids the technical complications of infinite-dimensional spaces and concentrates oneself on the main ideas and concepts. Some of these concepts have been originally worked out in the general setting of C^*-algebras, but we shall present here simpler proofs applying to the finite-dimensional case only.

Chapter 4

Quantum states

Wenn zwei getrennte Körper, die einzeln maximal bekannt sind, in eine Situation kommen, in der sie aufeinande einwirken, und sich wieder trennen, dann kommt regelmäßig das zustande, was ich eben Verschränkung unseres Wissens um die beiden Körper nannte.[1] (E. Schrödinger, 1935) [211]

In this chapter we review the basic definitions of pure and mixed states, entangled states, and the pure state decompositions and purifications of mixed states. Before that, we introduce in Sec. 4.1 some notation and define a few mathematical objects from the theory of operator algebras, which will be used repeatedly in this article. In Sec. 4.2 we discuss an extremely useful result from linear algebra, namely, the Schmidt decomposition.

In all what follows, capital letters A, B, etc., refer to quantum systems, $\mathcal{H}_\mathsf{A}$, $\mathcal{H}_\mathsf{B}$, etc., denote their Hilbert spaces, and $n_\mathsf{A} = \dim \mathcal{H}_\mathsf{A}$, $n_\mathsf{B} = \dim \mathcal{H}_\mathsf{B}$, etc., the dimensions of these spaces. These dimensions are always assumed to be finite. A bipartite system AB formed by putting together the systems A and B has Hilbert space given by the tensor product $\mathcal{H}_\mathsf{AB} = \mathcal{H}_\mathsf{A} \otimes \mathcal{H}_\mathsf{B}$. For instance, if A and B are two qubits with Hilbert spaces $\mathcal{H}_\mathsf{A} \simeq \mathcal{H}_\mathsf{B} \simeq \mathbb{C}^2$, the space of these two qubits is $\mathcal{H}_\mathsf{AB} = \mathbb{C}^2 \otimes \mathbb{C}^2 \simeq \mathbb{C}^4$. Similarly, $\mathcal{H}_{\mathsf{A}_1 \ldots \mathsf{A}_k} = \mathcal{H}_{\mathsf{A}_1} \otimes \cdots \otimes \mathcal{H}_{\mathsf{A}_k}$ is the Hilbert space of the multipartite system formed by putting together the systems $\mathsf{A}_1, \ldots, \mathsf{A}_k$. The tensor product vectors $|\psi_\mathsf{A}\rangle \otimes |\phi_\mathsf{B}\rangle \in \mathcal{H}_\mathsf{AB}$ will be denoted either by $|\psi_\mathsf{A} \otimes \phi_\mathsf{B}\rangle$ or, more often[2], by $|\psi_\mathsf{A}\rangle|\phi_\mathsf{B}\rangle$.

4.1 Quantum states and observables

A *state* of a quantum system with Hilbert space $\mathcal{H}$ is given by a density matrix ρ, that is, a non-negative operator on $\mathcal{H}$ with unit trace $\operatorname{tr} \rho = 1$. We write $\mathcal{E}(\mathcal{H})$ the convex cone formed by all states on $\mathcal{H}$. States will always be denoted by the letters ρ, σ, or τ, with subscripts referring to the corresponding system if necessary. The extreme points of the cone $\mathcal{E}(\mathcal{H})$ are the *pure states* $\rho_\psi = |\psi\rangle\langle\psi|$, with $|\psi\rangle \in \mathcal{H}$, $\|\psi\| = 1$ (here $|\psi\rangle\langle\psi|$ designates the rank-one orthogonal projector onto $\mathbb{C}|\psi\rangle$). The pure states can be identified with elements of the projective space $P\mathcal{H}$, that is, the set of equivalence classes of normalized vectors in $\mathcal{H}$ modulo a phase factor. The vectors $\mathrm{e}^{\mathrm{i}\theta}|\psi\rangle \in \mathcal{H}$ with $0 \leq \theta < 2\pi$ are called the representatives of $\rho_\psi \in P\mathcal{H}$. We will abusively write $|\psi\rangle$ instead of ρ_ψ, except when this may be a source of confusion. If ρ is a state of a bipartite system AB with Hilbert space $\mathcal{H}_\mathsf{AB} = \mathcal{H}_\mathsf{A} \otimes \mathcal{H}_\mathsf{B}$, the *reduced states* of A and B are defined by partial tracing ρ over the other subsystem. They are denoted by $\rho_A = \operatorname{tr}_B(\rho) \in \mathcal{E}(\mathcal{H}_\mathsf{A})$ and $\rho_B = \operatorname{tr}_A(\rho) \in \mathcal{E}(\mathcal{H}_\mathsf{B})$. These reduced states correspond to the marginals of a joint probability in classical probability theory.

The C^*-algebra of bounded linear operators from $\mathcal{H}$ to $\mathcal{H}'$ is denoted by $\mathcal{B}(\mathcal{H}, \mathcal{H}')$, and we write $\mathcal{B}(\mathcal{H}) = \mathcal{B}(\mathcal{H}, \mathcal{H})$. In our finite-dimensional setting, $\mathcal{B}(\mathcal{H}, \mathcal{H}')$ is the algebra of all $n' \times n$ finite complex matrices, with $\dim \mathcal{H} = n$ and $\dim \mathcal{H}' = n'$. The Hilbert-Schmidt scalar product on $\mathcal{B}(\mathcal{H}, \mathcal{H}')$ is defined by

$$\langle X \, , \, Y \rangle = \operatorname{tr}(X^* Y) \, , \tag{4.1}$$

where X^* denotes the adjoint operator of X. The associated norm is $\|X\|_2 = [\operatorname{tr}(X^* X)]^{\frac{1}{2}}$. The set of states $\mathcal{E}(\mathcal{H})$ can be endowed with the distances[3]

$$d_p(\rho, \sigma) = \|\rho - \sigma\|_p = \left[\operatorname{tr}(|\rho - \sigma|^p)\right]^{\frac{1}{p}} \tag{4.2}$$

[1] "When two separated bodies, which are maximally known individually, come into a situation in which they interact with each other and are then again separated, a state that I just called entanglement of our knowledge on both bodies regularly arises."

[2] As common in the physics literature we do not write the tensor product symbol $\otimes$ explicitly.

[3] We shall see in chapter 8 that there are other more natural distances on $\mathcal{E}(\mathcal{H})$ from a quantum information point of view.

with $p \geq 1$. Here $|X|$ denotes the non-negative operator $|X| = \sqrt{X^*X}$. When $p \to \infty$, $\|X\|_p$ converges to the operator norm $\|X\|_\infty = \|X\|$ of X, that is, the maximal eigenvalue of $|X|$. The Hölder inequality reads

$$\|X\|_p = \max_{Y, \|Y\|_q = 1} |\operatorname{tr}(XY)| \tag{4.3}$$

with $p > 1$ and $q = p/(p-1)$. This still holds for $p = 1$ and $q = \infty$, as can be shown by using the Cauchy-Schwarz inequality for the scalar product (4.1). In that case the maximum is achieved if and only if $YU|X|^{\frac{1}{2}} = e^{\mathrm{i}\theta}|X|^{\frac{1}{2}}$ with $\theta \in [0, 2\pi)$ and U a unitary such that $X = U|X|$ (polar decomposition).

A self-adjoint operator $O \in \mathcal{B}(\mathcal{H})$ is called an *observable*. The real vector space of all observables on $\mathcal{H}$ is denoted by $\mathcal{B}(\mathcal{H})_{\text{s.a.}}$. If AB is a bipartite system, one says that $O \in \mathcal{B}(\mathcal{H}_{\mathsf{AB}})_{\text{s.a}}$ is a *local observable* if either $O = A \otimes 1$ or $O = 1 \otimes B$, with $A \in \mathcal{B}(\mathcal{H}_{\mathsf{A}})_{\text{s.a.}}$ and $B \in \mathcal{B}(\mathcal{H}_{\mathsf{B}})_{\text{s.a.}}$. Here and in the following, 1 stands for the identity operator on $\mathcal{H}_{\mathsf{A}}$, $\mathcal{H}_{\mathsf{B}}$, or another space.

A linear map[4] $\mathcal{M} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ is positive if it transforms a non-negative operator into a non-negative operator. It is completely positive (CP) if the map

$$\mathcal{M} \otimes 1 : X \in \mathcal{B}(\mathcal{H} \otimes \mathbb{C}^m) \;\mapsto\; \sum_{k,l=1}^{m} \mathcal{M}(X_{kl}) \otimes |k\rangle\langle l| \in \mathcal{B}(\mathcal{H}' \otimes \mathbb{C}^m) \tag{4.4}$$

is positive for any integer $m \geq 1$.

Given two orthonormal bases $\{|i\rangle\}_{i=1}^{n_{\mathsf{A}}}$ of $\mathcal{H}_{\mathsf{A}}$ and $\{|j\rangle\}_{j=1}^{n_{\mathsf{B}}}$ of $\mathcal{H}_{\mathsf{B}}$, one can identify any operator $O : \mathcal{H}_{\mathsf{B}} \to \mathcal{H}_{\mathsf{A}}$ with a vector $|\widetilde{\Psi}_O\rangle \in \mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}}$ thanks to the bijection

$$O \mapsto |\widetilde{\Psi}_O\rangle = \sum_{i,j} \langle i|O|j\rangle |i\rangle|j\rangle \; . \tag{4.5}$$

This bijection is an isomorphism between the Hilbert spaces $\mathcal{B}(\mathcal{H}_{\mathsf{B}}, \mathcal{H}_{\mathsf{A}})$ (endowed with the scalar product (4.1)) and $\mathcal{H}_{\mathsf{AB}}$. Similarly, one can represent the linear map $\mathcal{M} : \mathcal{B}(\mathcal{H}_{\mathsf{B}}) \to \mathcal{B}(\mathcal{H}_{\mathsf{A}})$ by an operator $O_{\mathcal{M}}$ acting on $\mathcal{H}_{\mathsf{BB}} = \mathcal{H}_{\mathsf{B}} \otimes \mathcal{H}_{\mathsf{B}}$ with values in $\mathcal{H}_{\mathsf{AA}} = \mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{A}}$. The matrix elements of this operator in the product bases $\{|k\rangle|l\rangle\}_{k,l=1}^{n_{\mathsf{B}}}$ of $\mathcal{H}_{\mathsf{BB}}$ and $\{|i\rangle|j\rangle\}_{i,j=1}^{n_{\mathsf{A}}}$ of $\mathcal{H}_{\mathsf{AA}}$ are given by $(O_{\mathcal{M}})_{ij,kl} = \langle i|\mathcal{M}(|k\rangle\langle l|)|j\rangle$. This representation is an $*$-isomorphism between the C^*-algebras $\mathcal{B}(\mathcal{B}(\mathcal{H}_{\mathsf{B}}), \mathcal{B}(\mathcal{H}_{\mathsf{A}}))$ and $\mathcal{B}(\mathcal{H}_{\mathsf{BB}}, \mathcal{H}_{\mathsf{AA}})$. The so-called reshuffling operation [31] associates to $O_{\mathcal{M}}$ the operator $O_{\mathcal{M}}^{\mathcal{R}} \in \mathcal{B}(\mathcal{H}_{\mathsf{AB}})$ with matrix elements $(O_{\mathcal{M}}^{\mathcal{R}})_{ik,jl} = (O_{\mathcal{M}})_{ij,kl}$, which satisfies

$$\langle A \otimes B \, , \, O_{\mathcal{M}}^{\mathcal{R}} \rangle = \langle \widetilde{\Psi}_A | O_{\mathcal{M}} J | \widetilde{\Psi}_B \rangle = \langle A, \mathcal{M}(\overline{B}) \rangle \tag{4.6}$$

for any $A \in \mathcal{B}(\mathcal{H}_{\mathsf{A}})$ and $B \in \mathcal{B}(\mathcal{H}_{\mathsf{B}})$. Here J denotes the anti-unitary operator on $\mathcal{H}_{\mathsf{BB}}$ defined by $\langle k|\langle l|J|\Psi\rangle = \overline{\langle k|\langle l|\Psi\rangle}$ (complex conjugation in the canonical basis) and $\overline{B} = \sum_{k,l} \overline{\langle k|B|l\rangle} |k\rangle\langle l|$ is the operator associated to $J|\widetilde{\Psi}_B\rangle$ via the isomorphism (4.5). With these definitions, $\mathcal{M} : \mathcal{B}(\mathcal{H}_{\mathsf{B}}) \to \mathcal{B}(\mathcal{H}_{\mathsf{A}})$ is CP if and only if $O_{\mathcal{M}}^{\mathcal{R}} \geq 0$, that is, $O_{\mathcal{M}}^{\mathcal{R}}$ has non-negative eigenvalues[5].

The left and right multiplications $\mathcal{L}_X$ and $\mathcal{R}_X$ by $X \in \mathcal{B}(\mathcal{H})$ are the operators from $\mathcal{B}(\mathcal{H})$ into itself defined by[6]

$$\mathcal{L}_X(Y) = XY \quad , \quad \mathcal{R}_X(Y) = YX \quad , \forall \, Y \in \mathcal{B}(\mathcal{H}) \; . \tag{4.7}$$

They are represented on $\mathcal{B}(\mathcal{H} \otimes \mathcal{H})$ by local operators $X \otimes 1$ and $1 \otimes X^T$, respectively, where T stands for the transposition in the basis $\{|i\rangle\}$. Given two states ρ and $\sigma \in \mathcal{E}(\mathcal{H})$ with $\rho > 0$, the Araki relative modular operator $\Delta_{\sigma|\rho}$ is defined by [13]

$$\Delta_{\sigma|\rho}(Y) = \sigma Y \rho^{-1} = \mathcal{L}_\sigma \circ \mathcal{R}_{\rho^{-1}}(Y) \quad , \forall \, Y \in \mathcal{B}(\mathcal{H}) \; . \tag{4.8}$$

It is a self-adjoint non-negative operator on the Hilbert space $\mathcal{B}(\mathcal{H})$ (for the scalar product (4.1)).

[4] Operators acting on the vector space of observables $\mathcal{B}(\mathcal{H})_{\text{s.a.}}$ or on the whole algebra $\mathcal{B}(\mathcal{H})$ are always denoted by calligraphic letters.

[5] Actually, $O_{\mathcal{M}}^{\mathcal{R}} \geq 0$ is equivalent to $O_{\mathcal{M}}^{\mathcal{R}} = A^*A$ for some $A \in \mathcal{B}(\mathcal{H}_{\mathsf{AB}})$, that is, to $(O_{\mathcal{M}}^{\mathcal{R}})_{ik,jl} = \langle i|\mathcal{M}(|k\rangle\langle l|)|j\rangle = \sum_{p,q} \overline{A_{pq,ik}} A_{pq,jl}$ for all $i, j = 1, \ldots, n_{\mathsf{A}}$ and $k, l = 1, \ldots, n_{\mathsf{B}}$. Setting $A_{pq} = \sum_{i,k} \overline{A_{pq,ik}} |i\rangle\langle k|$, it follows that $O_{\mathcal{M}}^{\mathcal{R}} \geq 0$ if and only if $\mathcal{M}(X) = \sum_{pq} A_{pq} X A_{pq}^*$ for all $X \in \mathcal{B}(\mathcal{H}_{\mathsf{B}})$, which is equivalent to $\mathcal{M}$ being CP by the Kraus representation theorem (Theorem 5.2.3 below).

[6] In the C^*-algebra setting, the map $X \mapsto \mathcal{L}_X$ is the Gelfand-Neumark-Segal representation of the C^*-algebra [42].

4.2 The Schmidt decomposition

The following standard result is very useful in quantum information theory.

Theorem 4.2.1. (Schmidt decomposition) *Any pure state $|\Psi\rangle \in \mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}}$ of a bipartite system admits a decomposition*

$$|\Psi\rangle = \sum_{i=1}^{n} \sqrt{\mu_i} |\alpha_i\rangle |\beta_i\rangle \tag{4.9}$$

where $n = \min\{n_{\mathsf{A}}, n_{\mathsf{B}}\}$, $\mu_i \geq 0$, and $\{|\alpha_i\rangle\}_{i=1}^n$ (respectively $\{|\beta_i\rangle\}_{i=1}^n$) is an orthonormal family of $\mathcal{H}_{\mathsf{A}}$ (respectively of $\mathcal{H}_{\mathsf{B}}$). The μ_i and $|\alpha_i\rangle$ (respectively $|\beta_i\rangle$) are the eigenvalues and eigenvectors of the reduced state $\rho_{\mathsf{A}} = \mathrm{tr}_{\mathsf{B}}(|\Psi\rangle\langle\Psi|)$ (respectively $\rho_{\mathsf{B}} = \mathrm{tr}_{\mathsf{A}}(|\Psi\rangle\langle\Psi|)$). Thus, if the eigenvalues μ_i are non-degenerate then the decomposition (4.9) is unique.

The non-negative numbers μ_i are called the *Schmidt coefficients* of $|\Psi\rangle$. They satisfy $\sum_i \mu_i = \|\Psi\|^2 = 1$.

Proof. Let $\{|i\rangle\}_{i=1}^{n_{\mathsf{A}}}$ and $\{|j\rangle\}_{j=1}^{n_{\mathsf{B}}}$ be some fixed orthonormal bases of $\mathcal{H}_{\mathsf{A}}$ and $\mathcal{H}_{\mathsf{B}}$. By using the isomorphism $|\Psi\rangle \mapsto O_\Psi = \sum_{i,j} \langle i \otimes j|\Psi\rangle |i\rangle\langle j|$ between $\mathcal{H}_{\mathsf{AB}}$ and the space of $n_{\mathsf{A}} \times n_{\mathsf{B}}$ matrices (see Sec. 4.1), we observe that the decomposition (4.9) corresponds to the singular value decomposition of O_Ψ, that is, $O_\Psi = U_{\mathsf{A}} \sum_i \sqrt{\mu_i} |i\rangle\langle i| U_{\mathsf{B}}^*$ with μ_i the eigenvalues of $O_\Psi^* O_\Psi$ and U_{A} and U_{B} unitaries on $\mathcal{H}_{\mathsf{A}}$ and $\mathcal{H}_{\mathsf{B}}$. Then $U_{\mathsf{A}}|i\rangle = |\alpha_i\rangle$ and $U_{\mathsf{B}}|i\rangle = |\beta_i^*\rangle$ are eigenvectors of $O_\Psi O_\Psi^*$ and $O_\Psi^* O_\Psi$, respectively. Denoting by J is the complex conjugation in the basis $\{|j\rangle\}$ (see above), one has $|\beta_i\rangle = J|\beta_i^*\rangle$. □

The Schmidt decomposition can be generalized to mixed states by considering ρ as a vector in the Hilbert space $\mathcal{B}(\mathcal{H}_{\mathsf{A}}) \otimes \mathcal{B}(\mathcal{H}_{\mathsf{B}})$. Any $\rho \in \mathcal{E}(\mathcal{H}_{\mathsf{AB}})$ can be written as

$$\rho = \sum_{m=1}^{n^2} \sqrt{\mu_m} X_m \otimes Y_m \ , \tag{4.10}$$

where $\{X_m\}_{m=1}^{n_{\mathsf{A}}^2}$ and $\{Y_m\}_{m=1}^{n_{\mathsf{B}}^2}$ are orthonormal bases of $\mathcal{B}(\mathcal{H}_{\mathsf{A}})$ and $\mathcal{B}(\mathcal{H}_{\mathsf{B}})$ for the scalar product (4.1) and μ_m are the eigenvalues of the $n_{\mathsf{A}}^2 \times n_{\mathsf{A}}^2$ matrix $R \geq 0$ defined by

$$R_{ij,i'j'} = \Big\langle \rho\, |i\rangle\langle i'| \otimes 1 \ , \ |j\rangle\langle j'| \otimes 1\, \rho \Big\rangle \tag{4.11}$$

(the $R_{ij,i'j'}$ are the matrix elements in the orthonormal basis $\{|i\rangle\langle j|\}_{i,j=1}^{n_{\mathsf{A}}^2}$ of $\mathcal{B}(\mathcal{H}_{\mathsf{A}})$ of the operator playing the role of the reduced state in Theorem 4.2.1). Note that $\sum_m \mu_m = \mathrm{tr}(\rho^2) \leq 1$, with equality if and only if ρ is a pure state.

Remark 4.2.2. *Alternatively, the μ_m are the squares of the singular values of $\rho^{\mathcal{R}} \in \mathcal{B}(\mathcal{H}_{\mathsf{BB}}, \mathcal{H}_{\mathsf{AA}})$, where $\mathcal{R}$ is the reshuffling operation (Sec. 4.1), and X_m and Y_m are given in terms of the eigenvectors $|\chi_m\rangle$ and $|\psi_m\rangle$ of $\rho^{\mathcal{R}}(\rho^{\mathcal{R}})^*$ and $(\rho^{\mathcal{R}})^*\rho^{\mathcal{R}}$ by $X_m = \sum_{i,j}\langle i \otimes j|\chi_m\rangle |i\rangle\langle j|$ and $Y_m = \sum_{k,l} \overline{\langle k \otimes l|\psi_m\rangle} |k\rangle\langle l|$, respectively.*

Proof. Considering ρ as a vector in $\mathcal{B}(\mathcal{H}_{\mathsf{A}}) \otimes \mathcal{B}(\mathcal{H}_{\mathsf{B}})$ and introducing two orthonormal bases $\{A_p\}$ of $\mathcal{B}(\mathcal{H}_{\mathsf{A}})$ and $\{B_q\}$ of $\mathcal{B}(\mathcal{H}_{\mathsf{B}})$, according to the proof of Theorem 4.2.1, $\sqrt{\mu_m}$ are the singular values of the $n_{\mathsf{A}}^2 \times n_{\mathsf{B}}^2$ matrix $(\langle A_p \otimes B_q, \rho\rangle)_{p,q}$. Denote by $\{|\alpha_p\rangle\}$ and $\{|\beta_q\rangle\}$ the orthonormal bases of $\mathcal{H}_{\mathsf{AA}}$ and $\mathcal{H}_{\mathsf{BB}}$ associated to $\{A_p\}$ and $\{B_q\}$ via the isomorphism (4.5). The statement follows by choosing $A_p = |i\rangle\langle j|$ and $B_q = |k\rangle\langle l|$ and using the identity $\langle \alpha_p|\rho^{\mathcal{R}} J|\beta_q\rangle = \langle A_p \otimes B_q, \rho\rangle$, see (4.6). □

4.3 Purifications and pure state decompositions of mixed states

Definition 4.3.1. *Let ρ be an arbitrary state on $\mathcal{H}$ and $\mathcal{K}$ be another Hilbert space. A pure state $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{K}$ such that $\rho = \mathrm{tr}_{\mathcal{K}}(|\Psi\rangle\langle\Psi|)$ is called a purification of ρ on $\mathcal{H} \otimes \mathcal{K}$.*

In the language of C^*-algebras, a purification is an example of cyclic representation of a state [42]. An example of purification of ρ on $\mathcal{H} \otimes \mathcal{H}$ is

$$|\Psi\rangle = \sum_{k=1}^{n} \sqrt{p_k} |k\rangle |k\rangle \ , \tag{4.12}$$

where $\rho = \sum_k p_k |k\rangle\langle k|$ is a spectral decomposition of ρ. If $|\Psi\rangle$ and $|\Phi\rangle$ are two purifications of ρ on the same space $\mathcal{H} \otimes \mathcal{K}$, then there exists a unitary operator U acting on $\mathcal{K}$ such that $|\Phi\rangle = 1 \otimes U|\Psi\rangle$. In fact, one infers from the Schmidt decomposition that $|\Psi\rangle = \sum_k \sqrt{p_k}|k\rangle|f_k\rangle$ and $|\Phi\rangle = \sum_k \sqrt{p_k}|k\rangle|g_k\rangle$, where $\{|f_k\rangle\}_{k=1}^n$ and $\{|g_k\rangle\}_{k=1}^n$ are two orthonormal families of $\mathcal{K}$. Thus $|g_k\rangle = U|f_k\rangle$ for some unitary U.

We will often be interested in the sequel by families of quantum states of a system S, $\rho_i \in \mathcal{E}(\mathcal{H}_\mathsf{S})$, $i = 1, \ldots, m$, to which we attach some probabilities $\eta_i \geq 0$, $\sum_i \eta_i = 1$. Following the terminology employed by physicists in statistical physics, we call $\{\rho_i, \eta_i\}_{i=1}^m$ an *ensemble of quantum states* (or more simply an *ensemble*). A *convex decomposition* of ρ is an ensemble $\{\rho_i, \eta_i\}_{i=1}^m$ such that $\rho = \sum_i \eta_i \rho_i$. A *pure state decomposition* of ρ is a convex decomposition in terms of finitely many pure states $\rho_i = |\psi_i\rangle\langle\psi_i|$, i.e.

$$\rho = \sum_{i=1}^m \eta_i |\psi_i\rangle\langle\psi_i| \, . \tag{4.13}$$

If the vectors $|\psi_i\rangle$ are orthogonal, then (4.13) coincides with the spectral decomposition, but we will see that there are infinitely many other ways to decompose ρ. Physically, (4.13) describes a *state preparation*: it means that the system has been prepared in the pure state $|\psi_i\rangle$ with probability η_i. The non-uniqueness of the decomposition can be interpreted as follows. If a receiver is given two ensembles $\{|\psi_i\rangle, \eta_i\}_{i=1}^m$ and $\{|\phi_j\rangle, \xi_j\}_{j=1}^p$ corresponding to different state preparations of two identical systems in the same state ρ, then he cannot make any difference between them if he has no prior knowledge on the state preparation. Indeed, any measurement performed by him gives rise to the same distribution of outcomes for the two ensembles. In other words, the full information that the receiver can collect on the system via measurements is encoded in ρ, and *not* in the ensemble involved in the state preparation. This very important fact has consequences that are sometimes disconcerting to people unfamiliar with the conceptual aspects of quantum mechanics. For instance, if a preparer gives a maximally mixed state $\rho = 1/n$ to a receiver, the latter has no way to decide whether this state was prepared from n equiprobable orthonormal pure states (which are only known by the preparer) or if it was prepared by another procedure involving more than n states. It is also worth mentioning that the process transforming the ensemble $\{\rho_i, \eta_i\}_{i=1}^m$ into the average state $\rho = \sum_i \eta_i \rho_i$, which can be viewed as the inverse of a convex decomposition, corresponds physically to a loss of information about the state preparation.

Given a fixed orthonormal basis $\{|f_i\rangle\}_{i=1}^p$ of $\mathcal{K}$ with $p \geq \mathrm{ran}(\rho) = r$, there is a one-to-one correspondence between pure state decompositions of ρ containing at most p states and purifications of ρ on $\mathcal{H} \otimes \mathcal{K}$. Actually, given the pure state decomposition (4.13),

$$|\Psi\rangle = \sum_{i=1}^p \sqrt{\eta_i}|\psi_i\rangle|f_i\rangle \tag{4.14}$$

defines a purification of ρ on $\mathcal{H} \otimes \mathcal{K}$ (we have set $\eta_i = 0$ for $m < i \leq p$). Reciprocally, let $|\Psi\rangle$ be a purification of ρ on $\mathcal{H} \otimes \mathcal{K}$. Denote as before the eigenvalues and orthonormal eigenvectors of ρ by p_k and $|k\rangle$. As argued above, one can find a unitary U on $\mathcal{K}$ such that

$$|\Psi\rangle = \sum_{k=1}^r \sqrt{p_k}|k\rangle U|f_k\rangle = \sum_{i=1}^p \sum_{k=1}^r \sqrt{p_k}\langle f_i|U|f_k\rangle|k\rangle|f_i\rangle = \sum_{i=1}^p \sqrt{\eta_i}|\psi_i\rangle|f_i\rangle \tag{4.15}$$

with $\sqrt{\eta_i}|\psi_i\rangle = \sum_k \sqrt{p_k}\langle f_i|U|f_k\rangle|k\rangle$. Hence $|\Psi\rangle$ has the form (4.14). Taking the partial trace over $\mathcal{K}$, one can associate to it a unique pure state decomposition, which is given by (4.13).

Since two purifications $|\Psi\rangle$ and $|\Phi\rangle$ of the same state ρ are related by a local unitary U acting on the ancilla space $\mathcal{K}$, this implies that any two pure state decompositions $\rho = \sum_{i=1}^m \eta_i |\psi_i\rangle\langle\psi_i|$ and $\rho = \sum_{j=1}^p \xi_j |\phi_j\rangle\langle\phi_j|$ are related by

$$\sqrt{\xi_j}|\phi_j\rangle = \sum_{i=1}^{\max\{m,p\}} u_{ji}\sqrt{\eta_i}|\psi_i\rangle \, , \tag{4.16}$$

where (u_{ji}) is a unitary matrix with size $\max\{m, p\}$ (if $m < i \leq p$ we set as before $\eta_i = 0$).

4.4 Entangled and separable states

Let us consider a bipartite system AB. If this system is in a tensor product state $|\Psi_{\mathrm{sep}}\rangle = |\psi_\mathsf{A}\rangle|\phi_\mathsf{B}\rangle$ with $|\psi_\mathsf{A}\rangle \in \mathcal{H}_\mathsf{A}$ and $|\phi_\mathsf{B}\rangle \in \mathcal{H}_\mathsf{B}$, then the expectation value of the product of two local observables $A \otimes 1$ and $1 \otimes B$ coincides with the product of the expectations values, i.e.,

$$G_{AB}(|\Psi_{\mathrm{sep}}\rangle) = \langle\Psi_{\mathrm{sep}}|A \otimes B|\Psi_{\mathrm{sep}}\rangle - \langle\Psi_{\mathrm{sep}}|A \otimes 1|\Psi_{\mathrm{sep}}\rangle\langle\Psi_{\mathrm{sep}}|1 \otimes B|\Psi_{\mathrm{sep}}\rangle = 0 \, . \tag{4.17}$$

This means that the random outcomes of measurements of the local observables $A\otimes 1$ and $1\otimes B$ are uncorrelated. More generally, if one thinks of AB as a pair of particles located far apart (e.g. a photon pair shared by two observers Alice and Bob), this pair is in a product state if and only if there are no correlations between the results of arbitrary local measurements performed independently on each particle (for instance, if Alice sends her photon through a polarizer and then to a photodetector, and Bob does the same with his photon, the clicks of the two detectors will be uncorrelated whatever the polarizer angles). One says that $|\Psi_{\rm sep}\rangle = |\psi_{\mathsf{A}}\rangle|\phi_{\mathsf{B}}\rangle$ is a *separable state*. If the pure state $|\Psi\rangle \in \mathcal{H}_{\mathsf{AB}}$ is not a product state one says that it is *entangled*.

By applying the Schmidt decomposition, one sees that $|\Psi\rangle$ is separable if and only if all its Schmidt coefficients vanish except one, that is, if and only if its reduced states ρ_{A} and ρ_{B} are pure. In the opposite, if either ρ_{A} or ρ_{B} is proportional to the identity matrix (maximally mixed state), we say that $|\Psi\rangle$ is *maximally entangled*. Such states have the form

$$|\Psi_{\rm ent}\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |\alpha_i\rangle|\beta_i\rangle \ , \tag{4.18}$$

where $\{|\alpha_i\rangle\}_{i=1}^n$ and $\{|\beta_i\rangle\}_{i=1}^n$ are orthonormal families in $\mathcal{H}_{\mathsf{A}}$ and $\mathcal{H}_{\mathsf{B}}$ and $n = \min\{n_{\mathsf{A}}, n_{\mathsf{B}}\}$. For instance, denoting by $|0\rangle$ and $|1\rangle$ the canonical basis vectors of $\mathbb{C}^2$, the EPR (or Bell) states $|\Phi_\pm\rangle = (|0\rangle|0\rangle \pm |1\rangle|1\rangle)/\sqrt{2}$ and $|\Psi_\pm\rangle = (|0\rangle|1\rangle \pm |1\rangle|0\rangle)/\sqrt{2}$ are maximally entangled states of two qubits, and any maximally entangled two-qubit state is an EPR state, up to a local unitary transformation $U_{\mathsf{A}} \otimes U_{\mathsf{B}}$.

For mixed states, entanglement is no longer equivalent to being a product state. The "good" definition of mixed state entanglement is due to Werner [257].

Definition 4.4.1. *A mixed state ρ of a bipartite system AB is separable if it admits a pure state decomposition*

$$\rho = \sum_i \eta_i |\psi_i \otimes \phi_i\rangle\langle\psi_i \otimes \phi_i| \tag{4.19}$$

in terms of pure separable states $|\psi_i \otimes \phi_i\rangle \in \mathcal{H}_{\mathsf{AB}}$. If such a decomposition does not exist then ρ is entangled. The set of all separable states of AB forms a convex subset of $\mathcal{E}(\mathcal{H}_{\mathsf{AB}})$, which is denoted by $\mathcal{S}_{\mathsf{AB}}$.

It follows from the Carathéodory theorem that the number of pure product states in the decomposition (4.19) can always be chosen to be smaller or equal to $(n_{\mathsf{A}}n_{\mathsf{B}})^2 + 1$.

According to this definition, a state is separable if it *could* have been prepared from pure product states only. This does not mean that it has actually been prepared using such states. For example, if one prepares two qubits in the maximally entangled states $|\Phi_+\rangle$ and $|\Phi_-\rangle$ with equal probabilities, the corresponding state

$$\rho = \frac{1}{2}|\Phi_+\rangle\langle\Phi_+| + \frac{1}{2}|\Phi_-\rangle\langle\Phi_-| = \frac{1}{2}|0\rangle\langle 0| \otimes |0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| \otimes |1\rangle\langle 1| \tag{4.20}$$

is separable! This unexpected result is inherent to the ambiguity of the state preparation discussed in the preceding section. This quantum ambiguity unfortunately obliges us to look for all possible state preparations of a given mixed state ρ to decide whether ρ is entangled or not. This makes this problem highly non-trivial.

An explicit complete characterization of $\mathcal{S}_{\mathsf{AB}}$ is known for qubit-qubit and qubit-qutrit systems only, that is, for $(n_{\mathsf{A}}, n_{\mathsf{B}}) = (2,2)$, $(2,3)$, and $(3,2)$. In such a case, the Peres-Horodecki criterion [193, 127, 128] gives a necessary and sufficient condition for ρ to be entangled. This criterion is formulated in terms of the partial transpose. Given two orthonormal bases $\{|i\rangle\}$ of $\mathcal{H}_{\mathsf{A}}$ and $\{|k\rangle\}$ of $\mathcal{H}_{\mathsf{B}}$, the partial transpose $\rho^{T_{\mathsf{B}}}$ of ρ with respect to B has matrix elements in the product basis $\{|i\rangle|k\rangle\}$ given by

$$\langle i|\langle k|\rho^{T_{\mathsf{B}}}|j\rangle|l\rangle = \langle i|\langle l|\rho|j\rangle|k\rangle \ . \tag{4.21}$$

One defines similarly $\rho^{T_{\mathsf{A}}}$ and note that $\rho^{T_{\mathsf{A}}} = (\rho^{T_{\mathsf{B}}})^T$. It follows from Definition 4.4.1 that if ρ is separable then $\rho^{T_{\mathsf{A}}} \geq 0$ and $\rho^{T_{\mathsf{B}}} \geq 0$, i.e., $\rho^{T_{\mathsf{A}}}$ and $\rho^{T_{\mathsf{B}}}$ are states of AB. Thus, if $\rho^{T_{\mathsf{A}}}$ (or, equivalently, $\rho^{T_{\mathsf{B}}}$) has negative eigenvalues then ρ is necessarily entangled. Since the transpose is a positive but not CP map, such negative eigenvalues may indeed exist. However, if $n_{\mathsf{A}}n_{\mathsf{B}} > 6$, certain entangled states have non-negative partial transposes [128]. It is remarkable that this does not happen when $n_{\mathsf{A}}n_{\mathsf{B}} \leq 6$: then $\rho^{T_{\mathsf{A}}} \geq 0$ if and only if $\rho \in \mathcal{S}_{\mathsf{AB}}$ [127]. Two remarks should be made at this point. First, states with non-negative partial transposes cannot undergo entanglement distillation and therefore form an interesting subset of $\mathcal{E}(\mathcal{H}_{\mathsf{AB}})$ on their own, which contains $\mathcal{S}_{\mathsf{AB}}$ (see [130] for more detail). Second, extending the Peres criterion to all positive but not CP linear maps $\Lambda_{\mathsf{B}} : \mathcal{B}(\mathcal{H}_{\mathsf{B}}) \to \mathcal{B}(\mathcal{H}_{\mathsf{A}})$ (i.e., asking that $1 \otimes \Lambda_{\mathsf{B}}(\rho) \geq 0$ for any such map) yields a necessary and sufficient condition for entanglement,

valid whatever the space dimensions n_{A} and n_{B} [127]. Due to the lack of an explicit characterization of such maps (except for $(n_{\mathsf{A}}, n_{\mathsf{B}}) = (2,2)$ or $(3,2)$)[7], this condition is unfortunately not very helpful in general.

Let us also mention another necessary but not sufficient (even for two qubits) condition for entanglement, which relies on the Schmidt decomposition (4.10) for mixed states. By using the fact that $\sum_m \sqrt{\mu_m}$ defines a norm on $\mathcal{E}(\mathcal{H}_{\mathsf{AB}})$, one can show that if $\rho \in \mathcal{S}_{\mathsf{AB}}$ then $\sum_m \sqrt{\mu_m} \leq 1$ [59]. Hence $\sum_m \sqrt{\mu_m} > 1$ implies that ρ is entangled.

Once a state has been recognized as separable, it may be of relevance to determine its decomposition(s) into pure product states. This problem has been addressed in [265, 206, 252] for two qubits.

Definition 4.4.1 can be extended straightforwardly to multipartite systems $\mathsf{A}_1 \dots \mathsf{A}_k$. Then different kinds of entanglement can be defined according to the chosen partition of $\{\mathsf{A}_1, \dots, \mathsf{A}_k\}$. In this article we will not consider multipartite entanglement, which is a challenging subject in its own [110, 130].

[7] When $(n_{\mathsf{A}}, n_{\mathsf{B}}) = (2,2)$ or $(3,2)$, any positive map $\Lambda : \mathcal{B}(\mathcal{H}_{\mathsf{B}}) \to \mathcal{B}(\mathcal{H}_{\mathsf{A}})$ can be written as $\Lambda = \mathcal{M}_1 + \mathcal{M}_2 \circ T$, where $\mathcal{M}_1$ and $\mathcal{M}_2$ are CP and T is the transposition [266]. The fact that the partial transpose criterion is sufficient for entanglement follows from this characterization [127].

Chapter 5

Generalized measurements

Wenn wir experimentell feststellen, daß das Atom eben in den Zustand m wirklich übergegangen sei, so werden wir ihm zur Berechnung alles Folgenden nicht die Funktion $\sum_m c_{nm} S_m$, sondern eben die Funktion S_m mit unbestimmter Phase zuzuordnen haben.[1] (W. Heisenberg, 1927) [117]

In this chapter we review the notions of quantum operations and generalized measurements and give the basic theorems, namely, the Stinespring theorem, the Kraus decomposition, and the Neumark extension theorem. We then introduce a special type of measurement called the least square measurement and show its link with approximate reversals of quantum operations. We start by a physical description of a von Neumann measurement.

5.1 Physical realization of a measurement process

A measurement on a quantum system S is realized by coupling S with a measurement apparatus. This apparatus consists of a macroscopic pointer P interacting with an environment E playing the role of an infinite bath. One may think of P as the center of mass of the needle of a meter. The environment E then includes all the other degrees of freedom of the macroscopic apparatus. The coupling of the measured system S with the pointer transforms the initially uncorrelated state $|\psi\rangle \otimes |0\rangle$ of the composite system SP into an entangled state,

$$|\psi\rangle \otimes |0\rangle \longrightarrow |\Psi_{\mathsf{SP}}^{\rm ent}\rangle = \sum_{i,l} c_{il} |\alpha_{il}\rangle \otimes |i\rangle \ . \tag{5.1}$$

Our assumption that S and P are initially in pure states is made to simplify the foregoing discussion and can be easily relaxed. The states $|\alpha_{il}\rangle$ form an orthonormal basis of the system Hilbert space $\mathcal{H}_{\mathsf{S}}$ (measurement basis), which is the eigenbasis of the measured observable A, i.e., $A|\alpha_{il}\rangle = a_i|\alpha_{il}\rangle$. The index l labels if necessary the different orthogonal eigenstates of A with the same degenerate eigenvalue a_i. In ideal measurements $c_{il} = \langle\alpha_{il}|\psi\rangle$. The states $|i\rangle$ are the pointer states of the apparatus. After a sufficiently long coupling time between S and P, these states are macroscopically distinct and thus nearly orthogonal, $\langle i|j\rangle \simeq \delta_{ij}$ (hereafter δ_{ij} stands for the Kronecker symbol, equal to 1 if $i = j$ and zero otherwise). The transformation (5.1) is a unitary transformation, i.e., $|\Psi_{\mathsf{SP}}^{\rm ent}\rangle = U_{\mathsf{SP}}|\psi\rangle|0\rangle$ where U_{SP} is a unitary evolution operator on $\mathcal{H}_{\mathsf{SP}}$. One usually calls such a transformation the *pre-measurement* [102]. This unitary evolution induces quantum correlations between S and P, such that each eigenprojector $\Pi_i = \sum_l |\alpha_{il}\rangle\langle\alpha_{il}|$ of A is in one-to-one correspondence with a pointer state $|i\rangle$. The resulting state (5.1) is a superposition of macroscopically distinct states, sometimes referred to as a "Schrödinger cat state". The pointer states are singled out by their robustness against environment-induced decoherence. More precisely, if the pointer P is initially in the state $|i\rangle$, its interaction with the environment E does not entangle P and E. Letting P and E interact during a time t much larger than the decoherence time, the SP-entangled state $|\Psi_{\mathsf{SP}}^{\rm ent}\rangle$ is transformed into a statistical mixture in which all the coherences between the pointer states $|i\rangle$ have disappeared. After tracing out the environment degrees of freedom, the reduced state of SP is modified according to

$$|\Psi_{\mathsf{SP}}^{\rm ent}\rangle\langle\Psi_{\mathsf{SP}}^{\rm ent}| \longrightarrow \rho_{\mathsf{SP}}^{\rm p.m.} = \sum_{ikl} c_{ik}\overline{c_{il}} |\alpha_{ik}\rangle\langle\alpha_{il}| \otimes |i\rangle\langle i| = \sum_i \Pi_i \, \rho \, \Pi_i \otimes |i\rangle\langle i| \ , \tag{5.2}$$

[1] "If we determine experimentally that the atom has actually acquired the state m, then in the subsequent calculations we shall have to assign it the function S_m with an indeterminate phase, instead of $\sum_m c_{nm} S_m$."

$\rho = |\psi\rangle\langle\psi|$ being the initial system state. The final SP-state has no quantum correlations but is classically correlated: indeed, each pointer state $|i\rangle$ goes hand in hand with the system state

$$\rho_{S|i} = p_i^{-1} \Pi_i \, \rho \, \Pi_i \quad , \quad p_i = \mathrm{tr}(\Pi_i \rho) \; . \tag{5.3}$$

Concrete models for the pointer and its coupling with the system and the environment have been investigated in [4, 5, 223, 224]; in these works the aforementioned decoherence time and the time duration of the measurement are estimated in the more realistic situation where the two transformations (5.1) and (5.2) occur simultaneously. The readout of the pointer (that is, the observation of the position of the needle) cannot significantly alter the macroscopic state $|i\rangle$. It merely selects one of the measurement outcomes,

$$\textbf{outcome } i\text{:} \qquad \rho_{\mathsf{SP}}^{\text{p.m.}} \; \longrightarrow \; \rho_{\mathsf{SP}|i} = \rho_{\mathsf{S}|i} \otimes |i\rangle\langle i| \qquad \text{(wavepacket reduction)}. \tag{5.4}$$

After the measurement yielding the outcome i the measured system is in the conditional state $\rho_{S|i}$, and this outcome occurs with probability p_i (Born rule). The transformation (5.4) results from the knowledge of the random outcome, it should not be regarded as a true dynamical process. It is actually analog to a state preparation (see Sec. 4.3). In mathematical terms, it corresponds to a convex decomposition of $\rho_{\mathsf{SP}}^{\text{p.m.}} = \sum_i p_i \rho_{\mathsf{SP}|i}$.

We point out that recent progresses in the understanding of quantum measurement processes via dynamical models and their interpretation with a statistical physics viewpoint have been made by Allahverdyan, Balian, and Nieuwenhuizen [6].

5.2 Quantum operations

In the absence of readout of the measurement result, one does not know which state $\rho_{S|i}$ has been prepared and the system is after the measurement in the average state

$$\mathcal{M}_\Pi(\rho) = \sum_i \Pi_i \, \rho \, \Pi_i \; , \tag{5.5}$$

where ρ is the state before the measurement.

Since $\{\Pi_i\}$ is the spectral measure of the self-adjoint operator A, the Π_i form a family of projectors in $\mathcal{B}(\mathcal{H}_\mathsf{S})_{\text{s.a.}}$ satisfying $\Pi_i \Pi_j = \delta_{ij} \Pi_i$ and $\sum_i \Pi_i = 1$. We will refer in the sequel to such a family as an *orthonormal family of projectors*. It is easy to show that the map $\mathcal{M}_\Pi$ is CP (as a sum of CP maps) and trace-preserving. In quantum information, such maps are called *quantum operations*.

Definition 5.2.1. *A quantum operation $\mathcal{M} : \mathcal{B}(\mathcal{H}_\mathsf{S}) \to \mathcal{B}(\mathcal{H}'_\mathsf{S})$ is a trace-preserving CP map from $\mathcal{B}(\mathcal{H}_\mathsf{S})$ into $\mathcal{B}(\mathcal{H}'_\mathsf{S})$.*

A necessary and sufficient condition for a linear map $\mathcal{M} : \mathcal{B}(\mathcal{H}_\mathsf{S}) \to \mathcal{B}(\mathcal{H}'_\mathsf{S})$ to be CP is that it satisfies $\mathcal{M} \otimes 1(|\Psi_{\text{ent}}\rangle\langle\Psi_{\text{ent}}|) \geq 0$ for the maximally entangled state $|\Psi_{\text{ent}}\rangle = n_\mathsf{S}^{-1/2} \sum_k |k\rangle|k\rangle$ in $\mathcal{H}_\mathsf{S} \otimes \mathcal{H}_\mathsf{S}$, where $\{|k\rangle\}$ is an orthonormal basis of $\mathcal{H}_\mathsf{S}$. In fact, $\mathcal{M} \otimes 1(|\Psi_{\text{ent}}\rangle\langle\Psi_{\text{ent}}|)$ coincides with the operator $O^{\mathcal{R}}_{\mathcal{M}}$ defined in Sec. 4.1 up to a factor $1/n_\mathsf{S}$, and it has been argued above that $\mathcal{M}$ is CP if and only if $O^{\mathcal{R}}_{\mathcal{M}} \geq 0$.

A quantum operation is the quantum analog of a stochastic matrix $\mathcal{M}^{\text{clas}}$ giving the transition probabilities $q(j|i)$ of a classical Markov process,

$$\mathbf{p} = (p_1, \ldots, p_n) \; \mapsto \; \mathcal{M}^{\text{clas}} \mathbf{p} \quad \text{with} \quad (\mathcal{M}^{\text{clas}} \mathbf{p})_j = \sum_{i=1}^n q(j|i) \, p_i \quad , \quad q(j|i) \geq 0 \quad , \quad \sum_{j=1}^n q(j|i) = 1 \; . \tag{5.6}$$

Save for the wavepacket reduction (5.4), all physical dynamical processes on quantum systems are given by quantum operations[2]. Let a system S interact with another system E at times $t \geq 0$. If S and E are initially in a product state $\rho(0) \otimes \rho_\mathsf{E}(0)$ and SE can be considered as an isolated system, so that its dynamics is governed by the Schrödinger equation, then the reduced state of S at time t reads

$$\rho(t) = \mathrm{tr}_\mathsf{E}\big(e^{-\mathrm{i}tH_{\mathsf{SE}}} \rho(0) \otimes \rho_\mathsf{E}(0) e^{\mathrm{i}tH_{\mathsf{SE}}}\big) \; . \tag{5.7}$$

[2] In order to include the transformation (5.4), many authors define a more general notion of quantum operation by relaxing the trace-preserving condition and replacing it by $\mathrm{tr}(\mathcal{M}(\rho)) \leq 1$ for any $\rho \in \mathcal{E}(\mathcal{H})$. The state transformation is then given by the non-linear map $\rho \mapsto \mathcal{M}(\rho)/\mathrm{tr}(\mathcal{M}(\rho))$. Theorems 5.2.2 and 5.2.3 can be easily adapted to this more general definition. In particular, the Kraus decomposition (5.10) holds, with Kraus operators A_i satisfying $\sum_i A_i^* A_i \leq 1$.

Here $H_{\mathsf{SE}} = H_{\mathsf{S}} + H_{\mathsf{E}} + \lambda H_{\rm int}$ is the Hamiltonian of SE, where H_{S} and H_{E} are the Hamiltonians of S and E, $H_{\rm int}$ their coupling Hamiltonian, and λ the coupling constant. The time-evolved state (5.7) is related to the initial state $\rho(0)$ by a quantum operation $\mathcal{M}_t$, i.e., $\rho(t) = \mathcal{M}_t\rho(0)$. The Stinespring theorem says that any quantum operation $\mathcal{M}$ can be viewed as a reduced evolution of the system coupled to an auxiliary system (ancilla).

Theorem 5.2.2. (Stinespring [229]) *Let $\mathcal{M}$ be a quantum operation $\mathcal{B}(\mathcal{H}_{\mathsf{S}}) \to \mathcal{B}(\mathcal{H}_{\mathsf{S}})$. Then one can find an ancilla Hilbert space $\mathcal{H}_{\mathsf{E}}$, a state $|\epsilon_0\rangle \in \mathcal{H}_{\mathsf{E}}$, and a unitary operator U on $\mathcal{H}_{\mathsf{SE}}$ such that $\mathcal{M}(\rho) = {\rm tr}_{\mathsf{E}}(U\rho \otimes |\epsilon_0\rangle\langle\epsilon_0| U^*)$.*

It is appropriate at this point to review a few well-known facts from the theory of CP maps on C^*-algebras. The *adjoint* $\mathcal{M}^*$ with respect to the trace of $\mathcal{M} : \mathcal{B}(\mathcal{H}_{\mathsf{S}}) \to \mathcal{B}(\mathcal{H}'_{\mathsf{S}})$ is the map $\mathcal{M}^* : \mathcal{B}(\mathcal{H}'_{\mathsf{S}}) \to \mathcal{B}(\mathcal{H}_{\mathsf{S}})$ defined by ${\rm tr}[A\mathcal{M}(\rho)] = {\rm tr}[\mathcal{M}^*(A)\rho]$ for any $A \in \mathcal{B}(\mathcal{H}'_{\mathsf{S}})$ and $\rho \in \mathcal{B}(\mathcal{H}_{\mathsf{S}})$. If $\mathcal{M}$ is a quantum operation then $\mathcal{M}^*$ is also a CP map and is unity-preserving, $\mathcal{M}^*(1) = 1$. According to Stinespring's theorem, one has

$$\mathcal{M}^*(X) = \langle\epsilon_0|U^*X \otimes 1U|\epsilon_0\rangle \tag{5.8}$$

for any $X \in \mathcal{B}(\mathcal{H})$. It follows that $\mathcal{M}^*$ satisfies the Kadyson-Schwarz inequality

$$|\mathcal{M}^*(X)|^2 \leq \mathcal{M}^*(|X|^2) \; . \tag{5.9}$$

Theorem 5.2.3. (Kraus [149]) *A linear map $\mathcal{M}$ from $\mathcal{B}(\mathcal{H}_{\mathsf{S}})$ into itself is a quantum operation if and only if it admits the representation*

$$\mathcal{M}(\rho) = \sum_i A_i \rho A_i^* \; , \tag{5.10}$$

where $\{A_i\}$ is a countable family of operators on $\mathcal{H}_{\mathsf{S}}$ satisfying $\sum_i A_i^* A_i = 1$.

For infinite dimensional Hilbert spaces and in the more general C*-algebra setting, the Kraus decomposition holds under the additional assumption that $\mathcal{M}$ is normal, that is, ultra-weakly continuous. One usually deduces it from Stinespring's theorem. In our finite-dimensional setting, however, a simple direct proof of Theorem 5.2.3 exists (see Remark 5.2.4 below). One can then obtain the Stinespring theorem from the Kraus decomposition as follows. Let $\{|k\rangle\}_{k=1}^{n_{\mathsf{S}}}$ be an orthonormal basis of $\mathcal{H}_{\mathsf{S}}$ and $\mathcal{H}_{\mathsf{E}}$ be a (possibly infinite-dimensional) Hilbert space with orthonormal basis $\{|\epsilon_i\rangle\}$. Define the vectors $|\Psi_{k0}\rangle = \sum_i A_i|k\rangle|\epsilon_i\rangle$. Using $\sum_i A_i^* A_i = 1$, one finds that these vectors form an orthonormal family in $\mathcal{H}_{\mathsf{SE}}$, which can be completed so as to get an orthonormal basis $\{|\Psi_{kl}\rangle\}$. Then $\mathcal{M}^*(X) = \langle\epsilon_0|U^*X \otimes 1\, U|\epsilon_0\rangle$ for any $X \in \mathcal{B}(\mathcal{H}_{\mathsf{S}})$, where the unitary U on $\mathcal{H}_{\mathsf{SE}}$ is defined by $U|k\rangle|\epsilon_l\rangle = |\Psi_{kl}\rangle$ for any k and l.

Remark 5.2.4. *Any quantum operation $\mathcal{B}(\mathcal{H}_{\mathsf{S}}) \to \mathcal{B}(\mathcal{H}_{\mathsf{S}})$ with $\dim \mathcal{H}_{\mathsf{S}} = n_{\mathsf{S}} < \infty$ admits a Kraus decomposition (5.10) with at most n_{S}^2 operators A_i. Consequently, one can choose the ancilla space $\mathcal{H}_{\mathsf{E}}$ in Theorem 5.2.2 of dimension $\dim \mathcal{H}_{\mathsf{E}} = n_{\mathsf{S}}^2$.*

Sketch the proof [180]. To show that $\mathcal{M}$ has the form (5.10), consider the operator $B = \mathcal{M} \otimes 1(|\Psi_{\rm ent}\rangle\langle\Psi_{\rm ent}|)$ with $|\Psi_{\rm ent}\rangle = n_{\mathsf{S}}^{-1/2} \sum_k |k\rangle|k\rangle \in \mathcal{H}_{\mathsf{SS}}$ as above. Since $\mathcal{M}$ is CP, one has $B \geq 0$. Let $|\widetilde{\Phi}_i\rangle$ be orthogonal eigenvectors of B, normalized in such a way that $n_{\mathsf{S}} B = \sum_i |\widetilde{\Phi}_i\rangle\langle\widetilde{\Phi}_i|$. Then define the Kraus operators A_i as the operators associated to $|\widetilde{\Phi}_i\rangle$ by the isomorphism (4.5) between $\mathcal{B}(\mathcal{H}_{\mathsf{S}})$ and $\mathcal{H}_{\mathsf{SS}}$. □

It is important to realize that the Kraus decomposition is not unique. For indeed, if $\{A_i\}_{i=1}^p$ is a family of Kraus operators for $\mathcal{M}$ and $(u_{ji})_{i,j=1}^q$ is a unitary matrix of size $q \geq p$, then the operators

$$B_j = \sum_{i=1}^p \overline{u}_{ji} A_i \quad , \quad j = 1, \ldots, q \; , \tag{5.11}$$

define another family of Kraus operators for $\mathcal{M}$. Conversely, two families $\{A_i\}_{i=1}^p$ and $\{B_j\}_{j=1}^q$ of Kraus operators for $\mathcal{M}$ with $p \leq q < \infty$ are related to each other by (5.11). Actually, let B and $|\Psi_{\rm ent}\rangle$ be defined as in the Remark 5.2.4 above. Then $B = \sum_i |\widetilde{\mu}_i\rangle\langle\widetilde{\mu}_i| = \sum_j |\widetilde{\nu}_j\rangle\langle\widetilde{\nu}_j|$ with

$$|\widetilde{\mu}_i\rangle = n_{\mathsf{S}}^{-\frac{1}{2}} \sum_k (A_i|k\rangle)|k\rangle \quad , \quad |\widetilde{\nu}_j\rangle = n_{\mathsf{S}}^{-\frac{1}{2}} \sum_k (B_j|k\rangle)|k\rangle \; . \tag{5.12}$$

In view of the link (4.16) between pure state decompositions of a non-negative operator, one has $|\widetilde{\nu}_j\rangle = \sum_i \overline{u}_{ji}|\widetilde{\mu}_i\rangle$ with $(u_{ji})_{i,j=1}^q$ unitary. This implies (5.11).

Given a purification $|\Psi\rangle$ of ρ on $\mathcal{H}_\mathsf{S} \otimes \mathcal{H}_\mathsf{R}$ and a quantum operation $\mathcal{M} : \mathcal{B}(\mathcal{H}_\mathsf{S}) \to \mathcal{B}(\mathcal{H}'_\mathsf{S})$, it is natural to ask about purifications of $\mathcal{M}(\rho)$. A slight generalization of Theorem 5.2.2 ensures that there exist a vector $|\epsilon_0\rangle \in \mathcal{H}_\mathsf{E}$ and a unitary $U : \mathcal{H}_\mathsf{S} \otimes \mathcal{H}_\mathsf{E} \to \mathcal{H}'_\mathsf{S} \otimes \mathcal{H}'_\mathsf{E}$ such that $\mathcal{M}(\rho) = \mathrm{tr}_{\mathsf{E}'}(U\rho \otimes |\epsilon_0\rangle\langle\epsilon_0|U^*)$. Therefore,

$$|\Psi_\mathcal{M}\rangle = 1_\mathsf{R} \otimes U|\Psi\rangle|\epsilon_0\rangle = \sum_{k=1}^n \sum_{i=1}^p \sqrt{p_k}(A_i|k\rangle)|f_k\rangle|\epsilon'_i\rangle \tag{5.13}$$

is a purification of $\mathcal{M}(\rho)$ on $\mathcal{H}'_\mathsf{S} \otimes \mathcal{H}_\mathsf{R} \otimes \mathcal{H}'_\mathsf{E}$. In the second equality, $\{|k\rangle\}$ is an orthonormal eigenbasis of ρ, $\{|f_k\rangle\}$ is the orthonormal family of $\mathcal{H}_\mathsf{R}$ such that $|\Psi\rangle = \sum_k \sqrt{p_k}|k\rangle|f_k\rangle$, and $\{|\epsilon'_i\rangle\}$ is an orthonormal basis of $\mathcal{H}'_\mathsf{E}$ such that $U|k\rangle|\epsilon_0\rangle = \sum_i (A_i|k\rangle)|\epsilon'_i\rangle$ (see the expression of U in terms of the Kraus operators after Theorem 5.2.3).

5.3 Generalized measurements

For the quantum operation $\mathcal{M}_\Pi$ defined by (5.5), the orthogonal projectors Π_i form a family of Kraus operators. One may wonder if more general quantum operations, given by Kraus operators A_i which are not necessarily orthogonal projectors, correspond to some kind of measurements. The answer is yes: such operations can always be obtained by coupling the system S to an auxiliary system E (the ancilla) and subsequently performing a von Neumann measurement on E.

Theorem 5.3.1. (Neumark extension theorem) *Let $\{A_i\}_{i=1}^p$ be a finite family of operators satisfying $\sum_i A_i^* A_i = 1$. Then there exist a space $\mathcal{H}_\mathsf{E}$ with dimension $\dim \mathcal{H}_\mathsf{E} = p$, a pure state $|\epsilon_0\rangle \in \mathcal{H}_\mathsf{E}$, an orthonormal family $\{\pi_i^\mathsf{E}\}$ of projectors in $\mathcal{B}(\mathcal{H}_\mathsf{E})$, and a unitary operator U on $\mathcal{H}_\mathsf{SE}$ such that for any density matrix $\rho \in \mathcal{E}(\mathcal{H}_\mathsf{S})$,*

$$A_i \rho A_i^* = \mathrm{tr}_\mathsf{E}(1 \otimes \pi_i^\mathsf{E}\, U\rho \otimes |\epsilon_0\rangle\langle\epsilon_0|U^* 1 \otimes \pi_i^\mathsf{E}) \,. \tag{5.14}$$

Proof. Use the same arguments as in the above proof of Stinespring's theorem from Theorem 5.2.3, and define $\pi_i^\mathsf{E} = |\epsilon_i\rangle\langle\epsilon_i|$. □

Definition 5.3.2. *A generalized measurement is given by a family $\{M_i\}$ of non-negative operators M_i satisfying $\sum_i M_i = 1$ (positive operator valued measure, abbreviated as POVM) together with a family of operators $\{A_i\}$ such that $M_i = A_i^* A_i$. The conditional state $\rho_{S|i}$ given outcome i and the probability of this outcome read*

$$\rho_{S|i} = p_i^{-1} A_i \rho A_i^* \quad , \quad p_i = \mathrm{tr}(M_i \rho) \,. \tag{5.15}$$

According to Theorem 5.3.1, any generalized measurement can be realized by letting the system S interact with an ancilla E in the state $|\epsilon_0\rangle$ and subsequently performing a von Neumann measurement on E, that is, coupling E to a macroscopic apparatus with pointer P. The interaction between S and E first transforms the initial state $\rho_\mathsf{S} \otimes |\epsilon_0\rangle\langle\epsilon_0|$ into $\rho_\mathsf{SE} = U\rho_\mathsf{S} \otimes |\epsilon_0\rangle\langle\epsilon_0|U^*$, U being a unitary evolution operator on $\mathcal{H}_\mathsf{SE}$, and the subsequent von Neumann measurement leads to the wavepacket reduction for the system SP (compare with (5.3) and (5.4))

$$\textbf{outcome } i\text{:} \qquad \rho_\mathsf{SP} \to \rho_{\mathsf{SP}|i} = p_i^{-1}\, \mathrm{tr}_\mathsf{E}(1 \otimes \pi_i^\mathsf{E}\, \rho_\mathsf{SE} 1 \otimes \pi_i^\mathsf{E}) \otimes |i\rangle\langle i| = p_i^{-1} A_i \rho_\mathsf{S} A_i^* \otimes |i\rangle\langle i| \,, \tag{5.16}$$

where $p_i = \mathrm{tr}(1 \otimes \pi_i^\mathsf{E} \rho_\mathsf{SE}) = \mathrm{tr}(M_i \rho_\mathsf{S})$ is the probability of outcome i, in agreement with (5.15).

One has $A_i = U_i M_i^{1/2}$ (polar decomposition) for some unitary operator U_i depending on i. The conditional states $\rho_{\mathsf{S}|i}$ are thus characterized by the POVM $\{M_i\}$ up to unitary conjugations, which introduce a freedom in choosing the output state associated to each measurement outcome. For instance, if $M_i = |\widetilde{\mu}_i\rangle\langle\widetilde{\mu}_i|$ are of rank one then $A_i = |i\rangle\langle\widetilde{\mu}_i|$ for some arbitrary normalized vector $|i\rangle$ and the output conditional states are $\rho_{\mathsf{S}|i} = |i\rangle\langle i|$. One usually takes the vectors $|i\rangle$ to form an orthonormal basis (which can be identified to the pointer state basis of Sec. 5.1), in such a way that the states $\rho_{\mathsf{S}|i}$ be perfectly distinguishable (this happens if the $\rho_{\mathsf{S}|i}$ are orthogonal only, see chapter 6 below). One should keep in mind, however, that the probability $p_i = \langle\widetilde{\mu}_i|\rho|\widetilde{\mu}_i\rangle$ of outcome i is independent of the choice of $\{|i\rangle\}$. If one is interested only in functions of the post-measurement states $\rho_{\mathsf{S}|i}$ which are invariant under unitary conjugations (as, for instance, the von Neumann entropy), then the generalized measurement can be fully specified by the measurement operators M_i. Thanks to the Neumark extension theorem, these operators may be written as

$$M_i = A_i^* A_i = \langle\epsilon_0|U^* 1 \otimes \pi_i^\mathsf{E} U|\epsilon_0\rangle \,. \tag{5.17}$$

As stressed above, in the absence of read-out the state of the system after the measurement is the average of the conditional states,

$$\mathcal{M}(\rho) = \sum_i p_i \rho_{S|i} = \sum_i A_i \rho A_i^* \ , \tag{5.18}$$

in analogy with (5.5). This defines a quantum operation $\mathcal{M}$, the Kraus decomposition of which specifies the state preparation associated with the wavepacket reduction.

Writing the spectral decomposition of each operator M_i, one observes that

$$M_i = \sum_{k=1}^{r_i} |\widetilde{\mu}_{ik}\rangle\langle\widetilde{\mu}_{ik}| \quad , \quad \sum_i M_i = \sum_{i,k} |\widetilde{\mu}_{ik}\rangle\langle\widetilde{\mu}_{ik}| = 1 \ , \tag{5.19}$$

where $r_i = \mathrm{rank}(M_i)$ and $|\widetilde{\mu}_{ik}\rangle$ are unnormalized eigenvectors with norms equal to the square roots of the corresponding eigenvalues. The last condition in (5.19) implies that either $\{|\widetilde{\mu}_{ik}\rangle\}$ is an orthonormal basis, in which case $\{M_i\}$ is an orthonormal family of projectors (von Neumann measurement), or $\{|\widetilde{\mu}_{ik}\rangle\}$ is a non-orthogonal family containing more than n_{S} vectors, in which case at least two eigenvalues $\|\widetilde{\mu}_{ik}\|$ are strictly smaller than one and $\{M_i\}$ is not a von Neumann measurement.

The set of all POVMs is a convex set. Its boundary and extremal points have been studied in [66].

Remark 5.3.3. *An alternative version of Theorem 5.3.1 states that if $m = \sum_i r_i$ with $r_i = \mathrm{rank}(M_i)$, then there exist a space $\mathcal{H}_{\mathsf{E}}$ with dimension $m - n_{\mathsf{S}} + 1$, a state $|\epsilon_0\rangle \in \mathcal{H}_{\mathsf{E}}$, and a von Neumann measurement $\{\Pi_i^{\mathsf{SE}}\}$ on $\mathcal{H}_{\mathsf{SE}}$ such that*

$$M_i = \langle\epsilon_0|\Pi_i^{\mathsf{SE}}|\epsilon_0\rangle \ . \tag{5.20}$$

The interesting point is that the dimension of the ancilla space $\mathcal{H}_{\mathsf{E}}$ can be smaller than p in Theorem 5.3.1 (for instance $\dim \mathcal{H}_{\mathsf{E}} = p - n_{\mathsf{S}} + 1$ for rank-one operators M_i).

Sketch of the proof [192]. Note that $m \geq n_{\mathsf{S}}$ by the observation above. Define

$$|\zeta_{ik}\rangle = |\widetilde{\mu}_{ik}\rangle|\epsilon_0\rangle + \sum_{l=1}^{m-n_{\mathsf{S}}} c_{ik,l}|\phi\rangle|\epsilon_l\rangle \ , \tag{5.21}$$

where $|\widetilde{\mu}_{ik}\rangle$ is as in (5.19), $|\phi\rangle \in \mathcal{H}_{\mathsf{S}}$ is an arbitrary state, and $\{|\epsilon_l\rangle\}_{l=0}^{m-n_{\mathsf{S}}}$ is an orthonormal basis of $\mathcal{H}_{\mathsf{E}}$. The coefficients $c_{ik,l}$ may be chosen such that $\{|\zeta_{ik}\rangle\}$ is an orthonormal family of $\mathcal{H}_{\mathsf{SE}}$. To establish this statement, set $c_{ik,l} = \langle l|\widetilde{\mu}_{ik}\rangle$ for $m - n_{\mathsf{S}} < l \leq m$, with $\{|l\rangle\}_{l=m-n_{\mathsf{S}}+1}^{m}$ an orthonormal basis of $\mathcal{H}_{\mathsf{S}}$, and let $\mathbf{c}_l \in \mathbb{C}^m$ be the vector with components $c_{ik,l}$. Then $\mathbf{c}_l \cdot \mathbf{c}_{l'} = \delta_{ll'}$ for any $l, l' > m - n_{\mathsf{S}}$, as a result of $\sum_i M_i = 1$. One can choose the $(m - n_{\mathsf{S}})$ other vectors $\mathbf{c}_l$ in such a way that $(\mathbf{c}_1, \ldots, \mathbf{c}_m)$ forms a $m \times m$ unitary matrix. Then $\Pi_i^{\mathsf{SE}} = \sum_k |\zeta_{ik}\rangle\langle\zeta_{ik}|$ has the desired property. □

5.4 Connections between POVMs, quantum operations, and state ensembles

To each POVM one can associate a quantum operation and vice-versa. Similarly, there is a canonical way to associate to a quantum operation a state ensemble and vice-versa. These correspondences depend on an orthonormal basis $\{|i\rangle\}_{i=1}^m$ of a fictitious pointer P with m-dimensional space $\mathcal{H}_{\mathsf{P}}$. It has been already seen above that one can associate to a POVM $\{M_i\}_{i=1}^m$ on S a quantum operation with Kraus operators A_i such that $M_i = A_i^* A_i$. This operation implements the state changes in the measurement process in the absence of readout. If we imagine that S is coupled to P and that the measurement is performed on both S and P, one may consider the Kraus operators $A_{ik} = |k\rangle|i\rangle\langle\widetilde{\mu}_{ik}|$ such that $M_i = \sum_k A_{ik}^* A_{ik}$, where $\{|k\rangle\}_{k=1}^{n_{\mathsf{S}}}$ is an orthonormal basis of $\mathcal{H}_{\mathsf{S}}$ and $|\widetilde{\mu}_{ik}\rangle$ are the unnormalized eigenvectors of M_i in (5.19). Provided that there is no readout of the measurement on S, one may trace the post-measurement states over $\mathcal{H}_{\mathsf{S}}$. The conditional states of P are given by $\rho_{\mathsf{P}|i} = p_{ik}^{-1} \mathrm{tr}_{\mathsf{S}}(A_{ik}\rho A_{ik}^*) = |i\rangle\langle i|$ with $p_{ik} = \langle\widetilde{\mu}_{ik}|\rho|\widetilde{\mu}_{ik}\rangle$, and the corresponding probability is $p_i = \sum_k p_{ik} = \mathrm{tr}(M_i\rho)$. The state changes in the absence of readout are implemented by the quantum operation $\mathcal{M} : \mathcal{B}(\mathcal{H}_{\mathsf{S}}) \to \mathcal{B}(\mathcal{H}_{\mathsf{P}})$ defined by

$$\mathcal{M}(\rho) = \sum_i \mathrm{tr}(M_i\rho)|i\rangle\langle i| \ , \ \rho \in \mathcal{E}(\mathcal{H}_{\mathsf{S}}) \quad \Leftrightarrow \quad \mathcal{M}^*(|i\rangle\langle j|) = M_i\delta_{ij} \ , \ i, j = 1, \ldots, m \ . \tag{5.22}$$

Conversely, if $\mathcal{M}$ is a quantum operation $\mathcal{B}(\mathcal{H}_{\mathsf{S}}) \to \mathcal{B}(\mathcal{H}_{\mathsf{P}})$ then $M_i = \mathcal{M}^*(|i\rangle\langle i|)$ defines a POVM $\{M_i\}_{i=1}^m$ (actually, $M_i \geq 0$ by the positivity of $\mathcal{M}^*$ and $\sum_i M_i = \mathcal{M}^*(1) = 1$). Therefore, for a given orthonormal

basis $\{|i\rangle\}_{i=1}^m$ of $\mathcal{H}_\mathsf{P}$, there is a one-to-one correspondence between POVMs $\{M_i\}_{i=1}^m$ on $\mathcal{H}_\mathsf{S}$ and quantum operations $\mathcal{M} : \mathcal{B}(\mathcal{H}_\mathsf{S}) \to \mathcal{B}(\mathcal{H}_\mathsf{P})$ of the form (5.22).

A similar one-to-one correspondence can be found between state ensembles on $\mathcal{H}_\mathsf{S}$ with fixed probabilities $\{\eta_i\}_{i=1}^m$ and quantum operations $\mathcal{B}(\mathcal{H}_\mathsf{P}) \to \mathcal{B}(\mathcal{H}_\mathsf{S})$ such that $\mathcal{M}(|i\rangle\langle j|) = 0$ for $i \neq j$. This correspondence is given by

$$\rho_i = \mathcal{M}(|i\rangle\langle i|) \ , \ \ i = 1, \dots, m \ . \tag{5.23}$$

In fact, if $\mathcal{M} : \mathcal{B}(\mathcal{H}_\mathsf{P}) \to \mathcal{B}(\mathcal{H}_\mathsf{S})$ is a quantum operation then $\{\rho_i, \eta_i\}_{i=1}^m$ is clearly an ensemble on $\mathcal{H}_\mathsf{S}$. Conversely, if $\{\rho_i, \eta_i\}_{i=1}^m$ is an ensemble of m states, let us write the spectral decompositions $\rho_i = \sum_k p_{ik}|\psi_{ik}\rangle\langle\psi_{ik}|$. Then the operation with Kraus operators $A_{ik} = \sqrt{p_{ik}}|\psi_{ik}\rangle\langle i|$ has the required property.

5.5 Transpose operation and least square measurement

5.5.1 Recovery operation in quantum error correction

The notion of transpose operation was introduced by Ohya and Petz in their monograph [183]. It plays the role of an approximate reversal of a quantum operation, in a sense that will be made more precise below.

Definition 5.5.1. *Let $\mathcal{M} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ be a quantum operation and $\rho \in \mathcal{E}(\mathcal{H})$ be a state such that $\mathcal{M}(\rho) > 0$. The transpose operation of $\mathcal{M}$ for ρ is the quantum operation $\mathcal{R}_{\mathcal{M},\rho} : \mathcal{B}(\mathcal{H}') \to \mathcal{B}(\mathcal{H})$ with Kraus operators $R_i = \rho^{\frac{1}{2}} A_i^* \mathcal{M}(\rho)^{-\frac{1}{2}}$, where $\{A_i\}$ is a family of Kraus operators for $\mathcal{M}$. It is independent of the Kraus decomposition of $\mathcal{M}$. Actually, for any $\sigma \in \mathcal{E}(\mathcal{H}')$,*

$$\mathcal{R}_{\mathcal{M},\rho}(\sigma) = \rho^{\frac{1}{2}} \mathcal{M}^*\big(\mathcal{M}(\rho)^{-\frac{1}{2}} \sigma \mathcal{M}(\rho)^{-\frac{1}{2}}\big)\rho^{\frac{1}{2}} \ . \tag{5.24}$$

One easily checks that $\sum_i R_i^* R_i = 1$, so that $\mathcal{R}_{\mathcal{M},\rho}$ is indeed a quantum operation, and that $\mathcal{R}_{\mathcal{M},\rho} \circ \mathcal{M}(\rho) = \rho$. Furthermore, transposing twice amounts to do nothing, that is, the transpose of $\mathcal{R}_{\mathcal{M},\rho}$ for the state $\mathcal{M}(\rho)$ is equal to $\mathcal{M}$.

The operation $\mathcal{R}_{\mathcal{M},\rho}$ appears naturally in the context of quantum error correction. The problem of quantum error correction is to send a state ρ over a noisy quantum communication channel in such a way that ρ is resilient to the effect of the noise in the channel. The state ρ is encoded via a unitary transformation into a subspace $\mathcal{H}_C$ of the Hilbert space $\mathcal{H}$ of the quantum channel. The noise is described by some quantum operation $\mathcal{M}$.

Proposition 5.5.2. *Let $\mathcal{M}$ be a quantum operation on $\mathcal{B}(\mathcal{H})$ with Kraus operators $\{A_i\}$. Let Π_C denote the orthogonal projector onto a subspace $\mathcal{H}_C \subset \mathcal{H}$ and $\mathbb{E}_C : \rho \mapsto \Pi_C \, \rho \, \Pi_C$ be the conditional expectation onto the space of operators supported on $\mathcal{H}_C$. There exists a recovery quantum operation $\mathcal{R}$ on $\mathcal{B}(\mathcal{H})$ satisfying $\mathcal{R} \circ \mathcal{M} \circ \mathbb{E}_C = \mathbb{E}_C$ if and only if the following condition holds:*

$$\mathbb{E}_C(A_i^* A_j) = a_{ij} \Pi_C \ , \tag{5.25}$$

where (a_{ij}) is a self-adjoint matrix. If this condition is satisfied then for any ρ with support $\mathrm{ran}(\rho) \subset \mathcal{H}_C$, the transpose operation $\mathcal{R}_{\mathcal{M},\rho}$ is a recovery quantum operation.

We refer the reader to the book of Nielsen and Chuang [180] for a proof of the necessary and sufficient condition (5.25). Some bibliographic information on this topic can also be found there.

Proof of the second statement. By taking advantage of the non-uniqueness of the Kraus decomposition, (5.25) can be transformed into $\mathbb{E}_C(B_i^* B_j) = p_i \delta_{ij} \Pi_C$, where the Kraus operators B_i are given by (5.11) with $(u_{ij})(a_{ij})(u_{ij})^*$ the diagonal matrix with entries p_i. Together with the polar decomposition, this implies $B_j \Pi_C = \sqrt{p_j} W_j$ with $W_j = V_j \Pi_C$ satisfying $W_i^* W_j = \delta_{ij} \Pi_C$, the V_j being some unitary operators. Thus the subspaces $V_j \mathcal{H}_C$ are orthogonal for different j's and the restriction of $\sum_j W_j W_j^*$ to the subspace $\mathcal{V} = \oplus_j V_j \mathcal{H}_C$ equals the identity. If $\rho = \mathbb{E}_C(\rho)$ and the restriction of ρ to $\mathcal{H}_C$ is invertible, then $\mathcal{M}(\rho) = \sum_j p_j W_j \rho W_j^*$ and $\mathcal{M}(\rho)^{-1/2} = \sum_j W_j \rho^{-1/2} W_j^* / \sqrt{p_j}$, the last operator being defined on $\mathcal{V}$. A simple calculation then shows that $\mathcal{R}_{\mathcal{M},\rho} \circ \mathcal{M} \circ \mathbb{E}_C = \mathbb{E}_C$, as stated in the Proposition. □

5.5.2 Transpose operation as an approximate reverse operation

Since the condition (5.25) is not always fulfilled, it is natural to ask whether one can find an optimal imperfect recovery map, which would enable to recover a given ensemble $\{\rho_i, \eta_i\}$ subject to some noise with a maximal fidelity. A notion of fidelity has been introduced by Schumacher [213]. Its definition is as follows (for more detail and motivations from classical information theory, see [180]). Given a state $\rho \in \mathcal{E}(\mathcal{H}_\mathsf{S})$, consider a purification $|\Psi_\rho\rangle$ of ρ on $\mathcal{H}_\mathsf{S} \otimes \mathcal{H}_\mathsf{R}$, where R is a reference system with Hilbert space $\mathcal{H}_\mathsf{R} \simeq \mathcal{H}_\mathsf{S}$. For instance, $|\Psi_\rho\rangle$ can be given by (4.12). If ρ is a mixed state then $|\Psi_\rho\rangle$ is SR-entangled (Sec. 4.4). The *entanglement fidelity* of ρ quantifies how well this entanglement is preserved when the system S is subject to some noise modelized by a quantum operation $\mathcal{M}$ on $\mathcal{B}(\mathcal{H}_\mathsf{S})$. It is defined by

$$F_{\mathrm{e}}(\rho, \mathcal{M}) = \langle \Psi_\rho | \mathcal{M} \otimes 1 (|\Psi_\rho\rangle\langle\Psi_\rho|) |\Psi_\rho\rangle \; . \tag{5.26}$$

Since different purifications of ρ on $\mathcal{H}_\mathsf{SR}$ are related by unitaries acting on $\mathcal{H}_\mathsf{R}$, the right-hand side of (5.26) does not depend on the chosen purification. As a consequence of the positivity and the trace-preserving property of $\mathcal{M}$, one has $0 \leq F_{\mathrm{e}}(\rho, \mathcal{M}) \leq \mathrm{tr}_\mathsf{SR}[\mathcal{M} \otimes 1(|\Psi_\rho\rangle\langle\Psi_\rho|)] = \mathrm{tr}[\mathcal{M}(\rho)] = 1$. Plugging (4.12) and (5.10) into (5.26), a simple calculation yields

$$F_{\mathrm{e}}(\rho, \mathcal{M}) = \sum_j \big|\mathrm{tr}(A_j \rho)\big|^2 \; , \tag{5.27}$$

where $\{A_j\}$ is a family of Kraus operators for $\mathcal{M}$. Note that the sum in the right-hand side does not depend on the choice of Kraus decomposition (this follows from (5.11)), as it should be. For a pure state $\rho_\psi = |\psi\rangle\langle\psi|$, the entanglement fidelity reduces to the input-output fidelity $F(\rho_\psi, \mathcal{M}) = \langle\psi|\mathcal{M}(|\psi\rangle\langle\psi|)|\psi\rangle$. One infers from (5.27) that $F_{\mathrm{e}}(\rho, \mathcal{M})$ is a convex function of ρ.

Let us now consider an ensemble of states $\{\rho_i, \eta_i\}_{i=1}^m$. The corresponding average entanglement fidelity is defined by

$$\overline{F}_{\mathrm{e}}(\{\rho_i, \eta_i\}, \mathcal{M}) = \sum_i \eta_i F_{\mathrm{e}}(\rho_i, \mathcal{M}) \; . \tag{5.28}$$

This fidelity belongs to the interval $[0, 1]$.

Proposition 5.5.3. (Barnum and Knill [29]) *If the states ρ_i commute with $\rho = \sum_i \eta_i \rho_i$, then*

$$\overline{F}_{\mathrm{e}}\big(\{\rho_i, \eta_i\}, \mathcal{R}_{\mathcal{M},\rho} \circ \mathcal{M}\big) \geq \overline{F}_{\mathrm{e}}\big(\{\rho_i, \eta_i\}, \mathcal{R}_{\mathrm{opt}} \circ \mathcal{M}\big)^2 \; , \tag{5.29}$$

where $\mathcal{R}_{\mathcal{M},\rho}$ is the transpose operation of $\mathcal{M}$ for ρ and $\mathcal{R}_{\mathrm{opt}}$ the optimal recovery quantum operation $\mathcal{R}$ maximizing $\overline{F}_e(\{\rho_i, \eta_i\}, \mathcal{R} \circ \mathcal{M})$.

Hence, if the minimal fidelity error is $1 - \overline{F}_e(\{\rho_i, \eta_i\}, \mathcal{R}_{\mathrm{opt}} \circ \mathcal{M}) = \eta$, then the fidelity error by using $\mathcal{R}_{\rho,\mathcal{M}}$ as the recovery operation is at most twice larger than this minimal error.

Proof. Taking advantage of the non-uniqueness of the Kraus decomposition, one can choose for any fixed i some families $\{R_j^{\mathrm{opt}\,(i)}\}$ and $\{A_k^{(i)}\}$ of Kraus operators for $\mathcal{R}^{\mathrm{opt}}$ and $\mathcal{M}$ satisfying

$$\mathrm{tr}\big(R_j^{\mathrm{opt}\,(i)} A_k^{(i)} \rho_i\big) = 0 \quad , \quad j \neq k \; . \tag{5.30}$$

Actually, given any families $\{R_m^{\mathrm{opt}}\}$ for $\mathcal{R}^{\mathrm{opt}}$ and $\{A_l\}$ for $\mathcal{M}$, the operators $R_j^{\mathrm{opt}\,(i)} = \sum_m u_{jm}^{(i)} R_m^{\mathrm{opt}}$ and $A_k^{(i)} = \sum_l \overline{v}_{kl}^{(i)} A_l$ have the required property if $(u_{jm}^{(i)})$ and $(v_{kl}^{(i)})$ are the unitary matrices in the singular decomposition of $(\mathrm{tr}(R_m^{\mathrm{opt}} A_l \rho_i))$. Since $\{R_j^{\mathrm{opt}\,(i)} A_k^{(i)}\}$ is a Kraus family for $\mathcal{R}^{\mathrm{opt}} \circ \mathcal{M}$, one obtains from (5.27), (5.28), and (5.30)

$$\overline{F}_{\mathrm{e}}\big(\{\rho_i, \eta_i\}, \mathcal{R}^{\mathrm{opt}} \circ \mathcal{M}\big) = \sum_{i,j} \eta_i \big|\mathrm{tr}(R_j^{\mathrm{opt}(i)} A_j^{(i)} \rho_i)\big|^2 \; . \tag{5.31}$$

We first consider the case $\rho_\mathcal{M} = \mathcal{M}(\rho) > 0$. Without loss of generality, we may assume that $\mathrm{ran}(R_j^{\mathrm{opt}(i)}) \subset \mathrm{ran}\,\rho_i \subset \mathrm{ran}\,\rho$, so that the operators

$$X_{ij} = \eta_i^{\frac{1}{4}} \rho_\mathcal{M}^{-\frac{1}{4}} A_j^{(i)} \rho^{\frac{1}{4}} \rho_i^{\frac{1}{2}} \quad , \quad Y_{ij} = \eta_i^{\frac{1}{4}} \rho_\mathcal{M}^{-\frac{1}{4}} B_j^{(i)} \rho^{\frac{1}{4}} \rho_i^{\frac{1}{2}} \quad \text{and} \quad (B_j^{(i)})^* = \rho^{-\frac{1}{2}} R_j^{\mathrm{opt}\,(i)} \rho_\mathcal{M}^{\frac{1}{2}} \tag{5.32}$$

are well-defined. Since $[\rho_i, \rho] = 0$, one finds by using twice the Cauchy-Schwarz inequality

$$\begin{aligned}\overline{F}_{\rm e}(\{\rho_i,\eta_i\},\mathcal{R}^{\rm opt}\circ\mathcal{M})^2 &= \Big(\sum_{i,j}\big|{\rm tr}(Y_{ij}^*X_{ij})\big|^2\Big)^2 \leq \sum_{i,j}\big({\rm tr}(Y_{ij}^*Y_{ij})\big)^2\sum_{i,j}\big({\rm tr}(X_{ij}^*X_{ij})\big)^2 \\ &\leq \sum_{i,j,k}\big|{\rm tr}(Y_{ij}^*Y_{ik})\big|^2\sum_{i,j,k}\big|{\rm tr}(X_{ij}^*X_{ik})\big|^2 \,. \end{aligned} \tag{5.33}$$

The transpose operation $\mathcal{R}_{\rho,\mathcal{M}}$ has Kraus operators $R_j^{(i)} = \rho^{\frac{1}{2}}(A_j^{(i)})^*\rho_{\mathcal{M}}^{-\frac{1}{2}}$. As a result,

$$\overline{F}_{\rm e}\big(\{\rho_i,\eta_i\},\mathcal{R}_{\rho,\mathcal{M}}\circ\mathcal{M}\big) = \sum_{i,j,k}\eta_i\big|{\rm tr}(R_j^{(i)}A_k^{(i)}\rho_i)\big|^2 = \sum_{i,j,k}\big|{\rm tr}(X_{ij}^*X_{ik})\big|^2 \,. \tag{5.34}$$

The first sum in the last member of (5.33) is equal to $\overline{F}_{\rm e}(\{\rho_i,\eta_i\},\mathcal{R}^{\rm opt}\circ\mathcal{B})$, where $\mathcal{B}$ is the CP map defined by $\mathcal{B}(\sigma) = \sum_k B_k^{(i)}\sigma(B_k^{(i)})^*$ (note that $\mathcal{B}$ does not depend on i). Even if $\mathcal{B}$ is not trace-preserving, with the help of (5.26) this fidelity can be bounded from above by ${\rm tr}[\mathcal{R}^{\rm opt}\circ\mathcal{B}(\rho)]$, which equals unity thanks to the identity $\mathcal{B}(\rho) = \mathcal{M}(\rho)$. This yields the inequality (5.29). If $\rho_{\mathcal{M}}$ is not invertible, one approximates $\mathcal{M}$ by some quantum operations $\mathcal{M}_\varepsilon$ satisfying $\mathcal{M}_\varepsilon(\rho) > 0$ for $\varepsilon > 0$ and $\mathcal{M}_\varepsilon \to \mathcal{M}$ as $\varepsilon \to 0$, and obtains the result by continuity. □

5.5.3 Least square measurement

Let us consider an ensemble $\{\rho_i,\eta_i\}_{i=1}^m$ of states of the system S forming a convex decomposition of $\rho_{\rm out} = \sum_i \eta_i\rho_i$. For any i, we denote by $\rho_i = \sum_k p_{ik}|\psi_{ik}\rangle\langle\psi_{ik}|$ the spectral decomposition of ρ_i and set $\rho_i = A_iA_i^*$, where $A_i = \sqrt{\rho_i}U_i$ is defined up to a unitary U_i. Introducing as in Sec. 5.4 an arbitrary orthonormal basis $\{|k\rangle\}_{k=1}^{n_\mathsf{S}}$ of $\mathcal{H}_\mathsf{S}$ and a fictitious pointer with m-dimensional space $\mathcal{H}_\mathsf{P}$ and orthonormal basis $\{|i\rangle\}_{i=1}^m$, one can choose

$$A_i = \sum_{k=1}^{n_\mathsf{S}}\sqrt{p_{ik}}|\psi_{ik}\rangle\langle k|\langle i| \;\in\; \mathcal{B}(\mathcal{H}_{\mathsf{SP}},\mathcal{H}_\mathsf{S}) \,. \tag{5.35}$$

We remark that A_i is associated to a purification of $\rho_i\otimes|i\rangle\langle i|$ on $\mathcal{H}_{\mathsf{SP}}\otimes\mathcal{H}_\mathsf{S}$ via the isometry (4.5) between $\mathcal{B}(\mathcal{H}_{\mathsf{SP}},\mathcal{H}_\mathsf{S})$ and $\mathcal{H}_{\mathsf{SP}}\otimes\mathcal{H}_\mathsf{S}$, namely, $|\Psi_i\rangle = \sum_k\sqrt{p_{ik}}|\psi_{ik}\rangle|i\rangle|k\rangle$. Moreover, $|\Psi_{\rm out}\rangle = \sum_i\sqrt{\eta_i}|\Psi_i\rangle$ is a purification of $\rho_{\rm out}$ on the same space.

The *least square measurement* [3] associated to $\{\rho_i,\eta_i\}_{i=1}^m$ is given by the Kraus and measurement operators

$$R_i^{\rm lsm} = \sqrt{\eta_i}A_i^*\rho_{\rm out}^{-\frac{1}{2}} = \sum_k\sqrt{\eta_ip_{ik}}|k\rangle|i\rangle\langle\psi_{ik}|\rho_{\rm out}^{-\frac{1}{2}} \quad , \quad M_i^{\rm lsm} = \big|R_i^{\rm lsm}\big|^2 = \eta_i\rho_{\rm out}^{-\frac{1}{2}}\rho_i\rho_{\rm out}^{-\frac{1}{2}} \tag{5.36}$$

for $i = 1,\ldots,m$. One indeed checks that $\sum_i M_i^{\rm lsm} = 1$, so that (5.36) defines a generalized measurement in the sense of Definition 5.3.2. While the operators $M_i^{\rm lsm}$ and thus the outcome probabilities $q_i = {\rm tr}(M_i^{\rm lsm}\sigma_\mathsf{S})$ (here σ_S is the system state) only depend on $\{\rho_i,\eta_i\}$, the post-measurement states also depend on the choice of the basis $\{|i\rangle\}$, as highlighted in chapter 5. The conditional and average post-measurement states of the pointer P are

$$\textbf{outcome } i\text{:} \qquad \sigma_\mathsf{S} \;\mapsto\; \sigma_{\mathsf{P}|i} = q_i^{-1}\,{\rm tr}_\mathsf{S}(R_i^{\rm lsm}\sigma_\mathsf{S}(R_i^{\rm lsm})^*) = |i\rangle\langle i| \tag{5.37}$$

$$\textbf{no readout}\text{:} \qquad \sigma_\mathsf{S} \;\mapsto\; \sigma_\mathsf{P} = \mathcal{M}^{\rm lsm}(\sigma_\mathsf{S}) = \sum_{i=1}^m q_i\sigma_{\mathsf{P}|i} = \sum_{i=1}^m q_i|i\rangle\langle i| \,. \tag{5.38}$$

For a pure state ensemble $\{|\psi_i\rangle,\eta_i\}_{i=1}^m$, the least square measurement consists of rank-one measurement operators $M_i = |\widetilde{\mu}_i\rangle\langle\widetilde{\mu}_i|$ with $|\widetilde{\mu}_i\rangle = \sqrt{\eta_i}\rho_{\rm out}^{-\frac{1}{2}}|\psi_i\rangle$. The vectors $|\widetilde{\mu}_i\rangle$ have the following property [125, 83], which elucidates the name given to the measurement: they minimize the sum of the square norms $\||\widetilde{\mu}_i\rangle - \sqrt{\eta_i}|\psi_i\rangle\|^2$ under the constraint $\sum_i|\widetilde{\mu}_i\rangle\langle\widetilde{\mu}_i| = 1$. If the $|\psi_i\rangle$ are linearly independent and span $\mathcal{H}_\mathsf{S}$, so that $m = n$, then $\{|\widetilde{\mu}_i\rangle\}$ is an orthonormal basis of $\mathcal{H}_\mathsf{S}$. In that case $\{M_i^{\rm lsm}\}$ is a von Neumann measurement (see Sec. 5.3).

[3] This measurement bears several names: it was referred to as the "pretty good measurement" in [114] and is also called "square-root measurement" by many authors.

Remark 5.5.4. *The aforementioned property of a least square measurement can be stated as follows:*

$$\min_{\{|\widetilde{\mu}_i\rangle\}} \left\{ \sum_{i=1}^{m} \left\| |\widetilde{\mu}_i\rangle - \sqrt{\eta_i}|\psi_i\rangle \right\|^2 \right\} = n_{\mathsf{S}} + 1 - 2\,\mathrm{tr}(\rho_{\mathrm{out}}^{\frac{1}{2}}) \quad \textit{with} \quad \rho_{\mathrm{out}} = \sum_i \eta_i |\psi_i\rangle\langle\psi_i| \,, \tag{5.39}$$

the minimum being over all families $\{|\widetilde{\mu}_i\rangle\}_{i=1}^m$ *in* $\mathcal{H}_{\mathsf{S}}$ *such that* $\sum_i |\widetilde{\mu}_i\rangle\langle\widetilde{\mu}_i| = 1$*. This minimum is achieved if and only if* $|\widetilde{\mu}_i\rangle = \sqrt{\eta_i}\rho_{\mathrm{out}}^{-1/2}|\psi_i\rangle$ *(up to irrelevant phase factors).*

Sketch of the proof. [125, 83] Define $A = \sum_i \sqrt{\eta_i}|\psi_i\rangle\langle i|$ and $B = \sum_i |\widetilde{\mu}_i\rangle\langle i|$ in analogy with (5.35). Then observe that the sum to be minimized in (5.39) is equal to $\|A^* - B^*\|_2^2 = 1 + n_{\mathsf{S}} - 2\mathrm{Re}\ \mathrm{tr}(AB^*)$, and use (4.3). □

As suggested by this result, the least square measurement plays an important role in distinguishing quantum states drawn from a given ensemble. This point will be discussed in Sec. 6.3 below.

Let us recall from Sec. 5.4 that the relation $\rho_i = \mathcal{M}(|i\rangle\langle i|)$, where $\{|i\rangle\}_{i=1}^m$ is a fixed orthonormal basis of $\mathcal{H}_{\mathsf{P}}$, can be used to associate to a quantum operation $\mathcal{M} : \mathcal{B}(\mathcal{H}_{\mathsf{P}}) \to \mathcal{B}(\mathcal{H}_{\mathsf{S}})$ an ensemble $\{\rho_i, \eta_i\}_{i=1}^m$ on $\mathcal{H}_{\mathsf{S}}$. Conversely, if $\{\rho_i, \eta_i\}$ is an ensemble on $\mathcal{H}_{\mathsf{S}}$, the operation $\mathcal{M}$ with Kraus operators $A_{ik} = A_i|k\rangle = \sqrt{p_{ik}}|\Psi_{ik}\rangle\langle i|$ satisfies this relation (here A_i is the operator (5.35)). Similarly, the relation (5.22) establishes a one-to-one correspondence between POVMs $\{M_i\}$ on $\mathcal{H}_{\mathsf{S}}$ and quantum operations $\mathcal{R} : \mathcal{B}(\mathcal{H}_{\mathsf{S}}) \to \mathcal{B}(\mathcal{H}_{\mathsf{P}})$. It was recognized by Barnum and Knill [29] that *the least square measurement associated to the ensemble* $\{\rho_i = \mathcal{M}(|i\rangle\langle i|), \eta_i\}$ *is nothing but the measurement corresponding to the transpose operation* $\mathcal{R}_{\mathcal{M},\rho_{\mathrm{in}}}$ *of* $\mathcal{M}$ *for the state* $\rho_{\mathrm{in}} = \sum_i \eta_i |i\rangle\langle i|$. Actually, since $\mathcal{M}(\rho_{\mathrm{in}}) = \rho_{\mathrm{out}}$, according to the Definition 5.5.1,

$$R_{ik} = \rho_{\mathrm{in}}^{\frac{1}{2}} A_{ik}^* \rho_{\mathrm{out}}^{-\frac{1}{2}} = \sqrt{\eta_i p_{ik}}|i\rangle\langle\psi_{ik}|\rho_{\mathrm{out}}^{-\frac{1}{2}} \tag{5.40}$$

are Kraus operators for $\mathcal{R}_{\mathcal{M},\rho_{\mathrm{in}}}$. Thus

$$M_i^{\mathrm{lsm}} = \eta_i \rho_{\mathrm{out}}^{-\frac{1}{2}} \rho_i \rho_{\mathrm{out}}^{-\frac{1}{2}} = \sum_k R_{ik}^* R_{ik} = \mathcal{R}_{\mathcal{M},\rho_{\mathrm{in}}}^*(|i\rangle\langle i|) \,. \tag{5.41}$$

Conversely, it is immediate to verify that $\mathcal{R}_{\mathcal{M},\rho_{\mathrm{in}}}(\sigma) = \sum_{ik} R_{ik}\sigma R_{ik}^* = \sum_i \mathrm{tr}(M_i^{\mathrm{lsm}}\sigma)|i\rangle\langle i|$, hence $\mathcal{R}_{\mathcal{M},\rho_{\mathrm{in}}}$ is the operation associated to $\{M_i^{\mathrm{lsm}}\}$ by the relation (5.22).

Chapter 6

Quantum state discrimination

Theory determines what we observe (A. Einstein).

The carriers of information in quantum communication and quantum computing are quantum systems, and the information is encoded in the states of those systems. After processing the information, it is necessary to perform measurements in order to read out the result of the computation. In other words, one has to determine the output state of the system. If these possible outputs form a set of orthogonal states, that is, if they are given by m known density matrices ρ_i with orthogonal supports, then it is easy to devise a measurement which discriminates them without any error (a von Neumann measurement with projectors Π_i onto $\mathrm{ran}(\rho_i)$ will do the job). However, when the ρ_i are non-orthogonal a perfect discrimination is impossible. Indeed, if two non-orthogonal states $|\psi_1\rangle$ and $|\psi_2\rangle$ could be discriminated perfectly then one could duplicate those states by producing copies of $|\psi_i\rangle$ if the measurement outcome is $i = 1, 2$, without prior knowledge on which of the two states one actually possesses. This would contradict the no-cloning theorem[1]. Consequently, one can extract less information from an ensemble of non-orthogonal states than from an ensemble of orthogonal ones.

It is of interest to find the best measurement to distinguish non-orthogonal states ρ_i with the smallest possible failure probability. We study this state discrimination problem in this chapter. This is a quite important issue in quantum cryptography and in quantum communication in general. As emphasized in the introduction of this article, we aim at explaining some typical questions, providing examples, and establishing basic general results that will be used in the next chapters, rather than giving a full account on the subject. We refer the reader to the review articles [58, 37, 36] for more complete presentations. Measurements for distinguishing quantum states can also be optimized using other criteria than the minimal probability of equivocation. For instance, one can try to maximize the mutual information between the initial distribution of the state ensemble and the distribution of the measurement outcomes. This optimization problem, which plays an important role in the transmission of information in quantum channels, is briefly discussed at the end of this chapter.

Before entering into the detail of the theory, let us make a philosophical remark concerning the quantum-classical differences. Let us inquire about the quantum analog of the celebrated experiment in classical probability which consists of picking up randomly colored balls contained in an urn. In quantum mechanics, the readout of the system's state (the color of the ball in the classical analogy) is performed by a measurement perturbing the system. If the urn contains an ensemble of non-orthogonal states, we have just seen above that there is no way to identify with certainty which state from the ensemble has been picked up. Therefore, the starting assumption that the color of the ball is known once it has been extracted from the urn is not fulfilled in the quantum world and identifying these colors is already a non-trivial task!

6.1 Discriminating quantum states drawn from a given ensemble

We review in this section two strategies for discriminating non-orthogonal states, known as the ambiguous and unambiguous state discriminations. Let us consider an ensemble $\{\rho_i, \eta_i\}_{i=1}^m$ of states ρ_i with prior probabilities η_i. For instance, the ρ_i can be some states of the electromagnetic field encoding m symbols of a given alphabet,

[1] No unitary evolution on a system S initially in state $|\psi\rangle$ and a register R initially in state $|\phi\rangle$ can transform $|\Psi\rangle = |\psi\rangle|\phi\rangle$ into $|\Psi'\rangle = |\psi\rangle|\psi\rangle$ for any $|\psi\rangle$ belonging to a set of distinct non-orthogonal states, e.g. $|\psi\rangle \in \{|\psi_1\rangle, |\psi_2\rangle\}$. Actually, the scalar products $\langle\Psi_1|\Psi_2\rangle = \langle\psi_1|\psi_2\rangle$ and $\langle\Psi_1'|\Psi_2'\rangle = \langle\psi_1|\psi_2\rangle^2$ are different if $\langle\psi_1|\psi_2\rangle \neq 0, 1$. More generally, the no-cloning theorem tells us that one cannot duplicate unknown states by using any (not necessarily unitary) quantum evolution, except when these states pertain to a family of orthogonal states [28].

the ith symbol occurring with frequency η_i. In order to send a message, a sender prepares random states drawn from the ensemble and gives them to a receiver. To decode the message the latter must identify these states by performing measurements. He wants to find the measurement that minimizes the failure probability.

A first strategy, called *ambiguous (or minimal error) quantum state discrimination*, consists in looking for a generalized measurement with m outcomes yielding the maximal success probability $P_{\rm S} = \sum_i \eta_i p_{i|i}$, $p_{i|i}$ being the probability of the measurement outcome i given that the state is ρ_i. Here, the number of possible outcomes is chosen to be equal to the number of states in the ensemble. The conditional probability of the outcome j given the state ρ_i is (see Sec. 5.3)

$$p_{j|i} = \mathrm{tr}(M_j \rho_i) \tag{6.1}$$

so that the maximal success probability reads

$$P_{\rm S}^{\rm opt}(\{\rho_i, \eta_i\}) = \max_{\text{POVM } \{M_i\}} \left\{ \sum_{i=1}^{m} \eta_i \,\mathrm{tr}(M_i \rho_i) \right\}, \tag{6.2}$$

where the maximum is over all POVMs $\{M_i\}_{i=1}^m$.

A second strategy consists in seeking for a generalized measurement with $(m+1)$ outcomes enabling to identify perfectly each state ρ_i, but such that one of the outcomes leads to an inconclusive result. This strategy, originally proposed by Ivanovic [136] and further investigated by Dieks and Peres [78, 191], is called *unambiguous quantum state discrimination*. In other words, if the measurement outcome is $j \in \{1, \ldots, m\}$ then the receiver is certain that the state is ρ_j, whereas if $j = 0$ he does not know. This means that $p_{j|i} = p_{i|i}\delta_{ij}$ with $p_{i|i} > 0$, for any $i, j = 1, \ldots, m$. The probability of occurrence of the inconclusive outcome, $P_0 = \sum_i \eta_i p_{0|i}$, must be minimized. Since $p_{0|i} = 1 - p_{i|i}$, the success probability is obtained from the same formula (6.2) as for ambiguous discrimination, but with a maximum over all POVMs $\{M_j\}_{j=0}^m$ such that $\mathrm{tr}(M_j \rho_i) = p_{i|i}\delta_{ij}$ for $j \neq 0$. For pure states $\rho_i = |\psi_i\rangle\langle\psi_i|$, the rank-one measurement operators M_j satisfying this condition are

$$M_j = \frac{p_{j|j}}{|\langle \psi_j^* | \psi_j \rangle|^2} |\psi_j^*\rangle\langle\psi_j^*| \quad , \quad j = 1, \ldots, m \,, \tag{6.3}$$

with the dual normalized vectors $|\psi_j^*\rangle$ defined by $\langle \psi_j^* | \psi_i \rangle = \delta_{ij} \langle \psi_i^* | \psi_i \rangle$. The remaining problem is to find the values of the probabilities $p_{j|j}$ which maximize the success probability (6.2) under the constraint that $\{M_j\}_{j=0}^m$ is a POVM, that is,

$$M_0 = 1 - \sum_{j=1}^{m} \frac{p_{j|j}}{|\langle \psi_j^* | \psi_j \rangle|^2} |\psi_j^*\rangle\langle\psi_j^*| \;\geq\; 0 \,. \tag{6.4}$$

This is a non-trivial problem, which has been solved so far in particular cases only. Upper and lower bounds on the maximal success probability can be found in terms of the scalar products $\langle \psi_i | \psi_j \rangle$ (see e.g. [37]).

It is worth noting that unambiguous discrimination is not always possible. For instance, a pure state ensemble $\{|\psi_i\rangle, \eta_i\}$ with linearly dependent vectors $|\psi_i\rangle$ cannot be discriminated unambiguously [57]. Indeed, assume that $|\psi_{i_0}\rangle$ is a linear combination of the other states $|\psi_i\rangle$. Together with the no-error condition $p_{j|i} = p_{i|i}\delta_{ij}$, which is equivalent to $|\psi_i\rangle \in \ker M_j$ for any $j \notin \{0, i\}$, this means that $|\psi_{i_0}\rangle \in \ker(M_{i_0})$ and thus $p_{i_0|i_0} = 0$, in contradiction with the requirement $p_{i_0|i_0} > 0$. The same argument shows that one cannot discriminate unambiguously an ensemble of mixed states $\{\rho_i, \eta_i\}$ such that one state ρ_{i_0} has its support $\mathrm{ran}(\rho_{i_0})$ contained in the sum of the supports of the other states.

Ambiguous and unambiguous quantum state discriminations have many applications. For instance, the discrimination of two non-orthogonal states plays a central role in the quantum cryptography protocol proposed by Bennett in 1992 to distribute a secrete key between two parties [32]. We will not elaborate further on these applications. Let us also mention that other optimization schemes than those discussed above have been worked out [37, 36]. State discriminations have been implemented experimentally by using polarized photons in pure states (see [62] and references therein) and, more recently, in mixed states [171].

6.2 Ambiguous and unambiguous discriminations of two states

6.2.1 Ambiguous discrimination

The simplest example of ambiguous discrimination is the case of $m = 2$ states ρ_1 and ρ_2. Then the optimal success probability and measurement are easy to determine [119]. One starts by writing the measurement operator M_2 as $1 - M_1$ in the expression of the success probability,

$$P_{\rm S,a}^{\{M_i\}}(\{\rho_i, \eta_i\}) = \eta_1 \,\mathrm{tr}(M_1 \rho_1) + \eta_2 \,\mathrm{tr}(M_2 \rho_2) = \frac{1}{2}\big(1 - \mathrm{tr}\,\Lambda\big) + \mathrm{tr}(M_1 \Lambda) \tag{6.5}$$

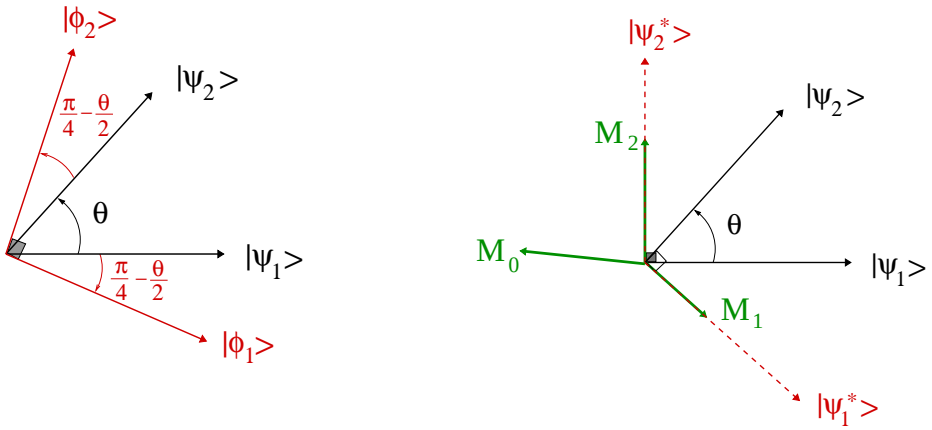

Figure 6.1: Optimal measurement $\{M_i^{\rm opt}\}$ in the discrimination of two non-orthogonal pure states $|\psi_1\rangle$ and $|\psi_2\rangle$ with equal prior probabilities $\eta_i = 1/2$. (a) For ambiguous discrimination, $\{M_i^{\rm opt}\}$ is the von Neumann measurement in the two orthogonal states $|\phi_1\rangle$ and $|\phi_2\rangle$ with $|\langle\phi_i|\psi_i\rangle| = \cos(\frac{\pi}{4}-\frac{\theta}{2})$, that is, it is the least square measurement associated to $\{|\psi_i\rangle, \eta_i\}$ (b) For unambiguous discrimination, if the maximal prior probability $\eta_{\max}$ is larger than $q_1 = 1/(1+\cos^2\theta)$, then the von Neumann measurement in the orthonormal basis $\{|\psi_1\rangle, |\psi_2^*\rangle\}$ (if $\eta_{\max} = \eta_2 > \eta_1$) or $\{|\psi_2\rangle, |\psi_1^*\rangle\}$ (if $\eta_{\max} = \eta_1 > \eta_2$) indicated by the red dashed vectors yields the smallest failure probability. Failure occurs when the outcome corresponds to the first vector in these two bases (inconclusive result). If $1-q_1 < \eta_1 < q_1$, a smaller failure probability is obtained by using the generalized measurement with rank-one operators M_i indicated schematically by the green vectors.

with $\Lambda = \eta_1\rho_1 - \eta_2\rho_2$. The maximum of $\mathrm{tr}(M_1\Lambda)$ over all M_1 satisfying $0 \le M_1 \le 1$ is achieved when M_1 is the spectral projector Π_1 associated to the positive eigenvalues $\lambda_1 \ge \cdots \ge \lambda_p > 0$ of the Hermitian matrix Λ. Consequently, the maximal success probability is given by the Helstrom formula

$$P_{\rm S,a}^{\rm opt}(\{\rho_i,\eta_i\}) = \frac{1}{2}\big(1+\mathrm{tr}\,|\Lambda|\big) \quad , \quad \Lambda = \eta_1\rho_1 - \eta_2\rho_2\ . \tag{6.6}$$

The optimal measurement is a von Neumann measurement $\{\Pi_1^{\rm opt}, 1-\Pi_1^{\rm opt}\}$ with $\Pi_1^{\rm opt}$ the projector onto the support of $\Lambda_+ = (\Lambda+|\Lambda|)/2$. If $\Lambda \ge 0$ the optimal measurement is $\{\Pi_1^{\rm opt}=1, \Pi_2^{\rm opt}=0\}$, meaning that no measurement can outperform the simple guess that the state is ρ_1 (a similar statement holds for ρ_2 if $\Lambda \le 0$). For pure states $\rho_i = |\psi_i\rangle\langle\psi_i|$, (6.6) reduces to

$$P_{\rm S,a}^{\rm opt}(\{|\psi_i\rangle,\eta_i\}) = \frac{1}{2}\Big(1+\sqrt{1-4\eta_1\eta_2|\langle\psi_1|\psi_2\rangle|^2}\Big) \tag{6.7}$$

and the optimal measurement consists of the rank-one eigenprojectors of Λ for the positive and negative eigenvalues. When $\eta_1 = \eta_2$, these are the projections onto the two orthogonal subspaces placed symmetrically with respect to $|\psi_1\rangle$ and $|\psi_2\rangle$, as represented in Fig. 6.1.

6.2.2 Unambiguous discrimination of two pure states

The power of generalized measurements is illustrated in the unambiguous discrimination of two pure states $|\psi_1\rangle$ and $|\psi_2\rangle$. Indeed, we will show that such measurements enable to distinguish quantum states better than von Neumann measurements[2]. Clearly, the Hilbert space $\mathcal{H}$ can be restricted to its two-dimensional subspace spanned by $|\psi_1\rangle$ and $|\psi_2\rangle$. The unambiguity condition implies $|\psi_1\rangle \in \ker M_2$ and $|\psi_2\rangle \in \ker M_1$, so that the measurement operators M_1 and M_2 are of rank one and given by (6.3). We can already observe at this point that the number of outcomes is larger than the space dimension, so that the unambiguous discrimination strategy cannot be realized with a von Neumann measurement.

The optimal success probability is given by [139]

$$P_{\rm S,u}^{\rm opt}(\{|\psi_i\rangle,\eta_i\}) = \begin{cases} 1-2\sqrt{\eta_1\eta_2}|\langle\psi_1|\psi_2\rangle| & \text{if } 1-q_1 \le \eta_1 \le q_1 \\ \eta_{\max}(1-|\langle\psi_1|\psi_2\rangle|^2) & \text{if } \eta_{\max} \ge q_1 \end{cases} \tag{6.8}$$

[2]This can be considered as the main physical motivation to introduce generalized measurements [194].

with $\eta_{\max} = \max\{\eta_1, \eta_2\}$ and $q_1 = 1/(1 + |\langle\psi_1|\psi_2\rangle|^2)$. It is instructive to establish this formula by using the Neumark extension theorem [37]. Thanks to Theorem 5.3.1, one can represent $\{M_j\}$ as a von Neumann measurement on the larger space $\mathcal{H} \otimes \mathcal{H}_{\mathsf{E}}$, with $\mathcal{H}_{\mathsf{E}} \simeq \mathbb{C}^3$. Let $\{A_j\}_{j=0}^2$ be the Kraus operators for the measurement and $|\epsilon_0\rangle$, U, and π_j^{E} be as in this theorem. We may assume that $\pi_j^{\mathsf{E}} = |j\rangle\langle j|$ are of rank one, where $\{|j\rangle\}_{j=0}^2$ is an orthonormal basis of $\mathcal{H}_{\mathsf{E}}$ (see the proof of Theorem 5.3.1). One writes

$$|\Psi_i'\rangle = U|\psi_i\rangle|\epsilon_0\rangle = \sum_{j=0}^{2} \sqrt{p_{j|i}}|\varphi_{j|i}\rangle|j\rangle \tag{6.9}$$

for $i = 1, 2$, where $\sqrt{p_{j|i}}|\varphi_{j|i}\rangle = \langle j|\Psi_i'\rangle \in \mathcal{H}$ are in general non-orthogonal for distinct j's and $\|\varphi_{j|i}\| = 1$. By (5.14) and (5.15) the unnormalized post-measurement states are $\widetilde{\rho}_{j|i} = \langle j|\Psi_i'\rangle\langle\Psi_i'|j\rangle = p_{j|i}|\varphi_{j|i}\rangle\langle\varphi_{j|i}|$, hence $p_{j|i}$ and $|\varphi_{j|i}\rangle$ can be interpreted as the probability of outcome j and the corresponding conditional state for the input state $|\psi_i\rangle$. Since we require $p_{2|1} = p_{1|2} = 0$, the unitarity of U imposes the conditions $p_{0|i} = 1 - p_{i|i}$ and $\langle\Psi_1'|\Psi_2'\rangle = \sqrt{p_{0|1}p_{0|2}}\langle\varphi_{0|1}|\varphi_{0|2}\rangle = \langle\psi_1|\psi_2\rangle$. The last relation implies that the probabilities $p_{0|i}$ satisfy

$$p_{0|1}p_{0|2} \geq p_{0|1}p_{0|2}|\langle\varphi_{0|1}|\varphi_{0|2}\rangle|^2 = \cos^2\theta\,, \tag{6.10}$$

where we have set $\cos\theta = |\langle\psi_1|\psi_2\rangle|$. Note that this bound could have been obtained directly from (6.4), which is easy to solve since we are dealing here with 2×2 matrices [37].

In order to maximize the success probability $P_{\mathrm{S}} = \sum_i \eta_i p_{i|i} = 1 - \sum_i \eta_i p_{0|i}$, we are looking for the smallest possible $p_{0|1}$ and $p_{0|2}$. For such $p_{0|i}$'s the inequality (6.10) is an equality. Assuming $\cos\theta > 0$, this holds whenever $|\varphi_{0|2}\rangle = \mathrm{e}^{\mathrm{i}\delta}|\varphi_{0|1}\rangle$ with $\delta = \arg\langle\psi_1|\psi_2\rangle$. Accordingly, the conditional post-measurement state for the inconclusive outcome is the same irrespective of the input state $|\psi_i\rangle$. This is physically meaningful since if this post-measurement state was depending on $|\psi_i\rangle$ then one could perform a new measurement on it to increase further the success probability. In summary, for the optimal measurement one has

$$|\Psi_i'\rangle = \sqrt{p_{i|i}}|\varphi_{i|i}\rangle|i\rangle + \sqrt{p_{0|i}}\,\mathrm{e}^{\mathrm{i}\delta_i}|\Phi_0\rangle \tag{6.11}$$

with $|\Phi_0\rangle = |\varphi_{0|1}\rangle|0\rangle$ and $\delta_1 = 0$, $\delta_2 = \delta$.

The failure probability

$$P_0 = \eta_1 p_{0|1} + \eta_2 \frac{\cos^2\theta}{p_{0|1}} \tag{6.12}$$

is easy to minimize as a function of $p_{0|1}$. The minimum is achieved for $p_{0|1}^{\mathrm{opt}} = \sqrt{\eta_2/\eta_1}\cos\theta$ and is equal to $P_0^{\mathrm{opt}} = 2\sqrt{\eta_1\eta_2}\cos\theta$. This yields the upper expression in (6.8). The restrictions on the values of η_1 come from the conditions $p_{0|1}^{\mathrm{opt}} \leq 1$ and $p_{0|2}^{\mathrm{opt}} \leq 1$. When $\eta_1 \leq 1 - q_1$, the minimum is achieved for $p_{0|1}^{\mathrm{opt}} = 1$ and $p_{0|2}^{\mathrm{opt}} = \cos^2\theta$, i.e., $p_{1|1}^{\mathrm{opt}} = 0$ and $p_{2|2}^{\mathrm{opt}} = \sin^2\theta$. In such a case only the state $|\psi_2\rangle$ can be identified with certainty, as $|\psi_1\rangle$ always produces an inconclusive outcome. Strictly speaking this does not correspond to an unambiguous discrimination. One can nevertheless determine the optimal measurement, characterized by $M_1^{\mathrm{opt}} = 0$ and by two orthogonal projectors $M_2^{\mathrm{opt}} = |\psi_2^*\rangle\langle\psi_2^*|$ and $M_0^{\mathrm{opt}} = |\psi_1\rangle\langle\psi_1|$, see (6.3). A similar statement holds when $\eta_1 \geq q_1$ by exchanging the indices 1 and 2. The corresponding success probability is given by the lower expression in (6.8).

These results are summarized in Fig. 6.1. As claimed above, when $1-q_1 < \eta_1 < q_1$ generalized measurements, obtained via a coupling of the system with an ancilla and a measurement on the latter, do better in decoding the message than a von Neumann measurement performed directly on the system.

6.2.3 Unambiguous discrimination of two mixed states

Let us now turn to the case of two mixed states ρ_1 and ρ_2. Such states cannot be unambiguously discriminated when $\operatorname{ran}\rho_1$ is contained in $\operatorname{ran}\rho_2$ or vice versa. By the unambiguity condition, $\operatorname{ran} M_1 \subset \ker\rho_2$ and $\operatorname{ran} M_2 \subset \ker\rho_1$. A trivial situation is when $\ker\rho_1 \perp \ker\rho_2$, in which case the optimal POVM is the von Neumann measurement with M_1 and M_2 equal to the projectors on $\ker\rho_2$ and $\ker\rho_1$, respectively. Then the minimal failure probability is $P_0^{\mathrm{opt}} = \mathrm{tr}[(\eta_1\rho_1 + \eta_2\rho_2)\Pi_0]$, Π_0 being the projector onto $\operatorname{ran}\rho_1 \cap \operatorname{ran}\rho_2$. One can as before restrict the Hilbert space so that $\operatorname{ran}\rho_1 + \operatorname{ran}\rho_2 = \mathcal{H}$. If $\operatorname{ran}\rho_1$ and $\operatorname{ran}\rho_2$ have co-dimension one in $\mathcal{H}$, then M_1 and M_2 are of rank one and take the form (6.3) with $|\psi_1^*\rangle \in \ker\rho_2$, $|\psi_2^*\rangle \in \ker\rho_1$, and $|\langle\psi_i^*|\psi_i\rangle|^2$ replaced

by $R_i = \langle\psi_i^*|\rho_i|\psi_i^*\rangle$. A simple generalization of (6.8) then yields [204]

$$P_{\rm S,u}^{\rm opt}(\{\rho_i,\eta_i\}) = P_{\rm S}^{\rm opt}(R_i,\eta_i) \equiv \begin{cases} \dfrac{\eta_1 R_1 + \eta_2 R_2 - 2\sqrt{\eta_1\eta_2 R_1 R_2}\cos\theta}{\sin^2\theta} & \text{if } \cos^2\theta < \min\Big\{\dfrac{\eta_1 R_1}{\eta_2 R_2}, \dfrac{\eta_2 R_2}{\eta_1 R_1}\Big\} \\ \max\{\eta_1 R_1, \eta_2 R_2\} & \text{otherwise} \end{cases} \tag{6.13}$$

with $\cos\theta = |\langle\psi_1^*|\psi_2^*\rangle|$. For kernels of dimensions $d_2 \geq d_1 > 1$, by a standard linear algebra argument one can construct two orthonormal bases $\{|\psi_{2k}^*\rangle\}_{k=1}^{d_1}$ of $\ker\rho_1$ and $\{|\psi_{1k}^*\rangle\}_{k=1}^{d_2}$ of $\ker\rho_2$ such that $\langle\psi_{1k}^*|\psi_{2l}^*\rangle = \delta_{kl}\cos\theta_k$, with $\theta_k \in [0,\pi/2]$. Let us take $M_i = \sum_k M_{ik}$ for $i = 1,2$, with $M_{ik} = m_{ik}|\psi_{ik}^*\rangle\langle\psi_{ik}^*|$. Optimizing $P_{\rm S,u}^{\{M_i\}}$ over the non-negative numbers m_{ik} under the constraint $1 - M_1 - M_2 \geq 0$ reduces to the optimization problem for rank-one measurement operators studied before (in fact, this constraint is equivalent to $1 - M_{1k} - M_{2k} \geq 0$ for $k = 1,\ldots,d_1$ and $1 - M_{1k} \geq 0$ for $d_1 < k \leq d_2$). This gives the lower bound [204]

$$P_{\rm S,u}^{\rm opt}(\{\rho_i,\eta_i\}) \geq \sum_{k=1}^{d_1} P_{\rm S}^{\rm opt}(R_{ik},\eta_i) + \eta_1 \sum_{d_1<k\leq d_2} R_{1k} \quad \text{with } R_{ik} = \langle\psi_{ik}^*|\rho_i|\psi_{ik}^*\rangle. \tag{6.14}$$

An upper bound can be obtained in terms of the fidelity between the states ρ_1 and ρ_2 defined by $F(\rho_1,\rho_2) = (\mathrm{tr}(|\sqrt{\rho_1}\sqrt{\rho_2}|))^2$ (see Proposition 6.5.2 and Remark 8.4.4 below) [204],

$$P_{\rm S,u}^{\rm opt}(\{\rho_i,\eta_i\}) \leq \begin{cases} 1 - 2\sqrt{\eta_1\eta_2 F(\rho_1,\rho_2)} & \text{if } F(\rho_1,\rho_2) < \dfrac{\eta_{\min}}{\eta_{\max}} \\ \eta_{\max}(1 - F(\rho_1,\rho_2)) & \text{otherwise.} \end{cases} \tag{6.15}$$

A nice application of two mixed state discrimination is the state comparison problem [27]. Consider two independent copies of a given system, the state of which is drawn from the pure state ensemble $\{|\psi_i\rangle, 1/2\}_{i=1,2}$. One would like to decide with the help of an appropriate measurement if the two copies are in the same state or not, without further information on the actual state of each copies. If $|\psi_1\rangle$ and $|\psi_2\rangle$ are not orthogonal, this can only be done with a probability of success $P_{\rm S,comp} < 1$. This amounts to discriminate the two mixed states

$$\begin{aligned} \rho_{\rm eq} &= \frac{1}{2}|\psi_1\otimes\psi_1\rangle\langle\psi_1\otimes\psi_1| + \frac{1}{2}|\psi_2\otimes\psi_2\rangle\langle\psi_2\otimes\psi_2| \\ \rho_{\rm diff} &= \frac{1}{2}|\psi_1\otimes\psi_2\rangle\langle\psi_1\otimes\psi_2| + \frac{1}{2}|\psi_2\otimes\psi_1\rangle\langle\psi_2\otimes\psi_1| \,. \end{aligned} \tag{6.16}$$

It is shown in [204] that for such mixed states of rank two, the lower and upper bounds in (6.14) and (6.15) coincide. A simple calculation (see Remark 8.4.4 below) then gives the optimal success probability [27]

$$P_{\rm S,comp}^{\rm opt} = 1 - |\langle\psi_1|\psi_2\rangle| \,. \tag{6.17}$$

6.3 Discrimination with least square measurements

How well does the least square measurement (Sec. 5.5.3) in discriminating ambiguously quantum states? More precisely, let

$$P_{\rm S,a}^{\rm lsm}(\{\rho_i,\eta_i\}) = \sum_i \eta_i \,\mathrm{tr}(\rho_i M_i^{\rm lsm}) \tag{6.18}$$

be the success probability in discriminating the states ρ_i by performing the least square measurement $\{M_i^{\rm lsm}\}$ associated to $\{\rho_i,\eta_i\}$. We would like to compare $P_{\rm S,a}^{\rm lsm}$ with the optimal success probability.

Let us first observe that if $\rho_i = \mathcal{M}(|i\rangle\langle i|)$, $\mathcal{M}$ being a quantum operation on $\mathcal{B}(\mathcal{H})$ and $\{|i\rangle\}_{i=1}^n$ a fixed orthonormal basis of $\mathcal{H}$, then $P_{\rm S,a}(\{\rho_i,\eta_i\})$ is related to the entanglement fidelity defined in Sec. 5.5.2. Recall that any ensemble $\{\rho_i,\eta_i\}_{i=1}^m$ with $m \leq n$ states can be obtained in this way from an operation $\mathcal{M} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ (since $m \leq n$ we can identify here the pointer space $\mathcal{H}_{\rm P}$ with a subspace of $\mathcal{H}$, see Sec. 5.4). To establish the relation with the average fidelity (5.28), consider a POVM $\{M_i\}_{i=1}^m$ with m measurement operators and let us associate to it the quantum operation $\mathcal{R}$ on $\mathcal{B}(\mathcal{H})$ defined by $\mathcal{R}^*(|i\rangle\langle j|) = M_i\delta_{ij}$. Then

$$P_{\rm S,a}^{\{M_i\}}\big(\{\rho_i,\eta_i\}\big) = \sum_{i=1}^m \eta_i \,\mathrm{tr}[\mathcal{R}^*(|i\rangle\langle i|)\rho_i] = \sum_{i=1}^m \eta_i \langle i|\mathcal{R}\circ\mathcal{M}(|i\rangle\langle i|)|i\rangle = \overline{F}_{\rm e}\big(\{|i\rangle,\eta_i\},\mathcal{R}\circ\mathcal{M}\big) \tag{6.19}$$

thanks to the equality of the entanglement fidelity with the input-output fidelity for pure states. In view of the one-to-one correspondence between POVMs with $m \leq n$ operators and quantum operations on $\mathcal{B}(\mathcal{H})$ we obtain the following relation between $P_{\mathrm{S,\,a}}^{\mathrm{opt}}$ and the maximal fidelity over all recovery operations $\mathcal{R}$ on $\mathcal{B}(\mathcal{H})$:

$$P_{\mathrm{S,\,a}}^{\mathrm{opt}}(\{\rho_i,\eta_i\}_{i=1}^m) = \max_{\mathcal{R}}\{\overline{F}_{\mathrm{e}}(\{|i\rangle,\eta_i\}_{i=1}^m,\mathcal{R}\circ\mathcal{M})\} \quad , \quad m \leq n \; . \tag{6.20}$$

Furthermore, the optimal measurement operators are given in terms of the optimal recovery operation $\mathcal{R}^{\mathrm{opt}}$ by $M_i^{\mathrm{opt}} = (\mathcal{R}^{\mathrm{opt}})^*(|i\rangle\langle i|)$. According to Proposition 5.5.3, taking $\mathcal{R}$ to be the transpose operation $\mathcal{R}_{\mathcal{M},\rho_{\mathrm{in}}}$ of $\mathcal{M}$ for the state $\rho_{\mathrm{in}} = \sum_i \eta_i |i\rangle\langle i|$ gives an entanglement fidelity larger than the square of the right-hand side of (6.20). But the measurement associated to $\mathcal{R}_{\mathcal{M},\rho_{\mathrm{in}}}$ is the least square measurement, i.e., $M_i^{\mathrm{lsm}} = \mathcal{R}^*_{\mathcal{M},\rho_{\mathrm{in}}}(|i\rangle\langle i|)$ (see Sec. 5.5.3). As a result, Proposition 5.5.3 yields the following inequality.

Corollary 6.3.1. *If $m \leq n = \dim \mathcal{H}$, then*

$$P_{\mathrm{S,a}}^{\mathrm{opt}}(\{\rho_i,\eta_i\}_{i=1}^m) \leq \left(P_{\mathrm{S,a}}^{\mathrm{lsm}}(\{\rho_i,\eta_i\}_{i=1}^m)\right)^{\frac{1}{2}} \; . \tag{6.21}$$

Thus, if the error probability for discriminating $\{\rho_i,\eta_i\}$ using the least square measurement is small, then it is at most twice the minimal error probability $P_{\mathrm{err,a}}^{\mathrm{opt}} = 1 - P_{\mathrm{S,a}}^{\mathrm{opt}}$, up to a small correction of the order of $(P_{\mathrm{S,a}}^{\mathrm{lsm}})^2$. Small error probabilities occur for almost orthogonal states. Therefore, for such states least square measurements are nearly optimal [114, 29].

It is worth mentioning that least square measurements are also asymptotically optimal for discriminating ambiguously equiprobable linearly independent pure states [125]. In addition, they optimally discriminate equiprobable states drawn from a symmetric ensemble, like for instance the states $\rho_i = U^{i-1}\rho_1(U^{i-1})^*$ related between themselves through conjugations by powers of a single unitary operator U satisfying $U^m = \pm 1$ (see [24, 26, 61, 83] and references therein). Necessary and sufficient conditions for the optimality of least square measurements in state discrimination have been investigated in [86, 207].

6.4 General results on ambiguous discrimination

Let $\{\rho_i,\eta_i\}_{i=1}^m$ be an ensemble of m states of a system with a n-dimensional Hilbert space $\mathcal{H}$. Hereafter we assume that $\eta_i > 0$ for all $i = 1,\ldots,m$, so that m is the actual number of states to discriminate. We denote by $\widetilde{\rho}_i = \eta_i\rho_i$ the unnormalized states with trace equal to the prior probability η_i. To shorten notation, the dependence of the success probability P_{S} on the ensemble is not written explicitly. The following proposition contains one of the few results in ambiguous discrimination applying to arbitrary ensembles.

Proposition 6.4.1. [123, 269, 84] *The optimal success probability in ambiguous state discrimination is given by*

$$P_{\mathrm{S,a}}^{\mathrm{opt}} = \inf_{\Upsilon \geq \widetilde{\rho}_i}\{\mathrm{tr}(\Upsilon)\} \; , \tag{6.22}$$

where the infimum is over all self-adjoint operators Υ satisfying $\Upsilon \geq \widetilde{\rho}_i$ for any $i = 1,\ldots,m$. Moreover, the POVM $\{M_i^{\mathrm{opt}}\}_{i=1}^m$ is optimal if and only if the operator $\Upsilon^{\mathrm{opt}} = \sum_i \widetilde{\rho}_i M_i^{\mathrm{opt}}$ satisfies the two conditions

(i) Υ^{opt} is self-adjoint;

(ii) $\Upsilon^{\mathrm{opt}} \geq \widetilde{\rho}_i$ for any $i = 1,\ldots,m$.

In such a case, the infimum in the right-hand side of (6.22) is attained for $\Upsilon = \Upsilon^{\mathrm{opt}}$.

The fact that (ii) is sufficient to ensure the optimality of $\{M_i^{\mathrm{opt}}\}$ is obvious from the relation

$$P_{\mathrm{S,a}}^{\mathrm{opt}} - P_{\mathrm{S,a}}^{\{M_i\}} = \sum_{i=1}^m \mathrm{tr}[(\Upsilon^{\mathrm{opt}} - \widetilde{\rho}_i)M_i] \; . \tag{6.23}$$

The necessary and sufficient conditions (i) and (ii) are due to Holevo [123], who derived them by considering a specific one-parameter family $\{M_i(\varepsilon)\}$ of POVMs such that $M_i(0) = M_i^{\mathrm{opt}}$ and by exploiting the fact that $\partial P_{\mathrm{S,a}}^{\{M_i(\varepsilon)\}}/\partial\varepsilon = 0$ for $\varepsilon = 0$ (see [119], chapter 4). Yuen, Kennedy, and Lax [269] proposed another derivation based on a duality argument in vector space optimization. We shall present below the related proof of Eldar, Megretski and Verghese [84].

Let us note that (i) and (ii) imply

$$(\Upsilon^{\rm opt} - \widetilde{\rho}_i)M_i^{\rm opt} = M_i^{\rm opt}(\Upsilon^{\rm opt} - \widetilde{\rho}_i) = 0 \quad , \quad i = 1, \ldots, m \; . \tag{6.24}$$

In fact, since $\sum_i {\rm tr}[(\Upsilon^{\rm opt} - \widetilde{\rho}_i)M_i^{\rm opt}] = 0$ and $\Upsilon^{\rm opt} - \widetilde{\rho}_i \geq 0$ by (ii), one deduces that $|(\Upsilon^{\rm opt} - \widetilde{\rho}_i)^{1/2}(M_i^{\rm opt})^{1/2}|^2 = 0$ (recall that $A \geq 0$ and ${\rm tr}(A) = 0$ imply $A = 0$). One concludes from this equality that $(\Upsilon^{\rm opt} - \widetilde{\rho}_i)M_i^{\rm opt} = 0$. It is easy to see by eliminating $\Upsilon^{\rm opt}$ that (6.24) is equivalent to

$$M_i^{\rm opt}(\widetilde{\rho}_i - \widetilde{\rho}_j)M_j^{\rm opt} = 0 \quad , \quad i, j = 1, \ldots, m \; . \tag{6.25}$$

The condition (6.25) automatically implies that $\Upsilon^{\rm opt}$ is self-adjoint. Hence a necessary and sufficient condition for $\{M_i^{\rm opt}\}$ to be optimal is given by conditions (ii) and (6.25).

Except in special cases such as ensembles of equiprobable states related by a symmetry [24, 26, 83, 61], it is difficult in practice to obtain the optimal measurement and success probability from the above necessary and sufficient conditions. Nevertheless, the formulas (6.22) and (6.24) are helpful for computing these quantities numerically. For indeed, the minimization task in (6.22) is simpler than the maximization in (6.2) and can be solved efficiently with the help of convex semidefinite programs [84].

Remark 6.4.2. *The necessary and sufficient optimality conditions (i) and (ii) of Proposition 6.4.1 are equivalent to the following condition:* ${\rm Re}\,\Upsilon^{\rm opt} \geq \widetilde{\rho}_i$ *for any* $i = 1, \ldots, m$, *see [124].*

Proof. The main idea is to show that the minimization problem in (6.22) is dual to the maximization problem in (6.2). More precisely, there exists a convex set $\Gamma \subset \mathcal{B}(\mathcal{H})_{\rm s.a.}$ such that

$$P_{\rm S,a}^{\{M_i\}} \leq {\rm tr}(\Upsilon) \quad , \quad \forall\, \{M_i\} \text{ POVM}, \; \forall\, \Upsilon \in \Gamma \; , \tag{6.26}$$

and the maximum of the left-hand member is equal to the minimum of the right-hand member, i.e., $P_{\rm S,a}^{\rm opt} = \min_{\Upsilon \in \Gamma} {\rm tr}(\Upsilon)$. The set Γ is defined by

$$\Gamma = \big\{ \Upsilon \in \mathcal{B}(\mathcal{H})_{\rm s.a.} \, ; \, \Upsilon \geq \widetilde{\rho}_i \, , \, i = 1, \ldots, m \big\} \; . \tag{6.27}$$

Then ${\rm tr}(\Upsilon) - P_{\rm S,a}^{\{M_i\}} = \sum_i {\rm tr}[(\Upsilon - \widetilde{\rho}_i)M_i] \geq 0$ for any $\Upsilon \in \Gamma$, so that (6.26) holds true. Let us now define the following convex subset Ω of the real vector space $\mathcal{B}(\mathcal{H})_{\rm s.a.} \times \mathbb{R}$:

$$(B, x) \in \Omega \quad \Leftrightarrow \quad B = \sum_{i=1}^m B_i - 1 \; , \; x = r - \sum_{i=1}^m {\rm tr}(B_i \widetilde{\rho}_i) \;\; \text{with} \;\; B_i \geq 0 \;\; \text{and} \;\; r > P_{\rm S,a}^{\rm opt} \; . \tag{6.28}$$

This space is endowed with the scalar product $\langle (B, x) \, , \, (C, y) \rangle = {\rm tr}(BC) + xy$. Since Ω is convex and does not contain $(0, 0)$, by the separating hyperplane theorem one can find a non-vanishing vector $(\Upsilon_a, a) \in \mathcal{B}(\mathcal{H})_{\rm s.a.} \times \mathbb{R}$ such that $\langle (\Upsilon_a, a) \, , \, (B, x) \rangle \geq 0$ for any $(B, x) \in \Omega$, that is

$${\rm tr}\Big[\Upsilon_a \Big(\sum_{i=1}^m B_i - 1 \Big)\Big] + a\Big(r - \sum_{i=1}^m {\rm tr}(B_i \widetilde{\rho}_i) \Big) \geq 0 \; . \tag{6.29}$$

Taking $B_i = t|\varphi\rangle\langle\varphi|$ if $i = k$ and zero otherwise, with $|\varphi\rangle \in \mathcal{H}$ and $t > 0$, and letting $t \to \infty$, we obtain $\langle\varphi|\Upsilon_a|\varphi\rangle - a\langle\varphi|\widetilde{\rho}_k|\varphi\rangle \geq 0$. But $|\varphi\rangle$ and k are arbitrary, hence

$$\Upsilon_a \geq a\widetilde{\rho}_i \quad , \quad i = 1, \ldots, m \; . \tag{6.30}$$

Similarly, taking $B_i = 0$ for all i and $r \to P_{\rm S,a}^{\rm opt}$, (6.29) yields

$$a P_{\rm S,a}^{\rm opt} \geq {\rm tr}(\Upsilon_a) \; . \tag{6.31}$$

From the same choice of B_i and $r \to \infty$ one gets $a \geq 0$. If $a = 0$ then $\Upsilon_a \geq 0$ and ${\rm tr}(\Upsilon_a) = 0$ by (6.30) and (6.31). This would imply $\Upsilon_a = 0$, in contradiction with $(\Upsilon_a, a) \neq (0, 0)$. Thus $a > 0$. The self-adjoint operator $\Upsilon^{\rm opt} = \Upsilon_a / a$ satisfies $\Upsilon^{\rm opt} \geq \widetilde{\rho}_i$ for all i (i.e., $\Upsilon^{\rm opt} \in \Gamma$) and ${\rm tr}(\Upsilon^{\rm opt}) \leq P_{\rm S,a}^{\rm opt}$, see (6.30) and (6.31). The converse of the last inequality follows from (6.26). Whence $P_{\rm S,a}^{\rm opt} = {\rm tr}(\Upsilon^{\rm opt}) = \min_{\Upsilon \in \Gamma} {\rm tr}(\Upsilon)$, as claimed in the proposition. This identity implies $\sum_i {\rm tr}[(\Upsilon^{\rm opt} - \widetilde{\rho}_i)M_i^{\rm opt}] = 0$ if $\{M_i^{\rm opt}\}$ is an optimal POVM. But all traces in

the sum are non-negative, thus they vanish and (6.24) is satisfied by the arguments given above to derive this equation. It results from (6.24) that $\Upsilon^{\rm opt} = \sum_i \widetilde{\rho}_i M_i^{\rm opt} = \sum_i M_i^{\rm opt} \widetilde{\rho}_i$. This concludes the proof. □

Let us consider the success probability

$$P_{\rm S,a}^{\rm opt\ v.N.}(\{\rho_i, \eta_i\}) = \max_{\{\Pi_i\}} \left\{ \sum_{i=1}^m \eta_i \, {\rm tr}(\Pi_i \rho_i) \right\} , \tag{6.32}$$

where the maximum is over all von Neumann measurements $\{\Pi_i\}_{i=1}^m$. A natural question is whether this probability may be equal to $P_{\rm S,a}^{\rm opt}$, i.e., whether the states ρ_i may be discriminated optimally with a von Neumann measurement. We have already argued above that this is not always the case, even for pure states. A simple consequence of Proposition 6.4.1 is that the equality holds for *linearly independent states*. The states ρ_i are called linearly independent if their eigenvectors $|\zeta_{ij}\rangle$ with non-zero eigenvalues form a linearly independent family $\{|\zeta_{ij}\rangle\}_{i=1,\ldots,m}^{j=1,\ldots,r_i}$ in $\mathcal{H}$ (here r_i is the rank of ρ_i). We say that they span the Hilbert space $\mathcal{H}$ if $\mathcal{H} = {\rm span}\{|\zeta_{ij}\rangle\}_{i=1,\ldots,m}^{j=1,\ldots,r_i}$. Without loss of generality one can restrict $\mathcal{H}$ to a subspace $\mathcal{H}'$ spanned by the ρ_i.

Corollary 6.4.3. [84] *Let $\{|\psi_i\rangle, \eta_i\}_{i=1}^m$ be an ensemble of pure states spanning $\mathcal{H}$. Then the optimal measurement operators $M_i^{\rm opt}$ in ambiguous state discrimination are of rank one. More generally, for any ensemble $\{\rho_i, \eta_i\}_{i=1}^m$ spanning $\mathcal{H}$, the optimal measurement operators have ranks ${\rm rank}(M_i^{\rm opt}) \leq {\rm rank}(\rho_i)$ for all $i = 1, \ldots, m$.*

Corollary 6.4.4. [85] *Let $\{\rho_i, \eta_i\}_{i=1}^m$ be an ensemble of linearly independent states spanning $\mathcal{H}$. Then an optimal measurement in ambiguous state discrimination is a von Neumann measurement with orthogonal projectors $M_i^{\rm opt} = \Pi_i^{\rm opt}$ of rank $r_i = {\rm rank}(\rho_i)$. In particular, the probabilities (6.2) and (6.32) are equal.*

Proof. Let us set $N_i^{\rm opt} = \Upsilon^{\rm opt} - \widetilde{\rho}_i$. The relation (6.24) implies ${\rm ran}\, M_i^{\rm opt} \subset \ker N_i^{\rm opt}$, hence ${\rm rank}(M_i^{\rm opt}) \leq \dim(\ker N_i^{\rm opt})$. Since the rank of the sum of two matrices is smaller or equal to the sum of their ranks, ${\rm rank}(\Upsilon^{\rm opt}) \leq {\rm rank}(N_i^{\rm opt}) + r_i$ and thus $\dim(\ker N_i^{\rm opt}) \leq \dim(\ker \Upsilon^{\rm opt}) + r_i$. But $\ker \Upsilon^{\rm opt} \subset [{\rm ran}(\rho_i)]^\perp$ for all i according to the condition (ii) of Proposition 6.4.1. Consequently, if the states ρ_i span $\mathcal{H}$ then $\ker \Upsilon^{\rm opt} = \{0\}$. This shows that ${\rm rank}(M_i^{\rm opt}) \leq r_i$. If furthermore the ρ_i are linearly independent, then $\sum_i r_i = n = \dim \mathcal{H}$. Introducing the spectral decomposition $M_i^{\rm opt} = \sum_k |\widetilde{\mu}_{ik}\rangle\langle\widetilde{\mu}_{ik}|$ with unnormalized vectors $|\widetilde{\mu}_{ik}\rangle$, $k = 1, \ldots, r_i$, and noting that the sum $\sum_{i,k} |\widetilde{\mu}_{ik}\rangle\langle\widetilde{\mu}_{ik}| = 1$ contains at most n terms, it follows that $\{|\widetilde{\mu}_{ik}\rangle\}$ is an orthonormal basis of $\mathcal{H}$. Thus $M_i^{\rm opt}$ are orthogonal projectors of rank r_i. □

6.5 Bounds on the maximal success probability

We now establish some inequalities satisfied by $P_{\rm S}^{\rm opt}$ for any number m of states to discriminate. A review of various upper bounds for ambiguous discrimination can be found in [200] (see also [237, 238] for upper and lower bounds of the same spirit as in (6.21)). We only discuss here the bounds involving the fidelity

$$F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1^2 = \left({\rm tr}[(\sqrt{\sigma}\rho\sqrt{\sigma})^{\frac{1}{2}}] \right)^2 . \tag{6.33}$$

The properties of this fidelity will be analyzed in the forthcoming chapter 8. Let us only mention here that $F(\rho, \sigma)$ is symmetric under the exchange of ρ and σ (actually, $\sqrt{\sigma}\rho\sqrt{\sigma}$ and $\sqrt{\rho}\,\sigma\sqrt{\rho}$ have the same non-zero eigenvalues) and reduces for pure states $\rho_\psi = |\psi\rangle\langle\psi|$ and $\sigma_\phi = |\phi\rangle\langle\phi|$ to the square modulus of the scalar product $\langle\psi|\phi\rangle$, i.e., $F(\rho_\psi, \sigma_\phi) = |\langle\psi|\phi\rangle|^2$. More generally, $F(\rho, \sigma)$ can be seen as a measure of non-orthogonality of ρ and σ.

The following lower and upper bounds on the maximum success probability $P_{\rm S,a}^{\rm opt}$ for ambiguous state discrimination are taken from Refs. [29] and [173], respectively[3].

Proposition 6.5.1. (Barnum and Knill [29], Montanaro [173]). *For any ensemble $\{\rho_i, \eta_i\}_{i=1}^m$, one has*

$$1 - \sum_{i>j} \sqrt{\eta_i \eta_j F(\rho_i, \rho_j)} \leq P_{\rm S,a}^{\rm opt}(\{\rho_i, \eta_i\}) \leq 1 - \sum_{i>j} \eta_i \eta_j F(\rho_i, \rho_j) \, . \tag{6.34}$$

[3]The upper bound is established in [29] (and is often reported in subsequent works) with an unnecessary extra factor of two in front of the sum (after correcting the obvious misprints in this reference).

The inequalities (6.34) make quantitative the intuitive fact that the more pairwise orthogonal are the states ρ_i, the larger is the success probability to discriminate them, and conversely.

Proof. Let $\rho_i = A_i A_i^*$, the operators A_i being, for instance, given by (5.35). Given a POVM $\{M_i\}$ with Kraus operators R_i (i.e., $M_i = R_i^* R_i$), we set

$$S_{ij} = \sqrt{\eta_j} R_i A_j \quad , \quad B_{ij} = \sqrt{\eta_i \eta_j} A_i^* A_j \; . \tag{6.35}$$

We view $S = (S_{ij})_{i,j=1}^m$ and $B = (B_{ij})_{i,j=1}^m$ as $m \times m$ matrices with values in $\mathcal{B}(\mathcal{H})$, which are related by $S^*S = B \geq 0$ (this follows from $\sum_i R_i^* R_i = 1$). Observe that

$$P_{\rm S,a}^{\{M_i\}} = \sum_j \eta_j \, {\rm tr}(M_j \rho_j) = 1 - \sum_{i \neq j} \eta_j \, {\rm tr}(M_i \rho_j) = 1 - \sum_{i \neq j} \|S_{ij}\|_2^2 \tag{6.36}$$

and

$$\eta_i \eta_j F(\rho_i, \rho_j) = \eta_i \eta_j \| \sqrt{\rho_i} \sqrt{\rho_j} \|_1^2 = \eta_i \eta_j \| U_i^* \sqrt{\rho_i} \sqrt{\rho_j} U_j \|_1^2 = \|B_{ij}\|_1^2 \; , \tag{6.37}$$

where $\| \cdot \|_{1,2}$ are the trace and Hilbert Schmidt norms. We have used in (6.37) the polar decomposition $A_i = \sqrt{\rho_i} U_i$ and the unitary invariance of these norms. The main idea to prove the first inequality in (6.34) is to bound from below the optimal success probability $P_{\rm S,a}^{\rm opt}$ by the success probability $P_{\rm S,a}^{\rm lsm}$ for discriminating the states with the least square measurement [29]. For the latter, the matrix S in (6.35) is the square root of B (in fact, according to (5.36), $S_{ij}^{\rm lsm} = \sqrt{\eta_i \eta_j} A_i^* \rho_{\rm out}^{-1/2} A_j$ so that $S^{\rm lsm} \geq 0$, and it has been argued above that $|S|^2 = B$). For instance, if the ρ_i are pure states $|\psi_i\rangle$, B and $S^{\rm lsm}$ can be identified with the scalar product matrices $(\langle \widetilde{\psi}_i | \widetilde{\psi}_j \rangle)_{i,j=1}^m$ and $(\langle \widetilde{\mu}_i | \widetilde{\psi}_j \rangle)_{i,j=1}^m$, respectively, with $|\widetilde{\psi}_i\rangle = \sqrt{\eta_i} |\psi_i\rangle$ and $|\widetilde{\mu}_i\rangle = \sqrt{\eta_i} \rho_{\rm out}^{-1/2} |\psi_i\rangle$, the latter being the vectors describing the least square measurement (Sec. 5.5.3). The identity $S^{\rm lsm} = \sqrt{B}$ then becomes evident from the definition of a POVM[4]. Therefore, in view of (6.36), $P_{\rm S,a}^{\rm opt} \geq P_{\rm S,a}^{\rm lsm} = 1 - \sum_{i \neq j} \|(\sqrt{B})_{ij}\|_2^2$. The lower bound in (6.34) comes from the following norm inequality proven in Appendix B: for any fixed $j = 1, \ldots, m$,

$$\sum_{i, i \neq j} \|(\sqrt{B})_{ij}\|_2^2 \leq \frac{1}{2} \sum_{i, i \neq j} \|B_{ij}\|_1 \; , \tag{6.38}$$

where the last sum is related to the fidelities by (6.37).

It remains to establish the upper bound. With the notation above, this bound takes the form

$$\frac{1}{2} \sum_{i \neq j} \|B_{ij}\|_1^2 \leq \sum_{i \neq j} \|S_{ij}\|_2^2 \; . \tag{6.39}$$

Fixing j again and introducing the notation $\| \cdot \|_{1/2}$ as in (4.2) (note that this is not a norm), if one can show that

$$\Big\| \sum_{i, i \neq j} |B_{ij}|^2 \Big\|_{\frac{1}{2}} \leq \sum_{i, i \neq j} \Big(\|S_{ij}\|_2^2 + \|S_{ji}\|_2^2 \Big) \tag{6.40}$$

then the required inequality (6.39) will be proven. Actually, by the inverse Minkowski inequality (1) in Appendix B one finds $\sum_i \|B_{ij}\|_1^2 = \sum_i \| |B_{ij}|^2 \|_{1/2} \leq \| \sum_i |B_{ij}|^2 \|_{1/2}$. In order to show (6.40), let us introduce the following $(m-1) \times (m-1)$ matrices with values in $\mathcal{B}(\mathcal{H})$:

$$\begin{array}{rclcrcl} C^{(j)} & = & \displaystyle\sum_{i, i \neq j} (S_{ji})^* \otimes |i\rangle\langle 1| & , & D^{(j)} & = & S_{jj} \otimes |1\rangle\langle 1| \\ E^{(j)} & = & \displaystyle\sum_{i \neq j} \sum_{k \neq j} (S_{ki})^* \otimes |i\rangle\langle k| & , & F^{(j)} & = & \displaystyle\sum_{k, k \neq j} S_{kj} \otimes |k\rangle\langle 1| \end{array} \tag{6.41}$$

(here $|i\rangle\langle k|$ stands for the matrix with vanishing entries except in the ith raw and kth column, which has a unit entry). An explicit calculation leads to

$$\big\| C^{(j)} D^{(j)} + E^{(j)} F^{(j)} \big\|_1^2 = \Big\| \sum_{i, i \neq j} |B_{ij}|^2 \Big\|_{\frac{1}{2}} \quad , \quad \big\| C^{(j)} \big\|_2^2 = \sum_{i, i \neq j} \|S_{ji}\|_2^2 \quad , \quad \big\| F^{(j)} \big\|_2^2 = \sum_{k, k \neq j} \|S_{kj}\|_2^2 \; . \tag{6.42}$$

[4]This remarkable identity has been singled out for pure states in [115]. The authors of this reference suggest to use it as a definition of the least square measurement.

Furthermore,

$$\left\|C^{(j)}\right\|_2^2 + \left\|D^{(j)}\right\|_2^2 + \left\|E^{(j)}\right\|_2^2 + \left\|F^{(j)}\right\|_2^2 = \sum_{i,k} \|S_{ik}\|_2^2 = \sum_k \eta_k \operatorname{tr}(\rho_k) = 1 \,. \tag{6.43}$$

We can now take advantage of the norm inequality (4) of Appendix B. Because of (6.43), this gives

$$\left\|C^{(j)}D^{(j)} + E^{(j)}F^{(j)}\right\|_1^2 \leq \left\|C^{(j)}\right\|_2^2 + \left\|F^{(j)}\right\|_2^2 \,. \tag{6.44}$$

We plug the equalities (6.42) into this result to obtain (6.40). This concludes the proof. □

Let us now turn to unambiguous discrimination. The following easy-to-derive bound generalizes the upper line in (6.15).

Proposition 6.5.2. [90] *The maximum success probability for unambiguous state discrimination is bounded by*

$$P_{\mathrm{S,u}}^{\mathrm{opt}}(\{\rho_i, \eta_i\}) \leq 1 - \left(\frac{2m}{m-1} \sum_{i>j} \eta_i \eta_j F(\rho_i, \rho_j)\right)^{\frac{1}{2}} . \tag{6.45}$$

Proof. The failure probability $P_0 = 1 - P_{\mathrm{S,u}}$ satisfies

$$P_0^2 = \left(\sum_{i=1}^m \eta_i \operatorname{tr}(M_0\rho_i)\right)^2 \geq \frac{m}{m-1} \sum_{i\neq j} \eta_i\eta_j \operatorname{tr}(M_0\rho_i)\operatorname{tr}(M_0\rho_j) \geq \frac{m}{m-1} \sum_{i\neq j} \eta_i\eta_j \left|\operatorname{tr}(U_{ij}\sqrt{\rho_i}M_0\sqrt{\rho_j})\right|^2 , \tag{6.46}$$

where U_{ij} are arbitrary unitary operators and the first and second bounds follow from the Cauchy-Schwarz inequality. Expressing M_0 as $1 - \sum_i M_i$ and using $\operatorname{ran} M_i \subset \ker \rho_j$ for $i \neq j$, one gets $\operatorname{tr}(U_{ij}\sqrt{\rho_i}M_0\sqrt{\rho_j}) = \operatorname{tr}(U_{ij}\sqrt{\rho_i}\sqrt{\rho_j})$. Using the formula $F(\rho_i, \rho_j) = \max_U |\operatorname{tr}(U\sqrt{\rho_i}\sqrt{\rho_j})|^2$ and maximizing over all unitaries U_{ij}, one obtains (6.45). □

One infers from the last two propositions and the Cauchy-Schwarz inequality that

Corollary 6.5.3. *The minimal failure probabilities $P_{\mathrm{err,a}}^{\mathrm{opt}} = 1 - P_{\mathrm{S,a}}^{\mathrm{opt}}$ and P_0^{opt} for discriminating m states ambiguously and unambiguously satisfy $P_0^{\mathrm{opt}} \geq 2P_{\mathrm{err,a}}^{\mathrm{opt}}/(m-1)$.*

In particular, as noted in [37], for two states P_0^{opt} is at least twice larger than $P_{\mathrm{err,a}}^{\mathrm{opt}}$.

6.6 The Holevo bound

Let us come back to the issue of encoding an input message A in an ensemble $\{\rho_i, \eta_i\}$ of quantum states and transmitting it to a receiver. From an information point of view, it makes sense to optimize the measurement in such a way as to maximize the mutual information between the input message A and the output message B reconstructed by the receiver (that is, B is the set of measurement outcomes). This mutual information is defined as [215]

$$I_{A:B} = H(A) + H(B) - H(A,B) \,, \tag{6.47}$$

where $H(A) = -\sum_i \eta_i \ln \eta_i$ is the Shannon entropy of the input message, $H(B) = -\sum_j p_j \ln p_j$ is the Shannon entropy of the measurement outcomes B with probabilities $p_j = \sum_i \eta_i \operatorname{tr}(M_j\rho_i)$, and $H(A,B) = -\sum_{i,j} p_{ij} \ln p_{ij}$ is the Shannon entropy of the joint process (A,B) with probabilities $p_{ij} = \eta_i p_{j|i} = \eta_i \operatorname{tr}(M_j\rho_i)$, see (6.1). One can show from the concavity of the logarithm that $I_{A:B} \geq 0$ and $I_{A:B} = 0$ if and only if A and B are independent.

The conditional Shannon entropies are defined by

$$H(B|A) = -\sum_i \eta_i \sum_j p_{j|i} \ln p_{j|i} \quad , \quad H(A|B) = -\sum_j p_j \sum_i \eta_{i|j} \ln \eta_{i|j} \,, \tag{6.48}$$

where $p_{j|i} = \operatorname{tr}(M_j\rho_i)$ is the conditional probability of the measurement outcome j given the state ρ_i and $\eta_{i|j}$ the conditional (*a posteriori*) probability that the state is ρ_i given the outcome j. The latter is given by the Bayes rule $\eta_{i|j} = \eta_i p_{j|i}/p_j$. The conditional entropy $H(A|B)$ represents the lack of knowledge of the receiver on the state of the ensemble that was sent to him, after he has performed the measurement. In general the measurement producing the lowest value of $H(A|B)$ is not a von Neumann measurement [74]. Thanks to the

well-known relation $H(A,B) = H(A) + H(B|A) = H(B) + H(A|B)$, the mutual information can be expressed in terms of these conditional entropies as [215],

$$I_{A:B} = H(A) - H(A|B) = H(B) - H(B|A) \; . \tag{6.49}$$

As $H(A|B) \geq 0$ one has $I_{A:B} \leq H(A)$, with equality if and only if B is a function of A. This means that if $I_{A:B}$ is maximal, i.e., $I_{A:B} = H(A)$, the receiver can reconstruct without any error the message A from his measurement outcomes. As stressed at the beginning of this chapter, this is never the case if A is encoded using non-orthogonal states ρ_i. Hence $I_{A:B} < H(A)$ for non-orthogonal states. The maximum

$$\max_{\text{POVM } \{M_i\}} \big\{ I_{A:B} \big\} \tag{6.50}$$

measures the maximal amount of information accessible to the receiver, that is, how well can he reconstruct the message. The determination of the optimal measurement maximizing $I_{A:B}$ appears to be a more difficult task than the minimization of the probability of error in state discrimination. However, one can place an upper bound on the maximal information (6.50) by means of the Holevo inequality

$$I_{A:B} \;\leq\; \chi_{\text{Holevo}} = S(\rho) - \sum_i \eta_i S(\rho_i) \quad , \quad \rho = \sum_i \eta_i \rho_i \; , \tag{6.51}$$

where $S(\rho) = -\operatorname{tr}(\rho \ln \rho)$ is the von Neumann entropy of ρ. The proof of this important result relies on the monotonicity of the quantum mutual information under certain quantum operations (see Remark 11.3.3 below). The positive number χ_{Holevo} is called the Holevo quantity. We will show below that $\chi_{\text{Holevo}} \leq H(\{\eta_i\})$ with equality if and only if the ρ_i have orthogonal supports (see (7.8)). We thus recover the aforementioned fact that for non-orthogonal states ρ_i the maximum (6.50) is smaller than the entropy $H(A)$ of the input message.

Chapter 7

Quantum entropies

Un certain désordre favorise la synthèse (M. Serres).

In this Chapter we give the definitions and main properties of the von Neumann entropy, the corresponding relative entropy, and the quantum Rényi relative entropies. For classical systems these entropies reduce to the Shannon entropy, the Kullback-Leibler divergence, and the Rényi divergences, respectively, which are central objects in classical information theory. To begin with we recall in Sec. 7.1 the standard properties of the von Neumann entropy. The most important result for our purpose is the monotonicity of the corresponding relative entropy with respect to quantum operations and the characterization of pairs of states which have the same relative entropy than their transformed states under a given operation. The proof of this result, which will be used later in chapter 11, is given in Sec. 7.2. We finally present in Sec. 7.3 the quantum version of the Rényi divergences introduced recently in [175, 260, 95]. This quantum version contains as special cases the von Neumann relative entropy and the logarithm of the fidelity (6.33). The fidelity and the closely related Bures distance will be the subject of chapter 8. Together with the von Neumann relative entropy, it plays a major role in our geometrical approach of quantum correlations (chapter 12). The generalization of this approach to the whole family formed by the relative Rényi entropies constitutes an interesting open problem that will not be deeply explored in this article. The reader may thus skip Sec. 7.3 in a first reading.

7.1 The von Neumann entropy

The entropy $H(\{p_k\}) = -\sum_k p_k \ln p_k$ introduced by Shannon in his two celebrated 1948 papers [215] quantifies the amount of information at our disposal on the state of a classical system. It vanishes when the state is perfectly known and takes its maximum value (equal to $\ln n$ if the system has n distinct possible states) when one has no information on this state at all, that is, if all possible states are equiprobable. The quantum analog of the Shannon entropy is the von Neumann entropy

$$S(\rho) = -\operatorname{tr}(\rho \ln \rho) \; . \tag{7.1}$$

This is a unitary invariant quantity, i.e., $S(U\rho U^*) = S(\rho)$ for U unitary. Moreover, S is additive for composite systems, i.e., $S(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{B}}) = S(\rho_{\mathsf{A}}) + S(\rho_{\mathsf{B}})$ for any states ρ_{A} and ρ_{B} of the systems A and B. Another important property of S is its strictly concavity[1], i.e., for any states ρ_0, ρ_1 and $0 \leq \eta \leq 1$ it holds $S((1-\eta)\rho_0 + \eta\rho_1) \geq (1-\eta)S(\rho_0) + \eta S(\rho_1)$, with equality if and only if $\rho_0 = \rho_1$ or $\eta \in \{0, 1\}$.

A much less trivial property of importance in quantum information theory is the so-called strong subadditivity

$$S(\rho_{\mathsf{AB}}) + S(\rho_{\mathsf{BC}}) - S(\rho_{\mathsf{ABC}}) - S(\rho_{\mathsf{B}}) \geq 0 \; , \tag{7.2}$$

where ρ_{ABC} is a state of ABC with marginals $\rho_{\mathsf{AB}} = \operatorname{tr}_{\mathsf{C}}(\rho_{\mathsf{ABC}})$, $\rho_{\mathsf{BC}} = \operatorname{tr}_A(\rho_{\mathsf{ABC}})$, and $\rho_{\mathsf{B}} = \operatorname{tr}_{\mathsf{AC}}(\rho_{\mathsf{ABC}})$. The inequality (7.2) was first proven by Lieb and Ruskai [154] by using a former work of Lieb [153] on the concavity of the map $\rho \mapsto \operatorname{tr}(K^* \rho^{1+\beta} K \rho^{-\beta})$ for $-1 \leq \beta \leq 0$ (see Lemma 7.3.2 below). Alternatively, (7.2) is a direct consequence of the monotonicity of the relative entropy (Theorem 7.2.1 below), which can be established by other means than Lieb's concavity theorem. Choosing $\mathcal{H}_{\mathsf{B}} = \mathbb{C}$, the strong subadditivity (7.2) implies that S is subadditive, i.e., $S(\rho_{\mathsf{AC}}) \leq S(\rho_{\mathsf{A}}) + S(\rho_{\mathsf{C}})$.

[1] This comes from the strict convexity of $f(x) = x \ln x$. Actually, it is not hard to prove that if f is strictly convex then the map $\rho \in \mathcal{E}(\mathcal{H}) \mapsto \operatorname{tr}[f(\rho)]$ is strictly convex [50].

As is well know in statistical physics, the von Neumann entropy $S(\rho)$ is the Legendre transform of the free energy $\Phi(\beta, H) = -\beta^{-1} \ln \operatorname{tr}(\mathrm{e}^{-\beta H})$. More precisely, one has (see [50], Theorem 2.13)

$$S(\rho) = \inf_{H \in \mathcal{B}(\mathcal{H})_{\text{s.a.}}} \big\{ \beta \operatorname{tr}(H\rho) - \beta \Phi(\beta, H) \big\} \quad , \quad \Phi(\beta, H) = \inf_{\rho \in \mathcal{E}(\mathcal{H})} \big\{ \operatorname{tr}(H\rho) - \beta^{-1} S(\rho) \big\} \, , \tag{7.3}$$

and the last infimum is attained if and only if ρ is the Gibbs state $\rho_\beta = \mathrm{e}^{-\beta H} / \operatorname{tr}(\mathrm{e}^{-\beta H})$. The free energy is a concave function of the energy observable H.

The following identity will be used repeatedly in chapters 10 and 11:

$$S(\rho_{\mathsf{A}}) = S(\rho_{\mathsf{B}}) \quad \text{ if } \rho_{\mathsf{A}} \text{ and } \rho_{\mathsf{B}} \text{ are the reduced states of the pure state } |\Psi_{\mathsf{AB}}\rangle \text{ of } \mathsf{AB}. \tag{7.4}$$

It is a consequence of Theorem 4.2.1, since if $|\Psi_{\mathsf{AB}}\rangle$ has Schmidt coefficients μ_i then $S(\rho_{\mathsf{A}}) = S(\rho_{\mathsf{B}}) = -\sum_i \mu_i \ln \mu_i$.

A last identity worthwhile mentioning here is

$$S(\rho) = \min_{\{|\psi_i\rangle, \eta_i\}} H(\{\eta_i\}) = \min_{\{|\psi_i\rangle, \eta_i\}} \Big\{ -\sum_{i=1}^m \eta_i \ln \eta_i \Big\} \, , \tag{7.5}$$

where the minimum is over all pure state decompositions of ρ. Furthermore, a decomposition minimizes $H(\{\eta_i\})$ if and only if it is a spectral decomposition of ρ. These statements can be justified as follows[2]. Let $\{|k\rangle, p_k\}_{i=1}^r$ be a spectral decomposition of ρ, with $r = \operatorname{ran}(\rho)$. An arbitrary pure state decomposition $\{|\psi_i\rangle, \eta_i\}_{i=1}^m$ of ρ has the form $\sqrt{\eta_i}|\psi_i\rangle = \sum_k u_{ik} \sqrt{p_k} |k\rangle$, where (u_{ik}) is a $m \times m$ unitary matrix and $m \geq r$ (see (4.16)). Setting $p_k = 0$ for $r < k \leq m$ one gets $\eta_i = \sum_k |u_{ik}|^2 p_k$. Since $f(x) = x \ln x$ is strictly convex, one finds

$$-H(\{\eta_i\}) = \sum_{i=1}^m \eta_i \ln \eta_i \leq \sum_{i,k=1}^m |u_{ik}|^2 p_k \ln p_k = \sum_{k=1}^r p_k \ln p_k = -S(\rho) \, , \tag{7.6}$$

so that $S(\rho) \leq H(\{\eta_i\})$. By strict convexity, the inequality in (7.6) is an equality if and only if for any i, there exists some $k_i \in \{1, \ldots, r+1\}$ such that $u_{ik} = 0$ when $k \notin I_i = \{k = 1, \ldots, m; p_k = p_{k_i}\}$. Thus $S(\rho) = H(\{\eta_i\})$ if and only if

$$\sqrt{\eta_i} |\psi_i\rangle = \sqrt{p_{k_i}} \sum_{k \in I_i} u_{ik} |k\rangle \tag{7.7}$$

are eigenvectors of ρ with eigenvalue $\eta_i = p_{k_i}$ (if $p_{k_i} \neq 0$). It remains to check that $\langle \psi_i | \psi_j \rangle = 0$ when $p_{k_i} = p_{k_j} \neq 0$. This comes from the unitarity of (u_{ik}). This yields the desired result. The inequality (7.6) can be easily generalized to get[3]

$$S(\rho) \leq H(\{\eta_i\}) + \sum_i \eta_i S(\rho_i) \tag{7.8}$$

for any ensemble $\{\rho_i, \eta_i\}$ forming a convex decomposition of ρ. Moreover, one has equality if and only if the ρ_i have orthogonal supports.

7.2 Relative entropy

A related quantity to the von Neumann entropy is the relative entropy introduced by Umegaki [244] and later extended by Araki [12] in the von Neumann algebra setting,

$$S(\rho||\sigma) = \begin{cases} \operatorname{tr}\big(\rho(\ln \rho - \ln \sigma)\big) & \text{if } \ker(\sigma) \subset \ker(\rho) \\ +\infty & \text{otherwise.} \end{cases} \tag{7.9}$$

Note that by taking $\sigma = 1/n$ proportional to the identity operator, $S(\rho||1/n) = \ln n - S(\rho)$ is the difference between the maximal and the von Neumann entropy of ρ. The relative entropy has the following properties:

(i) $S(\rho||\sigma) \geq 0$ with equality if and only if $\rho = \sigma$;

(ii) unitary invariance $S(U\rho U^*||U\sigma U^*) = S(\rho||\sigma)$ for any unitary U;

[2] An alternative proof can be found in [180].

[3] This follows from (7.5) by writing the spectral decompositions of the ρ_i (see [180], Sec. 11.3).

(iii) additivity for composite systems: $S(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{B}} || \sigma_{\mathsf{A}} \otimes \sigma_{\mathsf{B}}) = S(\rho_{\mathsf{A}} || \sigma_{\mathsf{A}}) + S(\rho_{\mathsf{B}} || \sigma_{\mathsf{B}})$;

(iv) joint convexity: if $0 \leq \eta \leq 1$ then $S((1-\eta)\rho_0 + \eta\rho_1 || (1-\eta)\sigma_0 + \eta\sigma_1) \leq (1-\eta)S(\rho_0||\sigma_0) + \eta S(\rho_1||\sigma_1)$.

The first property (i) follows from Klein's inequality, which states that if f is continuous and strictly convex, then $\mathrm{tr}[f(A) - f(B) - (A-B)f'(B)] \geq 0$, with equality if and only if $A = B$. Its proof can be found for instance in the excellent lecture notes of E.A. Carlen [50]. The properties (ii) and (iii) are immediate consequences of the cyclicity of the trace and the relation $\ln(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{B}}) = \ln \rho_{\mathsf{A}} \otimes 1 + 1 \otimes \ln \rho_{\mathsf{B}}$, as in the case of the von Neumann entropy. The last property (iv) can be deduced from the strong subadditivity (7.2) [157, 158]. It will be proven in Sec. 7.3. Let us point out that (i) implies the aforementioned subadditivity $S(\rho_{\mathsf{AC}}) \leq S(\rho_{\mathsf{A}}) + S(\rho_{\mathsf{C}})$ of the von Neumann entropy, with equality if and only if $\rho_{\mathsf{AC}} = \rho_{\mathsf{A}} \otimes \rho_{\mathsf{C}}$ is a product state (in fact, $S(\rho_{\mathsf{AC}} || \rho_{\mathsf{A}} \otimes \rho_{\mathsf{C}}) = S(\rho_{\mathsf{A}}) + S(\rho_{\mathsf{C}}) - S(\rho_{\mathsf{AC}})$).

Another fundamental property of $S(\rho||\sigma)$ is its monotonicity with respect to CP trace-preserving mappings. This monotonicity means that if one performs the same measurement on two states without readout of the outcomes, the pair of post-measurement states has a lower relative entropy than the pair of states before the measurement. This fact was first proven by Lindblad [158] (see also [12] and [241]). Notice that unlike the relative entropy, the von Neumann entropy is not monotonous with respect to non-projective measurements (see [180], Exercise 11.15). The following theorem provides a necessary and sufficient condition on the two states such that the monotonicity of the relative entropy is satisfied with equality. It is due to Petz [196].

Theorem 7.2.1. (Monotonicity of the relative entropy [196, 116]) *For any quantum operation $\mathcal{M} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ one has $S(\rho||\sigma) \geq S(\mathcal{M}(\rho)||\mathcal{M}(\sigma))$ for all states ρ, $\sigma \in \mathcal{E}(\mathcal{H})$. The inequality is an equality if and only if there exists a quantum operation $\mathcal{R} : \mathcal{B}(\mathcal{H}') \to \mathcal{B}(\mathcal{H})$ such that $\mathcal{R} \circ \mathcal{M}(\sigma) = \sigma$ and $\mathcal{R} \circ \mathcal{M}(\rho) = \rho$. This quantum operation is the transpose operation $\mathcal{R} = \mathcal{R}_{\mathcal{M},\sigma}$ defined in (5.24).*

Let us recall from Sec. 5.5.1 that the transpose operation $\mathcal{R}_{\mathcal{M},\sigma}$ is the quantum operation with Kraus operators

$$R_i = \sqrt{\sigma} A_i^* \mathcal{M}(\sigma)^{-1/2} \,, \tag{7.10}$$

where $\{A_i\}$ are some Kraus operators for $\mathcal{M}$. The conditions $\mathcal{R} \circ \mathcal{M}(\sigma) = \sigma$ and $\mathcal{R} \circ \mathcal{M}(\rho) = \rho$, which mean that ρ and σ can be recovered respectively from $\mathcal{M}(\rho)$ and $\mathcal{M}(\sigma)$ by means of the same quantum operation $\mathcal{R}$, is clearly sufficient to ensure the equality $S(\rho||\sigma) = S(\mathcal{M}(\rho)||\mathcal{M}(\sigma))$ if monotonicity holds true. It is remarkable that this is also a necessary condition, with $\mathcal{R} = \mathcal{R}_{\mathcal{M},\sigma}$ the approximate reversal of $\mathcal{M}$ introduced in the context of quantum error correction (Sec. 5.5).

We present below the derivation of this result given by Petz in Ref. [196], which also provides a nice and simple proof of the monotonicity. A completely different proof of the monotonicity, based on Lieb's concavity theorem as in Ref. [158, 50, 95], will be given in Sec. 7.3 in the more general setting of the Rényi entropies. It is noteworthy that Petz's derivation does neither rely on the Stinespring theorem nor on the Kraus decomposition (albeit it takes advantage of one of its consequence, namely, the Kadison-Schwarz inequality). It makes use of the theory of operator convex functions and of Araki's relative modular operators [13]. Let $\mathcal{M}$ be a quantum operation $\mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ and ρ and σ be two states of $\mathcal{E}(\mathcal{H})$ such that ρ and $\mathcal{M}(\rho)$ are invertible. One can define two relative modular operators by (see chapter 4)

$$\Delta_{\sigma|\rho}(B) = \sigma B \rho^{-1} \quad , \quad \Delta_{\mathcal{M}(\sigma)|\mathcal{M}(\rho)}(B') = \mathcal{M}(\sigma) B' \mathcal{M}(\rho)^{-1} \quad , \quad B \in \mathcal{B}(\mathcal{H}) \,,\, B' \in \mathcal{B}(\mathcal{H}') \,. \tag{7.11}$$

Proof. Let us set $\rho_{\mathcal{M}} = \mathcal{M}(\rho)$ and $\sigma_{\mathcal{M}} = \mathcal{M}(\sigma)$ and assume that ρ, σ, $\rho_{\mathcal{M}}$, and $\sigma_{\mathcal{M}}$ are invertible. In the whole proof these states are fixed, so to simplify notation we write Δ instead of $\Delta_{\sigma|\rho}$ and $\Delta_{\mathcal{M}}$ instead of $\Delta_{\sigma_{\mathcal{M}}|\rho_{\mathcal{M}}}$. We set $\xi = \rho^{\frac{1}{2}}$ and $\xi_{\mathcal{M}} = \rho_{\mathcal{M}}^{\frac{1}{2}}$. One can view these two operators as unit vectors in $\mathcal{B}(\mathcal{H})$ and $\mathcal{B}(\mathcal{H}')$, respectively, for the Hilbert-Schmidt scalar product $\langle \cdot, \cdot \rangle$. The first observation is that

$$S(\rho||\sigma) = \langle \xi \,,\, (\ln \rho - \ln \sigma)\xi \rangle = -\langle \xi \,,\, \ln(\Delta)\xi \rangle = \int_0^\infty \mathrm{d}t \left(\langle \xi \,,\, (\Delta + t)^{-1} \xi \rangle - (1+t)^{-1} \right) . \tag{7.12}$$

The third equality can be established, for instance, with the help of the first identity in (2) (see Appendix A). Therefore, in order to prove that $S(\rho||\sigma) \geq S(\rho_{\mathcal{M}}||\sigma_{\mathcal{M}})$, it suffices to show that for any $t > 0$,

$$\langle \xi_{\mathcal{M}} \,,\, (\Delta_{\mathcal{M}} + t)^{-1} \xi_{\mathcal{M}} \rangle \leq \langle \xi \,,\, (\Delta + t)^{-1} \xi \rangle \,. \tag{7.13}$$

To this end, let us consider the operator $\mathcal{C}_{\mathcal{M}}$ defined by

$$\mathcal{C}_{\mathcal{M}}(B' \xi_{\mathcal{M}}) = \mathcal{M}^*(B') \xi \quad , \quad B \in \mathcal{B}(\mathcal{H}') \,. \tag{7.14}$$

Since $\{B'\xi_{\mathcal{M}}; B' \in \mathcal{B}(\mathcal{H}')\}$ is equal to[4] $\mathcal{B}(\mathcal{H}')$ by the invertibility of $\rho_{\mathcal{M}}$, (7.15) defines an operator $\mathcal{C}_{\mathcal{M}}$ from $\mathcal{B}(\mathcal{H}')$ to $\mathcal{B}(\mathcal{H})$. Then

$$\mathcal{C}_{\mathcal{M}}^* \Delta \mathcal{C}_{\mathcal{M}} \leq \Delta_{\mathcal{M}} \,. \tag{7.15}$$

Actually, thanks to the Kadison-Schwarz inequality (5.9) and the relation $(\mathcal{M}^*(B'^*))^* = \mathcal{M}^*(B')$, one has

$$\begin{aligned} \big\langle \mathcal{C}_{\mathcal{M}}(B'\xi_{\mathcal{M}}) \,,\, \Delta\, \mathcal{C}_{\mathcal{M}}(B'\xi_{\mathcal{M}}) \big\rangle &= \mathrm{tr}\Big(|\mathcal{M}^*(B'^*)|^2 \sigma \Big) \\ &\leq \mathrm{tr}\Big(\mathcal{M}^*(B'B'^*) \sigma \Big) = \big\langle B'\xi_{\mathcal{M}} \,,\, \Delta_{\mathcal{M}} B'\xi_{\mathcal{M}} \big\rangle \,. \end{aligned} \tag{7.16}$$

One shows similarly that $\|\mathcal{C}_{\mathcal{M}}(B'\xi_{\mathcal{M}})\|_2 \leq \|B'\xi_{\mathcal{M}}\|_2$ for any $B' \in \mathcal{B}(\mathcal{H}')$, hence $\|\mathcal{C}_{\mathcal{M}}\| \leq 1$.

We now use the fact that the function $f(x) = (x+t)^{-1}$ is operator monotone-decreasing and operator convex. The definitions of operator monotone and operator convex functions are given in Appendix A. Together with the bound (7.15), this implies[5]

$$(\Delta_{\mathcal{M}} + t)^{-1} \leq (\mathcal{C}_{\mathcal{M}}^* \Delta \mathcal{C}_{\mathcal{M}} + t)^{-1} \leq \mathcal{C}_{\mathcal{M}}^* (\Delta + t)^{-1} \mathcal{C}_{\mathcal{M}} + t^{-1}(1 - \mathcal{C}_{\mathcal{M}}^* \mathcal{C}_{\mathcal{M}}) \,. \tag{7.17}$$

The last inequality follows by applying the Jensen-type inequality (4) for the operator convex function $g(x) = (x+t)^{-1} - t^{-1}$ satisfying $g(0) = 0$ and the contraction $\mathcal{C}_{\mathcal{M}}$. Since $\mathcal{C}_{\mathcal{M}}(\xi_{\mathcal{M}}) = \xi$ by (7.14) and $\mathcal{M}^*(1) = 1$, the inequality (7.17) entails

$$\big\langle \xi_{\mathcal{M}} \,,\, (\Delta_{\mathcal{M}} + t)^{-1} \xi_{\mathcal{M}} \big\rangle \leq \big\langle \xi \,,\, (\Delta + t)^{-1} \xi \big\rangle + t^{-1} \big(\mathrm{tr}(\rho_{\mathcal{M}}) - \mathrm{tr}(\rho) \big) \,. \tag{7.18}$$

The term proportional to t^{-1} vanishes because $\mathcal{M}$ is trace preserving, hence one obtains the desired bound (7.13). We have thus proven the monotonicity of the relative entropy.

In addition to its simplicity, the above proof offers the advantage that it easily yields a necessary and sufficient condition for having $S(\rho||\sigma) = S(\rho_{\mathcal{M}}||\sigma_{\mathcal{M}})$. Actually, this equality holds if and only if (7.13) is an equality, i.e.

$$\big\langle \xi_{\mathcal{M}} \,,\, (\Delta_{\mathcal{M}} + t)^{-1} \xi_{\mathcal{M}} \big\rangle = \big\langle \xi_{\mathcal{M}} \,,\, \big(\mathcal{C}_{\mathcal{M}}^* (\Delta + t)^{-1} \mathcal{C}_{\mathcal{M}} + t^{-1}(1 - \mathcal{C}_{\mathcal{M}}^* \mathcal{C}_{\mathcal{M}}) \big) \xi_{\mathcal{M}} \big\rangle \tag{7.19}$$

for all $t > 0$. But for any operators X, Y, and Z with Z invertible and $X \leq Y$, $\langle Z, XZ\rangle = \langle Z, YZ\rangle$ implies $XZ = YZ$. Hence we can infer from (7.17) and (7.19) that

$$(\Delta_{\mathcal{M}} + t)^{-1} \xi_{\mathcal{M}} = \mathcal{C}_{\mathcal{M}}^* (\Delta + t)^{-1} \xi \quad , \quad t > 0 \,, \tag{7.20}$$

where we have used the identity $\mathcal{C}_{\mathcal{M}}^* \mathcal{C}_{\mathcal{M}}(\xi_{\mathcal{M}}) = \xi_{\mathcal{M}}$ (in fact, the scalar product $\langle \mathcal{C}_{\mathcal{M}}(B'\xi_{\mathcal{M}}) \,,\, \mathcal{C}_{\mathcal{M}} \xi_{\mathcal{M}} \rangle$ is equal to $\langle B'\xi_{\mathcal{M}} \,,\, \xi_{\mathcal{M}} \rangle$ for any $B' \in \mathcal{B}(\mathcal{H}')$). Therefore,

$$\big\| \mathcal{C}_{\mathcal{M}}^* (\Delta + t)^{-1} \xi \big\|_2^2 = \big\langle (\Delta_{\mathcal{M}} + t)^{-2} \xi_{\mathcal{M}}, \xi_{\mathcal{M}} \big\rangle = \big\langle \mathcal{C}_{\mathcal{M}}^* (\Delta + t)^{-2} \xi, \xi_{\mathcal{M}} \big\rangle = \big\| (\Delta + t)^{-1} \xi \big\|_2^2 \,, \tag{7.21}$$

where the second equality is obtained by differentiating (7.20) with respect to t. Now, the identity $\|\mathcal{C}^*(X)\|_2 = \|X\|_2$ for $\mathcal{C}$ a contraction implies that $\mathcal{C}\mathcal{C}^*(X) = X$ (in fact, then the Cauchy-Schwarz inequality $\langle X \,, \mathcal{C}\mathcal{C}^*(X) \rangle \leq \|X\|_2 \|\mathcal{C}\mathcal{C}^*(X)\|_2 \leq \|X\|_2^2$ is an equality, so that $\mathcal{C}\mathcal{C}^*(X)$ must be proportional to X). We conclude that

$$\mathcal{C}_{\mathcal{M}} (\Delta_{\mathcal{M}} + t)^{-1} \xi_{\mathcal{M}} = \mathcal{C}_{\mathcal{M}} \mathcal{C}_{\mathcal{M}}^* (\Delta + t)^{-1} \xi = (\Delta + t)^{-1} \xi \tag{7.22}$$

for any $t > 0$. By means of the functional calculus, one deduces from this identity that

$$\mathcal{C}_{\mathcal{M}} \Delta_{\mathcal{M}}^{-\frac{1}{2}} \xi_{\mathcal{M}} = \Delta^{-\frac{1}{2}} \xi \,. \tag{7.23}$$

In view of the definitions (7.11) and (7.14) and as $\rho > 0$, the last formula gives $\mathcal{M}^*(\sigma_{\mathcal{M}}^{-\frac{1}{2}} \xi_{\mathcal{M}}) = \sigma^{-\frac{1}{2}} \xi$. By multiplying by the adjoint and using the Kadison-Schwarz inequality, we arrive at

$$\sigma^{-\frac{1}{2}} \rho\, \sigma^{-\frac{1}{2}} \leq \mathcal{M}^*(\sigma_{\mathcal{M}}^{-\frac{1}{2}} \rho_{\mathcal{M}} \sigma_{\mathcal{M}}^{-\frac{1}{2}}) \,, \tag{7.24}$$

that is, $\rho \leq \mathcal{R}_{\mathcal{M},\sigma}(\rho_{\mathcal{M}})$ with $\mathcal{R}_{\mathcal{M},\sigma}$ defined in (5.24). But $\mathrm{tr}[\rho] = \mathrm{tr}[\rho_{\mathcal{M}}] = \mathrm{tr}[\mathcal{R}_{\mathcal{M},\sigma}(\rho_{\mathcal{M}})]$, whence $\rho = \mathcal{R}_{\mathcal{M},\sigma}(\rho_{\mathcal{M}})$. The other equality $\sigma = \mathcal{R}_{\mathcal{M},\sigma}(\sigma_{\mathcal{M}})$ is obvious. Reciprocally, as stressed above, these two identities

[4] In the theory of C*-algebras, if this equality is true upon completion of $\{B'\xi_{\mathcal{M}}; B' \in \mathcal{B}'\}$ for the Hilbert-Schmidt norm one says that $(B' \in \mathcal{B}' \mapsto \mathcal{L}_{B'}, \xi_{\mathcal{M}})$ defines a cyclic representation of the algebra $\mathcal{B}'$ on the Hilbert space $\mathcal{B}(\mathcal{H}')$ [42].

[5] In [196] the last term in the right-hand side is omitted. This is a not correct as the Jensen-type inequality (4) cannot be applied for the function $f(x) = (x+t)^{-1}$, because it does not satisfy the condition $f(0) \leq 0$. Fortunately, this term disappears in (7.13) due to the trace-preserving property of $\mathcal{M}$ and the proof goes through.

imply $S(\rho||\sigma) = S(\rho_{\mathcal{M}}||\sigma_{\mathcal{M}})$ thanks to the monotonicity of the relative entropy and the fact that $\mathcal{R}_{\mathcal{M},\sigma}$ is a quantum operation. □

Let us end this section by pointing out that the strong subadditivity of the von Neumann entropy, the joint convexity of the relative entropy, and its monotonicity can be deduced from each other. For instance, the strong subadditivity (7.2) is a simple consequence of the monotonicity. Actually, one checks that

$$S(\rho_{\mathsf{AB}}) + S(\rho_{\mathsf{BC}}) - S(\rho_{\mathsf{ABC}}) - S(\rho_{\mathsf{B}}) = S(\rho_{\mathsf{ABC}}||\rho_{\mathsf{A}} \otimes \rho_{\mathsf{BC}}) - S(\mathcal{M}_C(\rho_{\mathsf{ABC}})||\mathcal{M}_C(\rho_{\mathsf{A}} \otimes \rho_{\mathsf{BC}})) \tag{7.25}$$

with $\mathcal{M}_{\mathsf{C}} : \rho \mapsto \mathrm{tr}_{\mathsf{C}}(\rho)$. It is easy to show that $\mathcal{M}_{\mathsf{C}}$ is a CP and trace-preserving map $\mathcal{B}(\mathcal{H}_{\mathsf{ABC}}) \to \mathcal{B}(\mathcal{H}_{\mathsf{AB}})$, therefore (7.2) follows from Theorem 7.2.1. With the help of this theorem it is also possible to characterize all states ρ_{ABC} such that (7.2) becomes an equality [116].

Conversely, Lindblad [157, 158] proves the monotonicity inequality from the strong subadditivity. The basic idea is to show that the strong subadditivity of the von Neumann entropy or the closely related Lieb concavity theorem imply the joint convexity (iv) of the relative entropy. The corresponding arguments are given in Sec. 7.3.2 below. One can then deduce the monotonicity of the relative entropy from its joint convexity (iv) with the help of Stinespring's theorem as follows [239, 263, 95]. Recall that if μ_H is the normalized Haar measure on the group $U(n)$ of $n \times n$ unitary matrices, then $\int \mathrm{d}\mu_H(U)\, UBU^* = n^{-1}\,\mathrm{tr}(B)$ for any $B \in \mathcal{B}(\mathcal{H})$ (in fact, all diagonal matrix elements of the left-hand side in an arbitrary basis are equal, as follows from the left-invariance $\mathrm{d}\mu_H(VU) = \mathrm{d}\mu_H(U)$ for $V \in U(n)$; as a result, this left-hand side is proportional to the identity matrix). We infer from Stinespring theorem 5.2.2 that

$$\mathcal{M}(\rho) \otimes (1/n_{\mathsf{E}}) = \int_{U(n_{\mathsf{E}})} \mathrm{d}\mu_H(U_{\mathsf{E}})\, (1 \otimes U_{\mathsf{E}}) U\rho \otimes |\epsilon_0\rangle\langle\epsilon_0| U^* (1 \otimes U_{\mathsf{E}}^*) \tag{7.26}$$

with U unitary on $\mathcal{H}_{\mathsf{SE}}$. Thanks to the additivity (iii), the joint convexity convexity (iv), and the unitary invariance (ii), we get

$$\begin{aligned} S(\mathcal{M}(\rho)||\mathcal{M}(\sigma)) &= S\big(\mathcal{M}(\rho) \otimes (1/n_{\mathsf{E}})\big|\big|\mathcal{M}(\sigma) \otimes (1/n_{\mathsf{E}})\big) \\ &\leq \int_{U(n_{\mathsf{E}})} \mathrm{d}\mu_H(U_{\mathsf{E}}) S\big((1 \otimes U_{\mathsf{E}}) U\rho \otimes |\epsilon_0\rangle\langle\epsilon_0| U^* (1 \otimes U_{\mathsf{E}}^*)\big|\big|(1 \otimes U_{\mathsf{E}}) U\sigma \otimes |\epsilon_0\rangle\langle\epsilon_0| U^* (1 \otimes U_{\mathsf{E}}^*)\big) \\ &= \int_{U(n_{\mathsf{E}})} \mathrm{d}\mu_H(U_{\mathsf{E}}) S(\rho||\sigma) = S(\rho||\sigma)\ . \end{aligned} \tag{7.27}$$

By the same argument, one can show a slightly more general result.

Proposition 7.2.2. *Let $f : \mathcal{E}(\mathcal{H}) \times \mathcal{E}(\mathcal{H}) \to \mathbb{R}$ be a unitary-invariant jointly convex function for any finite Hilbert space $\mathcal{H}$, which satisfies $f(\rho \otimes \tau, \sigma \otimes \tau) = f(\rho, \sigma)$ for all $\rho, \sigma \in \mathcal{E}(\mathcal{H})$ and $\tau \in \mathcal{E}(\mathcal{H}')$. Then f is monotonous with respect to quantum operations.*

7.3 Quantum relative Rényi entropies

7.3.1 Definitions

In the classical theory of information, other entropies than the Shannon entropy play a role when ergodicity breaks down or outside the asymptotic regime. The Rényi entropy depending on a parameter $\alpha > 0$ unifies these different entropies. In the quantum setting, it is defined as

$$S_\alpha(\rho) = (1-\alpha)^{-1} \ln \mathrm{tr}(\rho^\alpha)\ . \tag{7.28}$$

It is easy to show that $S_\alpha(\rho)$ converges to the von Neumann entropy $S(\rho)$ when $\alpha \to 1$ and that $S_\alpha(\rho)$ is a non-increasing function of α.

A first definition of the quantum relative Réyni entropy is

$$S_\alpha^{(\mathrm{n})}(\rho||\sigma) = (\alpha-1)^{-1} \ln(\mathrm{tr}[\rho^\alpha \sigma^{1-\alpha}]) \quad , \quad \alpha > 0\ ,\ \alpha \neq 1\ . \tag{7.29}$$

This entropy appears naturally in the context of the quantum hypothesis testing (Sec. 9.1 below). We shall discuss here a symmetrized version proposed recently by Müller-Lennert *et al.* [175] and by Wilde, Winter, and Yang [260]. It is given by

$$S_\alpha(\rho||\sigma) = (\alpha-1)^{-1} \ln \mathrm{tr}\big[\big(\sigma^{\frac{1-\alpha}{2\alpha}} \rho\, \sigma^{\frac{1-\alpha}{2\alpha}}\big)^\alpha\big] \tag{7.30}$$

if $\alpha \in (0,1)$ and $\mathrm{tr}(\sigma\rho) > 0$ or if $\alpha > 1$ and $\ker\sigma \subset \ker\rho$ (if none of these conditions are satisfied, one sets $S_\alpha(\rho||\sigma) = +\infty$). This relative entropy has been used in Ref. [260] to solve an important open problem related to the transmission of information in noisy quantum channels. It seems likely that much more applications in quantum information theory will be encountered in the future. The entropies S_α appeared recently as central objects in a very different context, namely, the quantum fluctuation relations in out-of-equilibrium statistical physics [141, 142]. A nice feature of the family $\{S_\alpha\}_{\alpha>0}$ is that it contains the von Neumann relative entropy, the fidelity entropy, and the max-entropy as special cases. Furthermore, S_α depends continuously and monotonously on α. The fidelity-entropy is obtained for $\alpha = 1/2$. It is given by $S_{1/2}(\rho||\sigma) = -\ln F(\rho,\sigma)$, where $F(\rho,\sigma)$ is the fidelity (6.33). The max-entropy is defined by

$$S_\infty(\rho||\sigma) = \lim_{\alpha\to\infty} S_\alpha(\rho||\sigma) = \ln\|\sigma^{-\frac{1}{2}}\rho\sigma^{-\frac{1}{2}}\| \,, \tag{7.31}$$

where $\|\cdot\|$ is the operator norm. The second equality follows from $\|A\|_\alpha \to \|A\|$ as $\alpha\to\infty$ (see Sec. 4.1). Finally, one recovers the von Neumann relative entropy (7.9) by letting $\alpha \to 1$,

$$S(\rho||\sigma) = \lim_{\alpha\to 1} S_\alpha(\rho||\sigma) \,. \tag{7.32}$$

To justify this statement, let us set $A(\alpha) = \sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}$. Explicit calculations show that

$$\begin{aligned} \frac{\mathrm{d}\,\mathrm{tr}[A(\alpha)^\alpha]}{\mathrm{d}\alpha} &= \mathrm{tr}[A(\alpha)^\alpha \ln A(\alpha)] + \alpha\,\mathrm{tr}\Big[A(\alpha)^{\alpha-1}\frac{\mathrm{d}A}{\mathrm{d}\alpha}\Big] \\ \frac{\mathrm{d}A}{\mathrm{d}\alpha} &= -\frac{1}{2\alpha^2}\Big(\ln(\sigma)A(\alpha) + A(\alpha)\ln(\sigma)\Big) \,. \end{aligned} \tag{7.33}$$

Consequently, $S_\alpha(\rho||\sigma) \to (\mathrm{d}\ln\mathrm{tr}[A(\alpha)^\alpha]/\mathrm{d}\alpha)_{\alpha=1} = \mathrm{tr}(\rho\ln\rho - \rho\ln\sigma)$ as $\alpha\to 1$. Note that a similar result holds for the unsymmetrized Rényi entropy (7.29), i.e., $S(\rho||\sigma) = \lim_{\alpha\to1} S_\alpha^{(\mathrm{n})}(\rho||\sigma)$. Let us also emphasize that

$$S_\alpha(\rho||\sigma) \le S_\alpha^{(\mathrm{n})}(\rho||\sigma) \tag{7.34}$$

by the Lieb-Thirring trace inequality (3).

For commuting matrices $\rho = \sum p_k|k\rangle\langle k|$ and $\sigma = \sum_k q_k|k\rangle\langle k|$, both $S_\alpha(\rho||\sigma)$ and $S_\alpha^{(\mathrm{n})}(\rho||\sigma)$ reduce to the classical Réyni divergence

$$S_\alpha^{\mathrm{clas}}(\mathbf{p}||\mathbf{q}) = (\alpha-1)^{-1}\ln\Big(\sum_{k=1}^n p_k^\alpha q_k^{1-\alpha}\Big) \,, \tag{7.35}$$

which is non-negative for $\alpha > 0$ by the Hölder inequality.

7.3.2 Main properties

It is shown in this subsection that the Rényi relative entropy $S_\alpha(\rho||\sigma)$ satisfies the same properties (i-iv) as the von Neumann relative entropy in Sec. 7.2 for any $\alpha \in [1/2, 1]$. For $0 < \alpha < \infty$ we define the α-fidelity by

$$F_\alpha(\rho||\sigma) = \|\rho^{\frac{1}{2}}\sigma^{\frac{\beta}{2}}\|_{2\alpha}^2 = \|\sigma^{\frac{\beta}{2}}\rho\sigma^{\frac{\beta}{2}}\|_\alpha = \mathrm{e}^{-\beta S_\alpha(\rho||\sigma)} \quad \text{with} \quad \beta = \frac{1-\alpha}{\alpha} \,. \tag{7.36}$$

Here, we have used the notation $\|A\|_{2\alpha} = (\mathrm{tr}[(A^*A)^\alpha])^{\frac{1}{2\alpha}}$ even if this does not correspond to a norm when $0 < \alpha < 1/2$.

Theorem 7.3.1. *For any $\alpha > 0$, one has*

(i) $S_\alpha(\rho||\sigma) \ge 0$ with equality if and only if $\rho = \sigma$;

(ii) $S_\alpha(\rho||\sigma)$ is unitary invariant;

(iii) $S_\alpha(\rho||\sigma)$ is additive for composite systems;

(iv) $F_\alpha(\rho||\sigma)^\alpha$ is jointly concave for $\alpha \in [1/2, 1)$ and jointly convex for $\alpha > 1$. In particular, $S_\alpha(\rho||\sigma)$ is jointly convex for $\alpha \in [1/2, 1]$;

(v) if $\alpha \ge 1/2$ then $S_\alpha(\rho||\sigma) \ge S_\alpha(\mathcal{M}(\rho)||\mathcal{M}(\sigma))$ for any quantum operation $\mathcal{M}$ on $\mathcal{B}(\mathcal{H})$.

The statements (i-iii), as well as (iv-v) for a restricted range of α, namely $\alpha \in (1,2]$, have been established in [175, 260]. The justification of (iv-v) in full generality is due to Frank and Lieb [95].

Proof. The unitary invariance (ii) and additivity (iii) are evident and also hold for the α-fidelity. We now argue that the non-negativity (i) and the monotonicity (iv) can be deduced from the convexity/concavity property (iv). Thanks to Proposition 7.2.2, (iv) implies that if $\alpha \in [1/2, 1)$ then $F_\alpha(\mathcal{M}(\rho)||\mathcal{M}(\sigma)) \geq F_\alpha(\rho||\sigma)$ for any quantum operation $\mathcal{M}$, and the reverse inequality holds true if $\alpha > 1$. The monotonicity of S_α for $\alpha \geq 1/2$ then follows immediately (the case $\alpha = 1$ is obtained by continuity, see (7.32)). Let $\{|k\rangle\}$ be an orthonormal basis of $\mathcal{H}$ and $\mathcal{M}_\Pi$ be the quantum operation (5.5) associated to the von Neumann measurement $\{\Pi_k = |k\rangle\langle k|\}$. The monotonicity entails

$$S_\alpha(\rho||\sigma) \geq S_\alpha(\mathcal{M}_\Pi(\rho)||\mathcal{M}_\Pi(\sigma)) = S_\alpha^{\rm clas}(\mathbf{p}||\mathbf{q}) \,, \tag{7.37}$$

where $\mathbf{p}$ and $\mathbf{q}$ are the vectors with components $p_k = \langle k|\rho|k\rangle$ and $q_k = \langle k|\sigma|k\rangle$. Since the classical Rényi divergence (7.35) is non-negative and vanishes if and only if $\mathbf{p} = \mathbf{q}$, we deduce from (7.37) that $S_\alpha(\rho||\sigma) \geq 0$, with equality if and only if $\langle k|\rho|k\rangle = \langle k|\sigma|k\rangle$ for all k. The orthonormal basis $\{|k\rangle\}$ being arbitrary, this justifies the assertion (i) for $\alpha \geq 1/2$. To show this assertion for $\alpha \in (0, 1/2)$, we argue as in [175] that

$$S_\alpha(\rho||\sigma) \geq S_\alpha(\mathcal{M}_\Pi(\rho)||\sigma) = S_\alpha^{\rm clas}(\mathbf{p}||\mathbf{q}) \tag{7.38}$$

with $0 < \alpha < 1$, $\mathcal{M}_\Pi$ being as before associated with the von Neumann $\{\Pi_k = |k\rangle\langle k|\}$ but with $\{|k\rangle\}$ an orthonormal eigenbasis of σ. Actually, let $\alpha \in (0,1)$ and let us set $A(\beta) = \sigma^{\frac{\beta}{2}}\rho\sigma^{\frac{\beta}{2}}$ with $\beta = \alpha^{-1} - 1$. By virtue of the Jensen type inequality (8) of Appendix A, one has

$$\big(\mathcal{M}_\Pi(A(\beta))\big)^\alpha \geq \mathcal{M}_\Pi\big(A(\beta)^\alpha\big) \tag{7.39}$$

due to the operator concavity of $f(x) = x^\alpha$. Hence, by the trace-preserving property of $\mathcal{M}_\Pi$ and the identity $\sigma^{\frac{\beta}{2}}\mathcal{M}_\Pi(\rho)\sigma^{\frac{\beta}{2}} = \mathcal{M}_\Pi(A(\beta))$,

$$\begin{aligned} S_\alpha(\rho||\sigma) &= (\alpha-1)^{-1}\ln \mathrm{tr}\big[\mathcal{M}_\Pi\big(A(\beta)^\alpha\big)\big] \\ &\geq (\alpha-1)^{-1}\ln \mathrm{tr}\big[\big(\mathcal{M}_\Pi\big(A(\beta)\big)\big)^\alpha\big] = S_\alpha(\mathcal{M}_\Pi(\rho)||\sigma) \,. \end{aligned} \tag{7.40}$$

This proves (7.38) and thus the non-negativity of S_α for $\alpha \in (0,1)$. Observe that $S_\alpha(\rho||\sigma) = S_\alpha(\mathcal{M}_\Pi(\rho)||\sigma)$ if and only if (7.39) holds with equality, that is, $\langle k|A(\beta)|k\rangle^\alpha = \langle k|A(\beta)^\alpha|k\rangle$ for all k. By the strict concavity of $f(x) = x^\alpha$, $\{|k\rangle\}$ must then be an eigenbasis of $A(\beta)$, and thereby also of ρ. Thus ρ and σ commute and $S_\alpha(\rho||\sigma)$ coincides with the classical Rényi divergence $S_\alpha^{\rm clas}(\mathbf{p}||\mathbf{q})$. By the aforementioned properties of $S_\alpha^{\rm clas}(\mathbf{p}||\mathbf{q})$, it follows from (7.38) that $S_\alpha(\rho||\sigma) = 0$ implies $\mathbf{p} = \mathbf{q}$ and thus $\rho = \sigma$.

It remains to show the statement (iv) of the theorem. Following [95], we obtain (iv) with the help of a duality formula for $F_\alpha(\rho,\sigma)$ and of Lieb's concavity and Ando's convexity theorems. We omit here the proof of these two important theorems, which can be found in [50] (see also [180] for the Lieb theorem). The duality formula will be shown at the end this subsection.

Lemma 7.3.2. (Lieb's concavity and Ando's convexity theorem [9, 153]) *For any $K \in \mathcal{B}(\mathcal{H})$ and any $\beta \in [-1,1]$, the function $(R,S) \mapsto \mathrm{tr}(K^*R^qKS^{-\beta})$ on $\mathcal{B}(\mathcal{H})_+ \times \mathcal{B}(\mathcal{H})_+$ is jointly concave in (R,S) if $-1 \leq \beta \leq 0$ and $0 \leq q \leq 1+\beta$ and is jointly convex in (R,S) if $0 \leq \beta \leq 1$ and $1+\beta \leq q \leq 2$.*

Lemma 7.3.3. (Duality formula for the α-fidelity [95]) *If $\alpha \in (0,1)$ (that is, $\beta = \alpha^{-1} - 1 > 0$) then*

$$F_\alpha(\rho,\sigma)^\alpha = \inf_{H \geq 0}\Big\{\alpha\,\mathrm{tr}(H\rho) + (1-\alpha)\,\mathrm{tr}\big[(\sqrt{H}\sigma^{-\beta}\sqrt{H})^{-\frac{1}{\beta}}\big]\Big\} \,. \tag{7.41}$$

If $\alpha > 1$ (that is, $-1 < \beta < 0$), the same identity holds but with the infimum replaced by a supremum.

Given Lemma 7.3.3, if one can show that, for a fixed operator $B \in \mathcal{B}(\mathcal{H})$, the function

$$g_{B,\beta}(\sigma) = \mathrm{tr}\big[(B^*\sigma^{-\beta}B)^{-\frac{1}{\beta}}\big] \tag{7.42}$$

is concave in σ when $-1 \leq \beta \leq 1$, $\beta \neq 0$, it will follow that $F_\alpha(\rho||\sigma)^\alpha$ is jointly concave for $\alpha \in [1/2, 1)$ (i.e., $0 < \beta \leq 1$) and jointly convex for $\alpha > 1$ (i.e., $-1 < \beta < 0$), thereby proving Theorem 7.3.1. We first assume $-1 \leq \beta < 0$. For any operator $Y \geq 0$, let us set

$$h_Y(X) = \mathrm{tr}(YX^{1+\beta}) - (1+\beta)\,\mathrm{tr}(X) \tag{7.43}$$

with $X \in \mathcal{B}(\mathcal{H})_+$. Given two self-adjoint matrices Y and Z, it is known that (see [38], Problem III.6.14)

$$\sum_{i=1}^n y_{n-i} z_i \leq \mathrm{tr}(YZ) \leq \sum_{i=1}^n y_i z_i \ , \tag{7.44}$$

where $y_1 \geq y_2 \geq \cdots \geq y_n$ and $z_1 \geq z_2 \geq \cdots \geq z_n$ are the eigenvalues of Y and Z in non-increasing order. Therefore,

$$\sup_{X \geq 0} \{h_Y(X)\} = \max_{\mathbf{x}} \Big\{ \sum_{i=1}^n \big(y_i x_i^{1+\beta} - (1+\beta) x_i \big) \Big\} = -\beta \sum_{i=1}^n y_i^{-\frac{1}{\beta}} = -\beta \, \mathrm{tr}\big(Y^{-\frac{1}{\beta}}\big) \ , \tag{7.45}$$

the maximum in the second member being over all vectors $\mathbf{x} \in \mathbb{R}_+^n$. Similarly, it follows from (7.44) that if $0 < \beta \leq 1$ then $\inf_{X \geq 0} \{h_Y(X)\} = -\beta \, \mathrm{tr}(Y^{-\frac{1}{\beta}})$. Plugging $Y = B^* \sigma^{-\beta} B$ into these identities, one finds

$$g_{B,\beta}(\sigma) = \sup_{X \geq 0} \Big\{ -\beta^{-1} \big(\mathrm{tr}(B^* \sigma^{-\beta} B X^{1+\beta}) - (1+\beta) \, \mathrm{tr}(X) \big) \Big\} \quad , \quad -1 \leq \beta < 0 \ \text{ or } \ 0 < \beta \leq 1. \tag{7.46}$$

Let us introduce the 2×2 block matrices

$$K = \begin{pmatrix} 0 & 0 \\ B^* & 0 \end{pmatrix} \quad , \quad S = \begin{pmatrix} \sigma & 0 \\ 0 & X \end{pmatrix} \ . \tag{7.47}$$

A simple calculation gives

$$\mathrm{tr}(B^* \sigma^{-\beta} B X^{1+\beta}) = \mathrm{tr}_{\mathcal{H} \otimes \mathbb{C}^2}(K^* S^{1+\beta} K S^{-\beta}) \ . \tag{7.48}$$

By Lemma 7.3.2, the right-hand side of (7.48) is concave (respectively convex) in S when $-1 \leq \beta < 0$ (respectively $0 < \beta \leq 1$). As a result, the left-hand side is jointly concave (convex) in (σ, X). But the maximum over X of a jointly concave function $f(\sigma, X)$ is concave in σ. Thanks to (7.46), we may conclude that $g_{B,\beta}(\sigma)$ is concave in σ for all $\beta \in [-1, 1]$, $\beta \neq 0$. The proof of Theorem 7.3.1 is now complete. □

Let us come back to the duality formula (7.41). We observe in passing that this formula bears some similarity with the variational formula (7.3) for the von Neumann entropy.

Proof of lemma 7.3.3. Since $\sigma^{-\frac{\beta}{2}} H \sigma^{-\frac{\beta}{2}}$ has the same non-zero eigenvalues as $\sqrt{H} \sigma^{-\beta} \sqrt{H}$, the quantity inside the infimum in (7.41) is equal to

$$g(H) = \alpha \, \mathrm{tr}(H\rho) + (1-\alpha) \, \mathrm{tr}\big[(\sigma^{-\frac{\beta}{2}} H \sigma^{-\frac{\beta}{2}})^{-\frac{1}{\beta}} \big] \ . \tag{7.49}$$

Differentiating the right-hand side with respect to the matrix elements of H in the some orthonormal basis $\{|i\rangle\}$ and using the relation $\partial \, \mathrm{tr}[f(B)] / \partial B_{ij} = f'(B)_{ji}$ with $f(x)$ a C^1-function, we get

$$\frac{\partial g(H)}{\partial H_{ij}} = \alpha \Big(\rho - \sigma^{-\frac{\beta}{2}} (\sigma^{-\frac{\beta}{2}} H \sigma^{-\frac{\beta}{2}})^{-\frac{1}{\beta}-1} \sigma^{-\frac{\beta}{2}} \Big)_{ji} \ . \tag{7.50}$$

Hence $g(H)$ has an extremum if and only if $H = \widehat{H} = \sigma^{\frac{\beta}{2}} (\sigma^{\frac{\beta}{2}} \rho \sigma^{\frac{\beta}{2}})^{\alpha-1} \sigma^{\frac{\beta}{2}} \geq 0$. But

$$g(\widehat{H}) = \mathrm{tr}[(\sigma^{\frac{\beta}{2}} \rho \sigma^{\frac{\beta}{2}})^{\alpha}] = F_\alpha(\rho || \sigma)^\alpha \ . \tag{7.51}$$

As $B \in \mathcal{B}(\mathcal{H})_+ \mapsto \mathrm{tr}(B^p)$ is convex for $p \geq 1$ or $p \leq 0$, $g(H)$ is convex if $\alpha \in (0,1)$ (i.e., $-\beta^{-1} < 0$) and concave if $\alpha > 1$ (i.e., $-\beta^{-1} > 1$). It follows that $g(\widehat{H})$ is a minimum for $\alpha \in (0,1)$ and a maximum for $\alpha > 1$. □

Let us point out that it follows from Lemma 7.3.2 that the normal-ordered Rényi entropy (7.29) is also jointly convex for $\alpha \in (0,1)$. Taking $\alpha \to 1$ and recalling that $S_\alpha^{(\mathrm{n})}(\rho||\sigma) \to S(\rho||\sigma)$, this gives a direct proof the joint convexity of the relative von Neumann entropy $S(\rho||\sigma)$ from the Lieb concavity theorem, as noted by Lindblad [157, 158]. Combined with Proposition 7.2.2, this leads to a completely different justification of the monotonicity of $S(\rho||\sigma)$ in Theorem 7.2.1 than that presented in Sec. 7.2. It would be interesting to look for a generalization of the arguments of Petz in Sec. 7.2 to the case of the α-entropies.

7.3.3 Monotonicity in α

As stated above, a very nice feature of the α-entropy (7.30) is that, like the classical Rényi divergence, it is monotonous in α. This leads in particular to some bound between the relative von Neumann entropy and the fidelity (see (8.32) below).

Proposition 7.3.4. [175] *For any $\rho, \sigma \in \mathcal{E}(\mathcal{H})$, $S_\alpha(\rho||\sigma)$ is a non-decreasing function of α on $(0, \infty)$.*

Proof. One first derive the following identity similar to (7.46):

$$\big(g_{B,-\alpha^{-1}}(\sigma)\big)^{\frac{1}{\alpha}} = \big\|B^*\sigma^{1/\alpha}B\big\|_\alpha = \sup_{\tau\geq 0, \mathrm{tr}(\tau)=1} \mathrm{tr}\big(B^*\sigma^{1/\alpha}B\,\tau^{1-1/\alpha}\big) \quad , \quad \alpha \geq 1\,. \tag{7.52}$$

If $0 < \alpha \leq 1$ the supremum has to be replaced by an infimum. When $\alpha \geq 1$ this identity is nothing but a rewriting of the Hölder's inequality (4.3). The derivation for $\alpha \in (0, 1)$ relies on (7.44) and follows the same lines as for the derivation of (7.46) (apart from the fact that we substituted β by $-1/\alpha$), but one must introduce a Lagrange multiplier to account for the constraint $\mathrm{tr}(\tau) = 1$. Applying the relation (7.52) for $B = \sigma^{-\frac{1}{2}}\rho^{\frac{1}{2}}$ and plugging the identity $\|\sigma^{\frac{\beta}{2}}\rho\sigma^{\frac{\beta}{2}}\|_\alpha = \|\rho^{\frac{1}{2}}\sigma^\beta\rho^{\frac{1}{2}}\|_\alpha$ into (7.36), we are led to

$$S_\alpha(\rho||\sigma) = \sup_{\tau\in\mathcal{E}(\mathcal{H})} \big\{-\beta^{-1}\ln F_\alpha(\rho||\sigma;\tau)\big\} \quad , \quad F_\alpha(\rho||\sigma;\tau) = \mathrm{tr}\big(\rho^{\frac{1}{2}}\sigma^\beta\rho^{\frac{1}{2}}\tau^{-\beta}\big) = \big\langle \xi\,,\, \Delta^\beta_{\sigma|\tau}\xi\big\rangle\,, \tag{7.53}$$

for any $\alpha > 0$, $\alpha \neq 1$. In the last identity $\xi = \rho^{\frac{1}{2}}$ and we have introduced the relative modular operator, see (4.8). For any fixed $\tau \in \mathcal{E}(\mathcal{H})$, one finds

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\Big(-\beta^{-1}\ln F_\alpha(\rho||\sigma;\tau)\Big) = -\frac{1}{\beta^2 F_\alpha(\rho||\sigma;\tau)}\Big(\langle\xi\,,\,\Delta^\beta_{\sigma|\tau}\ln(\Delta^\beta_{\sigma|\tau})\xi\rangle - \langle\xi\,,\,\Delta^\beta_{\sigma|\tau}\xi\rangle\ln\langle\xi\,,\,\Delta^\beta_{\sigma|\tau}\xi\rangle\Big)\,. \tag{7.54}$$

The Jensen inequality applied to the convex function $f(x) = x\ln x$ implies that the quantity inside the parenthesis in the right-hand side is non-negative. Thus $-\beta^{-1}F_\alpha(\rho||\sigma;\tau)$ is a non-increasing function of β. This is true for any density matrix τ, thus one infers from (7.53) that $\alpha \mapsto S_\alpha(\rho||\sigma)$ is non-decreasing. □

Chapter 8

The Bures distance and Uhlmann fidelity

As time goes on, it becomes increasingly evident that the rules which the mathematician finds interesting are the same as those which Nature has chosen (P.A. Dirac, 1939).

In this chapter we study the Bures distance on the set of quantum states $\mathcal{E}(\mathcal{H})$. This distance is Riemannian and monotonous with respect to quantum operations. It is a simple function of the fidelity (6.33). Its metric coincides with the quantum Fisher information quantifying the best achievable precision in the parameter estimation problem discussed in Sec. 9.2. The material of this chapter (as well as of chapter 9) is completely independent from that of chapters 10 and 11, so it is possible at this point to proceed directly to chapter 10. The reading of Secs. 8.1–8.4 is, however, recommended before going through chapter 12 devoted to the geometrical measures of quantum correlations, where the Bures distance plays the key role. The chapter is organized as follows. Sec. 8.1 contains a short discussion on contractive (i.e., monotonous) distances. It is argued there that the distances induced by the $\|\cdot\|_p$-norm are not contractive save for $p = 1$. The definition and main properties of the Bures distance are given in Secs. 8.2–8.4. The Bures metric is determined in Sec.8.5. Finally, Sec. 8.6 contains the proof of an important result of Petz on the characterization of all Riemannian contractive metrics on $\mathcal{E}(\mathcal{H})$ for finite-dimensional Hilbert spaces $\mathcal{H}$.

8.1 Contractive and convex distances

In order to quantify how far are two states ρ and σ it is necessary to define a distance on the set $\mathcal{E}(\mathcal{H})$ of quantum states. One has a priori the choice between many distances. The most common ones are the L^p-distances defined by (4.2). In quantum information theory it seems, however, natural to impose the following requirement.

Definition 8.1.1. *A distance d on the sets of quantum states is contractive if for any finite Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$, any quantum operation $\mathcal{M} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$, and any ρ, $\sigma \in \mathcal{E}(\mathcal{H})$, it holds*

$$d(\mathcal{M}(\rho), \mathcal{M}(\sigma)) \leq d(\rho, \sigma) \, . \tag{8.1}$$

A contractive distance is in particular invariant under unitary conjugations, i.e.

$$d\big(U\rho U^*, U\sigma U^*\big) = d(\rho, \sigma) \quad \text{if } U \text{ is unitary} \tag{8.2}$$

(in fact, $\rho \mapsto U\rho U^*$ is an invertible quantum operation on $\mathcal{B}(\mathcal{H})$). For such a distance, if a generalized measurement is performed on a system, two states are closer from each other after the measurement than before it, and if the system is subject to a unitary evolution the distance between the time-evolved states remains unchanged.

For $p > 1$, the distances d_p (in particular, the Hilbert-Schmidt distance d_2) are not contractive. A counter-example for two qubits is obtained [187] by taking $\mathcal{M}(\rho) = A_1\rho A_1^* + A_2\rho A_2^*$ with

$$A_1 = \sigma_+ \otimes 1 \quad , \quad A_2 = \sigma_+\sigma_- \otimes 1 \quad , \quad \rho = \frac{1}{2} \otimes \sigma_+\sigma_- \quad , \quad \sigma = \frac{1}{2} \otimes \sigma_-\sigma_+ \tag{8.3}$$

(here $\sigma_+ = |1\rangle\langle 0|$ is the raising operator and $\sigma_- = \sigma_+^*$). Then $\|\mathcal{M}(\rho) - \mathcal{M}(\sigma)\|_p = 2^{1/p}$ is larger than $\|\rho - \sigma\|_p = 2^{2/p-1}$.

Proposition 8.1.2. [205] *The trace distance d_1 is contractive.*

Proof. : Let $R = \rho - \sigma = R_+ - R_-$ with $R_\pm = (|R| \pm R)/2 = \pm RP_\pm \geq 0$ the positive and negative parts of R (here P_+ and P_- are the spectral projectors of R on $[0, \infty)$ and $(-\infty, 0)$). Then $\|R\|_1 = \mathrm{tr}(R_+ + R_-) = 2\,\mathrm{tr}(R_+)$ because $\mathrm{tr}(R) = \mathrm{tr}(R_+) - \mathrm{tr}(R_-) = 0$. Since $\mathcal{M}$ is trace preserving and CP, one has $\|\mathcal{M}(R)\|_1 = 2\,\mathrm{tr}[\mathcal{M}(R)_+]$ and $\mathcal{M}(R)_+ = (\mathcal{M}(R_+) - \mathcal{M}(R_-))_+ \leq \mathcal{M}(R_+)$. Thus $\|\mathcal{M}(R)\|_1 \leq 2\,\mathrm{tr}[\mathcal{M}(R_+)] = 2\,\mathrm{tr}[R_+] = \|R\|_1$. □

A distance d on $\mathcal{E}(\mathcal{H})$ is *jointly convex* if for any state ensembles $\{\rho_i, p_i\}$ and $\{\sigma_i, p_i\}$ with the same probabilities p_i,

$$d\Big(\sum_i p_i\rho_i, \sum_i p_i\sigma_i\Big) \leq \sum_i p_i d(\rho_i, \sigma_i)\ . \tag{8.4}$$

Since they are associated to a norm, the distances d_p are jointly convex for any $p \geq 1$.

8.2 The Bures distance

We now introduce the Bures distance d_{B}. This distance is contractive like d_1. It was first considered by Bures in the context of infinite products of von Neumann algebras [49] (see also [11]) and was later studied in a series of papers by Uhlmann [240, 242, 243]. Uhlmann used it to define parallel transport and related it to the fidelity generalizing the usual fidelity $|\langle\psi|\phi\rangle|^2$ between pure states. Indeed, d_{B} is a extension to mixed states of the Fubini-Study distance on the projective space $P\mathcal{H}$ of pure states,

$$d_{\mathrm{FS}}\big(\rho_\psi, \sigma_\phi\big) = \inf_{|\psi\rangle, |\phi\rangle} \big\||\psi\rangle - |\phi\rangle\big\| = \big(2 - 2|\langle\psi|\phi\rangle|\big)^{\frac{1}{2}}\ , \tag{8.5}$$

where the infimum in the second member is over all representatives $|\psi\rangle$ of $\rho_\psi \in P\mathcal{H}$ and $|\phi\rangle$ of $\sigma_\phi \in P\mathcal{H}$ (i.e., $\rho_\psi = |\psi\rangle\langle\psi|$ and $\sigma_\phi = |\phi\rangle\langle\phi|$). Observe that the third member is independent of these representatives. For two mixed states ρ and σ in $\mathcal{E}(\mathcal{H})$, one can define analogously [242, 132]

$$d_{\mathrm{B}}(\rho, \sigma) = \inf_{A,B} d_2(A - B)\ , \tag{8.6}$$

where the infimum is over all Hilbert-Schmidt matrices A and B satisfying $AA^* = \rho$ and $BB^* = \sigma$. Such matrices are given by $A = \sqrt{\rho}V$ and $B = \sqrt{\sigma}W$ for some unitaries V and W (polar decompositions). If $\rho = \rho_\psi$ and $\sigma = \sigma_\phi$ are pure states, then $A = |\psi\rangle\langle\mu|$ and $B = |\phi\rangle\langle\nu|$ with $\|\mu\| = \|\nu\| = 1$, so that (8.6) reduces to the Fubini-Study distance (8.5).

For mixed states ρ and σ, the right-hand side of (8.6) is given by

$$\big(2 - 2\sup_U \mathrm{Re}\ \mathrm{tr}(U\sqrt{\rho}\sqrt{\sigma})\big)^{\frac{1}{2}} \tag{8.7}$$

with a supremum over all unitaries $U = WV^*$. This supremum is equal to $\|\sqrt{\rho}\sqrt{\sigma}\|_1$ and is attained if and only if $UU_0|\sqrt{\rho}\sqrt{\sigma}|^{\frac{1}{2}} = |\sqrt{\rho}\sqrt{\sigma}|^{\frac{1}{2}}$, where U_0 is such that $\sqrt{\rho}\sqrt{\sigma} = U_0|\sqrt{\rho}\sqrt{\sigma}|$ (see Sec. 4.1). Equivalently, the infimum in (8.6) is attained if and only if the parallel transport condition $A^*B \geq 0$ holds. We obtain the following equivalent definition of d_{B}.

Definition 8.2.1. *For any states $\rho, \sigma \in \mathcal{E}(\mathcal{H})$,*

$$d_{\mathrm{B}}(\rho, \sigma) = \big(2 - 2\sqrt{F(\rho, \sigma)}\big)^{\frac{1}{2}} \tag{8.8}$$

where the Uhlmann fidelity is defined by

$$F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1^2 = \Big(\mathrm{tr}\big[(\sqrt{\sigma}\rho\sqrt{\sigma})^{\frac{1}{2}}\big]\Big)^2\ . \tag{8.9}$$

The fidelity $F(\rho,\sigma)$ is symmetric in (ρ,σ) and belongs to the interval $[0,1]$. It is clearly a generalization of the usual pure state fidelity $F(|\psi\rangle,|\phi\rangle) = |\langle\psi|\phi\rangle|^2$. If σ_ϕ is pure, then

$$F(\rho,\sigma_\phi) = \langle\phi|\rho|\phi\rangle \tag{8.10}$$

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

It is immediate on (8.6) that $d_{\rm B}$ is positive and symmetric, and $d_{\rm B}(\rho,\sigma) = 0$ if and only if $\rho = \sigma$. The triangle inequality is more difficult to show. It can be established with the help of the following astonishing theorem.

Theorem 8.2.2. (Uhlmann [240]) *Let ρ, $\sigma \in \mathcal{E}(\mathcal{H})$ and $|\Psi\rangle$ be a purification of ρ on the space $\mathcal{H}\otimes\mathcal{K}$, with $\dim\mathcal{K} \geq \dim\mathcal{H}$. Then*

$$F(\rho,\sigma) = \max_{|\Phi\rangle} |\langle\Psi|\Phi\rangle|^2 \tag{8.11}$$

where the maximum is over all purifications $|\Phi\rangle$ of σ on $\mathcal{H}\otimes\mathcal{K}$.

Proof. We give here a simple proof due to Josza [144]. Let us first assume $\mathcal{K} \simeq \mathcal{H}$. Let $|\Psi\rangle$ and $|\Phi\rangle$ be purifications of ρ and σ on $\mathcal{H}\otimes\mathcal{H}$, respectively. As it has been noticed in Sec. 4.3, by the Schmidt decomposition these purifications can always be written as

$$|\Psi\rangle = \sum_{k=1}^n \sqrt{p_k}|k\rangle|f_k\rangle \quad , \quad |\Phi\rangle = \sum_{k=1}^n \sqrt{q_k}(U|k\rangle)|g_k\rangle \ , \tag{8.12}$$

where $\rho = \sum_k p_k|k\rangle\langle k|$ and $\sigma = \sum_k q_k U|k\rangle\langle k|U^*$ are spectral decompositions of ρ and σ, U is a unitary operator on $\mathcal{H}$, and $\{|f_k\rangle\}_{i=1}^n$ and $\{|g_k\rangle\}_{i=1}^n$ are two orthonormal bases of $\mathcal{H}$. Defining the unitaries V and W on $\mathcal{H}$ by $|f_k\rangle = V|k\rangle$ and $|g_k\rangle = W|k\rangle$ for any $k = 1,\ldots,n$, we have

$$|\Psi\rangle = \sqrt{\rho}\otimes V|\Sigma\rangle \quad , \quad |\Phi\rangle = \sqrt{\sigma}\,U\otimes W|\Sigma\rangle \quad \text{with} \quad |\Sigma\rangle = \sum_{k=1}^n |k\rangle|k\rangle \ . \tag{8.13}$$

The vector $|\Sigma\rangle$ is the vector associated to the identity operator on $\mathcal{B}(\mathcal{H})$ by the isomorphism (4.5). For any $X, Y \in \mathcal{B}(\mathcal{H})$, one obtains by setting $O = X^T\otimes Y$ in (4.6) and noting that $\mathrm{tr}(O^{\mathcal{R}}) = \mathrm{tr}(XY)$ that

$$\mathrm{tr}(XY) = \langle\Sigma|X^T\otimes Y|\Sigma\rangle \tag{8.14}$$

(here X^T is the transpose of X in the basis $\{|k\rangle\}$). Introducing the unitary $U_0 = V^*WU^T$, this gives

$$\sup_{|\Phi\rangle} |\langle\Phi|\Psi\rangle| = \sup_W |\langle\Sigma|U^*\sqrt{\sigma}\sqrt{\rho}\otimes W^*V|\Sigma\rangle| = \sup_{U_0} |\,\mathrm{tr}(\sqrt{\rho}\sqrt{\sigma}\,U_0^*)| = \|\sqrt{\rho}\sqrt{\sigma}\|_1 \ . \tag{8.15}$$

The last equality comes from (4.3). This proves the desired result. The supremum is achieved by choosing $|\Phi\rangle$ as in (8.12) with $U = U_0^T(W^*)^TV^T$, U_0 being a unitary in the polar decomposition of $\sqrt{\rho}\sqrt{\sigma}$.

If $\mathcal{K}$ has a dimension m larger than n, we extend ρ and σ to a space $\mathcal{H}' \simeq \mathcal{K}$ by adding to them new orthonormal eigenvectors $|k\rangle$ and $U|k\rangle$ with zero eigenvalues $p_k = q_k = 0$, $k = n+1,\ldots,m$. This does not change the fidelity $F(\rho,\sigma)$, thus $F(\rho,\sigma) = \max_{|\Phi'\rangle} |\langle\Psi'|\Phi'\rangle|^2$, where $|\Psi'\rangle$ is a purification of $\rho' = \sum_{k=1}^m p_k|k\rangle\langle k| = \rho$ on $\mathcal{H}'\otimes\mathcal{H}'$, and similarly for $|\Phi'\rangle$. But $|\Psi'\rangle$ and $|\Phi'\rangle$ have the form (8.12), hence they belong to $\mathcal{H}\otimes\mathcal{K}$. □

Let ρ, σ, and τ be three states of $\mathcal{E}(\mathcal{H})$ and $|\Psi\rangle$ be a purification of ρ on $\mathcal{H}\otimes\mathcal{H}$. According to Theorem 8.2.2, there exists a purification $|\Phi\rangle$ of σ on $\mathcal{H}\otimes\mathcal{H}$ such that $F(\rho,\sigma) = |\langle\Psi|\Phi\rangle|^2$. One can choose the arbitrary phase factor of $|\Phi\rangle$ in such a way that $\langle\Psi|\Phi\rangle \geq 0$, whence $\sqrt{F(\rho,\sigma)} = \langle\Psi|\Phi\rangle$. Similarly, there exists a purification $|\chi\rangle$ of τ such that $\sqrt{F(\sigma,\tau)} = \langle\Phi|\chi\rangle \geq 0$. In view of (8.8) and (8.11),

$$\begin{aligned} d_{\rm B}(\rho,\tau) &\leq \big(2-2|\langle\Psi|\chi\rangle|\big)^{\frac12} \\ &\leq \big(2-2\mathrm{Re}\,\langle\Psi|\chi\rangle|\big)^{\frac12} = \big\||\Psi\rangle-|\chi\rangle\big\| \\ &\leq \big\||\Psi\rangle-|\Phi\rangle\big\| + \big\||\Phi\rangle-|\chi\rangle\big\| = \big(2-2\langle\Psi|\Phi\rangle|\big)^{\frac12} + \big(2-2\langle\Phi|\chi\rangle|\big)^{\frac12} \ , \end{aligned} \tag{8.16}$$

showing that $d_{\rm B}$ satisfies the triangle inequality $d_{\rm B}(\rho,\tau) \leq d_{\rm B}(\rho,\sigma) + d_{\rm B}(\sigma,\tau)$.

Corollary 8.2.3. *The map $(\rho,\sigma)\mapsto d_{\rm B}(\rho,\sigma)$ defines a distance $d_{\rm B}$ on quantum states, with values in $[0,\sqrt{2}]$. This distance is contractive. Moreover, $d_{\rm B}^2$ is jointly convex.*

Note that $d_{\rm B}$ is not jointly convex. One gets a counter-example by choosing $\rho_0=\sigma_0=|0\rangle\langle 0|$, $\rho_1=|1\rangle\langle 1|$, $\sigma_1=|2\rangle\langle 2|$, and $p_0=p_1=1/2$, $\{|0\rangle,|1\rangle,|2\rangle\}$ being an orthonormal family in $\mathcal{H}$.

It is clear on (8.9) that $F(\rho,\sigma)=0$ if and only if ρ and σ have orthogonal supports, $\operatorname{ran}\rho\perp\operatorname{ran}\sigma$. Therefore, two states ρ and σ have a maximal distance $d_{\rm B}(\rho,\sigma)=1$ if they are orthogonal and thus perfectly distinguishable.

Proof. We have already established above that $d_{\rm B}$ satisfies all the axioms of a distance. To show the contractivity, it is enough to check that for any quantum operation $\mathcal{M}:\mathcal{B}(\mathcal{H})\to\mathcal{B}(\mathcal{H}')$ and any states $\rho,\sigma\in\mathcal{E}(\mathcal{H})$,

$$F(\mathcal{M}(\rho),\mathcal{M}(\sigma))\geq F(\rho,\sigma)\;. \tag{8.17}$$

This property of the fidelity is a consequence of the contractivity of the relative Rényi entropy for $\alpha=1/2$ (Theorem 7.3.1*(v)*). It is, however, instructive to re-derive this result from Theorem 8.2.2. According to this theorem, there exist some purifications $|\Psi\rangle$ and $|\Phi\rangle$ of ρ and σ on $\mathcal{H}\otimes\mathcal{K}$ such that $F(\rho,\sigma)=|\langle\Psi|\Phi\rangle|^2$. Now, thanks to (5.13) one obtains some purifications $|\Psi_{\mathcal{M}}\rangle=1_{\mathcal{K}}\otimes U|\Psi\rangle|\epsilon_0\rangle$ of $\mathcal{M}(\rho)$ and $|\Phi_{\mathcal{M}}\rangle=1_{\mathcal{K}}\otimes U|\Phi\rangle|\epsilon_0\rangle$ of $\mathcal{M}(\sigma)$ on $\mathcal{K}\otimes\mathcal{H}'\otimes\mathcal{H}'_{\mathsf{E}}$, with $|\epsilon_0\rangle\in\mathcal{H}_{\mathsf{E}}$ and $U:\mathcal{H}\otimes\mathcal{H}_{\mathsf{E}}\to\mathcal{H}'\otimes\mathcal{H}'_{\mathsf{E}}$ unitary. Thus

$$F(\mathcal{M}(\rho),\mathcal{M}(\sigma))\geq|\langle\Psi_{\mathcal{M}}|\Phi_{\mathcal{M}}\rangle|^2=|\langle\Psi|\Phi\rangle|^2=F(\rho,\sigma)\;. \tag{8.18}$$

The joint convexity of $d_{\rm B}^2$ is a consequence of the bound[1]

$$\sqrt{F\Big(\sum_i p_i\rho_i,\sum_i q_i\sigma_i\Big)}\geq\sum_i\sqrt{p_iq_i}\sqrt{F(\rho_i,\sigma_i)}\;, \tag{8.19}$$

where $\{\rho_i,p_i\}$ and $\{\sigma_i,q_i\}$ are arbitrary ensembles in $\mathcal{E}(\mathcal{H})$. Note that the statement (8.19) is slightly more general than the joint concavity of $\sqrt{F(\rho,\sigma)}$ proven in Sec. 7.3 (Theorem 7.3.1*(iv)*). To show that (8.19) is true, we introduce as before some purifications $|\Psi_i\rangle$ of ρ_i and $|\Phi_i\rangle$ of σ_i on $\mathcal{H}\otimes\mathcal{H}$ such that $\sqrt{F(\rho_i,\sigma_i)}=\langle\Psi_i|\Phi_i\rangle$. Let us define the vectors

$$|\Psi\rangle=\sum_i\sqrt{p_i}|\Psi_i\rangle|\epsilon_i\rangle\quad,\quad|\Phi\rangle=\sum_i\sqrt{p_i}|\Phi_i\rangle|\epsilon_i\rangle \tag{8.20}$$

in $\mathcal{H}\otimes\mathcal{H}\otimes\mathcal{H}_{\mathsf{E}}$, where $\mathcal{H}_{\mathsf{E}}$ is an auxiliary Hilbert space and $\{|\epsilon_i\rangle\}$ is an orthonormal basis of $\mathcal{H}_{\mathsf{E}}$. Then $|\Psi\rangle$ and $|\Phi\rangle$ are purifications of $\rho=\sum_i p_i\rho_i$ and $\sigma=\sum_i q_i\sigma_i$, respectively. One infers from Theorem 8.2.2 that

$$\sqrt{F(\rho,\sigma)}\geq|\langle\Psi|\Phi\rangle|=\sum_i\sqrt{p_iq_i}\langle\Psi_i|\Phi_i\rangle=\sum_i\sqrt{p_iq_i}\sqrt{F(\rho_i,\sigma_i)}\;. \tag{8.21}$$

This complete the proof of the corollary. □

Remark 8.2.4. *A consequence of (5.26) and (8.10) and of the monotonicity of the fidelity F with respect to partial trace operations (see (8.17)) is that the entanglement fidelity $F_{\rm e}(\rho,\mathcal{M})$ of a state ρ with respect to a quantum operation $\mathcal{M}$ satisfies*

$$F_{\rm e}(\rho,\mathcal{M})\leq F(\rho,\mathcal{M}(\rho))\;. \tag{8.22}$$

Remark 8.2.5. *As the fidelity satisfies $F(\rho\otimes\rho',\sigma\otimes\sigma')=F(\rho,\sigma)F(\rho',\sigma')$, the Bures distance increases by taking tensor products, $d_{\rm B}(\rho\otimes\rho',\sigma\otimes\sigma')\geq d_{\rm B}(\rho,\sigma)$ for any $\rho,\sigma\in\mathcal{E}(\mathcal{H})$, $\rho',\sigma'\in\mathcal{E}(\mathcal{H}')$, with equality if and only if $\rho'=\sigma'$. This has to be contrasted with the trace distance, which does not enjoy this property.*

In the two following sections we collect some important properties of the Bures distance. We refer the reader to the monographs [31, 180] for a list of names to which these properties should be attached.

[1]Note that one cannot replace $\sqrt{F}$ by F in this inequality, that is, $F(\rho,\sigma)$ is not jointly concave (one can take the same counter-example as that given above for $d_{\rm B}$). However, by a slight modification of the proof of Corollary 8.2.3 one can show that $\rho\mapsto F(\rho,\sigma)$ and $\sigma\mapsto F(\rho,\sigma)$ are concave. In their book [180], Nielsen and Chuang define the fidelity as the square root of (8.9). This must be kept in mind when comparing the results in this monograph with those of this article.

8.3 Bures distance and statistical distance in classical probability

The restriction of a distance d on $\mathcal{E}(\mathcal{H})$ to all density matrices commuting with a given state ρ_0 defines a distance on the simplex $\mathcal{E}_{\rm clas} = \{\mathbf{p} \in \mathbb{R}^n_+; \sum_i p_i = 1\}$ of classical probabilities on the finite space $\{1, 2, \ldots, n\}$. In particular, if ρ and σ are two commuting states with spectral decompositions $\rho = \sum_k p_k |k\rangle\langle k|$ and $\sigma = \sum_k q_k |k\rangle\langle k|$, then

$$d_1(\rho, \sigma) = d_1^{\rm clas}(\mathbf{p}, \mathbf{q}) = \sum_{k=1}^n |p_k - q_k|$$

is the ℓ^1-distance, and

$$d_{\rm B}(\rho, \sigma) = d_{\rm H}^{\rm clas}(\mathbf{p}, \mathbf{q}) = \Big(\sum_{k=1}^n (\sqrt{p_k} - \sqrt{q_k})^2 \Big)^{\frac{1}{2}} = \Big(2 - 2 \sum_{k=1}^n \sqrt{p_k q_k} \Big)^{\frac{1}{2}} \tag{8.23}$$

is the Hellinger distance. A distance closely related to $d_{\rm H}^{\rm clas}$ is the so-called statistical distance $\Theta^{\rm clas}(\mathbf{p}, \mathbf{q}) = \arccos(1 - d_{\rm H}^{\rm clas}(\mathbf{p}, \mathbf{q})^2/2)$, i.e., the angle between the vectors $\mathbf{x} = (\sqrt{p_k})_{k=1}^n$ and $\mathbf{y} = (\sqrt{q_k})_{k=1}^n$ on the unit sphere. Given two non-commuting states ρ and σ, one can consider the distance $d^{\rm clas}(\mathbf{p}, \mathbf{q})$ between the outcome probabilities $\mathbf{p}$ and $\mathbf{q}$ of a measurement performed on the system in states ρ and σ, respectively. It is natural to ask whether there is a relation between $d(\rho, \sigma)$ and the supremum of $d^{\rm clas}(\mathbf{p}, \mathbf{q})$ over all measurements.

Proposition 8.3.1. *For any $\rho, \sigma \in \mathcal{E}(\mathcal{H})$,*

$$d_1(\rho, \sigma) = \sup_{\{M_i\}} d_1^{\rm clas}(\mathbf{p}, \mathbf{q}) \quad , \quad d_{\rm B}(\rho, \sigma) = \sup_{\{M_i\}} d_{\rm H}^{\rm clas}(\mathbf{p}, \mathbf{q}) \, , \tag{8.24}$$

where the suprema are over all POVMs $\{M_i\}$ and $p_i = {\rm tr}(M_i \rho)$ (respectively $q_i = {\rm tr}(M_i \sigma)$) is the probability of the measurement outcome i in the state ρ (respectively σ). Moreover, the suprema are achieved for von Neumann measurements with rank-one projectors $M_i = |i\rangle\langle i|$.

Proof. We leave the justification of the first identity to the reader. It can be obtained by following similar arguments as in the proof of Proposition 8.1.2 (see [180]). Let us show the second identity. Given a POVM $\{M_i\}$, by taking advantage of the definition (8.9) of the fidelity, the polar decomposition $\sqrt{\rho}\sqrt{\sigma} = U|\sqrt{\rho}\sqrt{\sigma}|$, and the identity $\sum_i M_i = 1$, one gets

$$\sqrt{F(\rho, \sigma)} = \sum_i {\rm tr}(U^* \sqrt{\rho} \sqrt{M_i} \sqrt{M_i} \sqrt{\sigma}) \; \leq \; \sum_i \sqrt{p_i q_i} \, . \tag{8.25}$$

The upper bound comes from the Cauchy-Schwarz inequality. It remains to show that this bound can be attained for an appropriate choice of POVM. The Cauchy-Schwarz inequality holds with equality if and only if $\sqrt{M_i}\sqrt{\rho}U = \lambda_i \sqrt{M_i}\sqrt{\sigma}$ with $\lambda_i \in \mathbb{C}$. Assuming $\sigma > 0$ and observing that $\sqrt{\rho}U = \sigma^{-\frac{1}{2}}|\sqrt{\rho}\sqrt{\sigma}|$, this identity can be recast as

$$\sqrt{M_i}(R - \lambda_i) = 0 \quad \text{ with } \quad R = \sigma^{-\frac{1}{2}} |\sqrt{\rho}\sqrt{\sigma}| \sigma^{-\frac{1}{2}} \, . \tag{8.26}$$

Let $R = \sum_i r_i |i\rangle\langle i|$ be a spectral projection of the non-negative matrix R. Taking M_i to be the von Neumann projector $M_i = |i\rangle\langle i|$ and $\lambda_i = r_i$, we find that (8.26) is satisfied for all i. Thus $\sqrt{F(\rho, \sigma)}$ is equal to the right-hand side of (8.25). If σ is not invertible it can be approached by invertible density matrices $\sigma_\varepsilon = (1 - \varepsilon)\sigma + \varepsilon$, $\varepsilon > 0$, and the result follows by continuity. □

Much as for the quantum relative Rényi entropies (Sec. 7.3), one may define another distance on $\mathcal{E}(\mathcal{H})$ which also reduces to the Hellinger distance $d_{\rm H}^{\rm clas}$ for commuting matrices, by setting

$$d_{\rm H}(\rho, \sigma) = d_2(\sqrt{\rho}, \sqrt{\sigma}) = \Big(2 - 2\sqrt{F^{(\rm n)}_{\frac{1}{2}}(\rho || \sigma)} \Big)^{\frac{1}{2}} \, , \tag{8.27}$$

where $F^{(\rm n)}_\alpha(\rho||\sigma)$ is the fidelity associated to the normal-ordered α-entropy (7.29), namely,

$$F^{(\rm n)}_\alpha(\rho||\sigma) = \Big({\rm tr}\big[\rho^\alpha \sigma^{1-\alpha}\big] \Big)^{\frac{1}{\alpha}} = {\rm e}^{-\beta S^{(\rm n)}_\alpha(\rho||\sigma)} \quad , \quad \beta = \frac{1-\alpha}{\alpha} \, . \tag{8.28}$$

This distance is sometimes called the quantum Hellinger distance. Thanks to Lieb's concavity theorem (Lemma 7.3.2), $F^{(\rm n)}_\alpha(\rho||\sigma)^\alpha$ is jointly concave in (ρ, σ) for all $\alpha \in (0, 1)$. Consequently, the square Hellinger distance $d_{\rm H}(\rho, \sigma)^2$ is jointly convex, just as $d_{\rm B}(\rho, \sigma)^2$. From Proposition 7.2.2 one then deduces that $d_{\rm H}$ is contractive. It is worth noting that $d_{\rm H}$ does not coincide with the Fubini-study distance (8.5) for pure states (in fact, one finds $F^{(\rm n)}_{1/2}(\rho_\psi||\sigma_\phi) = |\langle\psi|\phi\rangle|^4$). For any $\rho, \sigma \in \mathcal{E}(\mathcal{H})$, one finds by comparing (8.6) and (8.27) that $d_{\rm B}(\rho, \sigma) \leq d_{\rm H}(\rho, \sigma)$.

8.4 Comparison of the Bures and trace distances

The next result shows that the Bures and trace distances $d_{\rm B}$ and d_1 are equivalent and gives optimal bounds of d_1 in terms of $d_{\rm B}$.

Proposition 8.4.1. *For any $\rho, \sigma \in \mathcal{E}(\mathcal{H})$, one has*

$$d_{\rm B}(\rho,\sigma)^2 \leq d_1(\rho,\sigma) \leq 2\Big\{1 - \Big(1 - \frac{1}{2} d_{\rm B}(\rho,\sigma)^2\Big)^2\Big\}^{\frac{1}{2}} . \tag{8.29}$$

The lower bound has been first proven by Araki [11] in the C^*-algebra setting. We shall justify it from Proposition 8.3.1 as in Ref. [180]. The upper bound is saturated for pure states, as shown in the proof below. Note that this bound implies that $d_1(\rho,\sigma) \leq 2 d_{\rm B}(\rho,\sigma)$.

Proof. We first argue that if $\rho_\psi = |\psi\rangle\langle\psi|$ and $\sigma_\phi = |\phi\rangle\langle\phi|$ are pure states, then $d_1(\rho_\psi,\sigma_\phi) = 2\sqrt{1 - F(\rho_\psi,\sigma_\phi)}$ and thus the upper bound in (8.29) is an equality. Actually, let $|\phi\rangle = \cos\theta|\psi\rangle + {\rm e}^{{\rm i}\delta}\sin\theta|\psi^\perp\rangle$, where $\theta, \delta \in [0, 2\pi)$ and $|\psi^\perp\rangle$ is a unit vector orthogonal to $|\psi\rangle$. Since $\rho_\psi - \sigma_\phi$ has non-vanishing eigenvalues $\pm\sin\theta$, one has $d_1(\rho_\psi,\sigma_\phi) = 2|\sin\theta|$. But $F(\rho_\psi,\sigma_\phi) = \cos^2\theta$, hence the aforementioned statement is true. It then follows from Theorem 8.2.2 and from the contractivity of the trace distance with respect to partial trace operations (Proposition 8.1.2) that for arbitrary ρ and $\sigma \in \mathcal{E}(\mathcal{H})$,

$$d_1(\rho,\sigma) \leq 2\sqrt{1 - F(\rho,\sigma)} . \tag{8.30}$$

To bound $d_1(\rho,\sigma)$ from below, we use Proposition 8.3.1 and consider a generalized measurement $\{M_i\}$ such that $\sqrt{F(\rho,\sigma)} = \sum_i \sqrt{p_i q_i}$ with $p_i = {\rm tr}(\rho M_i)$ and $q_i = {\rm tr}(\sigma M_i)$. This yields

$$d_{\rm B}(\rho,\sigma)^2 = \sum_i (\sqrt{p_i} - \sqrt{q_i})^2 \leq \sum_i |p_i - q_i| \leq d_1(\rho,\sigma) , \tag{8.31}$$

where the last inequality comes from Proposition 8.3.1 again. □

The following bound on the relative entropy can be obtained from (7.36), (7.32), and Proposition 7.3.4

$$S(\rho||\sigma) \geq -2\ln\Big(1 - \frac{1}{2} d_{\rm B}(\rho,\sigma)^2\Big) \geq -\ln\Big(1 - \frac{1}{4} d_1(\rho,\sigma)^2\Big) . \tag{8.32}$$

Remark 8.4.2. *By taking advantage of the inequality $F(\rho,\sigma) \geq {\rm tr}(\rho\sigma)$, which follows from (8.9) and the norm inequality $\|A\|_1 \geq \|A\|_2$, one can establish another bound on $S(\rho||\sigma)$ in terms of the fidelity, which reads [231]*

$$S(\rho||\sigma) \geq -S(\rho) - \ln F(\rho,\sigma) . \tag{8.33}$$

Remark 8.4.3. *The formula*

$$F(\rho,\sigma) = \frac{1}{4} \inf_{H>0} \big\{ {\rm tr}(H\rho) + {\rm tr}(H^{-1}\sigma) \big\}^2 = \inf_{H>0} \big\{ {\rm tr}(H\rho)\, {\rm tr}(H^{-1}\sigma) \big\} \tag{8.34}$$

can be easily proven with the help of Lemma 7.3.3 and Theorem 8.2.2. The last expression is due to Alberti [2].

Remark 8.4.4. *We are now in position to show without much effort several results of Sec. 6.2.*

(a) The upper bound (6.15) on the optimal success probability $P_{\rm S,u}^{\rm opt}$ in unambiguous discrimination of two mixed states can be established from Uhlmann's theorem, formula (6.8), and the fact that $P_{\rm S,u}^{\rm opt}(\{\rho_i,\eta_i\}) \leq P_{\rm S,u}^{\rm opt}(\{|\Psi_i\rangle,\eta_i\})$, where $|\Psi_i\rangle$ is a purification of ρ_i for any i [204].

(b) It is instructive to derive in the special case of $m = 2$ states the lower bound on $P_{\rm S,a}^{\rm opt}$ given in Proposition 6.5.1 by using the Helstrom formula (6.6), the fact that ${\rm tr}(|\Lambda|) \geq \sum_i |\langle i|\Lambda|i\rangle|$ for any orthonormal basis $\{|i\rangle\}$, and Proposition 8.3.1 [37].

(c) The Uhlmann theorem gives an efficient way to calculate the fidelity between the two states (6.16) (the result is $F(\rho_{\rm eq},\rho_{\rm diff}) = |\langle\psi_1|\psi_2\rangle|^2$).

8.5 Bures and quantum Hellinger metrics, quantum Fisher information

Recall that a Riemannian metric on $\mathcal{E}(\mathcal{H})$ is a map g which associates to each $\rho \in \mathcal{E}(\mathcal{H})$ a scalar product g_ρ on the tangent space to $\mathcal{E}(\mathcal{H})$ at ρ. For any state ρ on $\mathcal{H}$, this tangent space can be identified with the (real) vector space $\mathcal{B}(\mathcal{H})_{\text{s.a.}}$ of self-adjoint operators on $\mathcal{H}$. A metric g defines a Riemannian distance d, which is such that the square distance $\mathrm{d}s^2 = d(\rho, \rho + \mathrm{d}\rho)^2$ between two infinitesimally close states ρ and $\rho + \mathrm{d}\rho$ is given by

$$\mathrm{d}s^2 = g_\rho(\mathrm{d}\rho, \mathrm{d}\rho) \; . \tag{8.35}$$

The Hilbert-Schmidt distance d_2 is obviously Riemannian: its metric is constant and given by the scalar product (4.1). In contrast, the trace distance d_1 is not Riemannian.

Let us show that the Bures distance d_{B} is Riemannian and determine its metric g_{B}. It is convenient to introduce a small parameter $t \in \mathbb{R}$. According to Definition 8.2.1 one has

$$d_{\mathrm{B}}(\rho, \rho + t\,\mathrm{d}\rho)^2 = 2 - 2\,\mathrm{tr}(A(t)) \quad , \quad A(t) = \big(\sqrt{\rho}(\rho + t\mathrm{d}\rho)\sqrt{\rho}\big)^{\frac{1}{2}} \; . \tag{8.36}$$

The scalar product $(g_{\mathrm{B}})_\rho$ will be given in terms of the eigenvectors $|k\rangle$ and eigenvalues p_k of ρ in the spectral decomposition $\rho = \sum_k p_k |k\rangle\langle k|$. Using the notation $\dot{A}(t) = \mathrm{d}A/\mathrm{d}t$, $\ddot{A}(t) = \mathrm{d}^2A/\mathrm{d}t^2$, and the identity $A(t)^2 = \sqrt{\rho}(\rho + t\mathrm{d}\rho)\sqrt{\rho}$, one finds

$$\begin{aligned} \dot{A}(0)A(0) + A(0)\dot{A}(0) &= \sqrt{\rho}\,\mathrm{d}\rho\,\sqrt{\rho} \\ \ddot{A}(0)A(0) + 2\dot{A}(0)\dot{A}(0) + A(0)\ddot{A}(0) &= 0 \end{aligned} \tag{8.37}$$

The first equation yields

$$(p_k + p_l)\langle k|\dot{A}(0)|l\rangle = \sqrt{p_k p_l}\langle k|\mathrm{d}\rho|l\rangle \; . \tag{8.38}$$

Since $\mathrm{tr}(\mathrm{d}\rho) = 0$, it follows that $\mathrm{tr}[\dot{A}(0)] = 0$. Assume that $A(0) = \rho$ is invertible. Multiplying the second equation in (8.37) by $A(0)^{-1}$ and taking the trace, one verifies that

$$\mathrm{tr}[\ddot{A}(0)] = -\,\mathrm{tr}\big[\dot{A}(0)^2 A(0)^{-1}\big] = -\sum_{k,l=1}^{n} p_k^{-1}\big|\langle k|\dot{A}(0)|l\rangle\big|^2 = -\sum_{k,l=1}^{n} \frac{p_l|\langle k|\mathrm{d}\rho|l\rangle|^2}{(p_k + p_l)^2} \; . \tag{8.39}$$

Thus, going back to (8.36) we arrive at

$$d_{\mathrm{B}}(\rho, \rho + t\mathrm{d}\rho)^2 = -\,\mathrm{tr}[\ddot{A}(0)]t^2 + \mathcal{O}(t^3) = (g_{\mathrm{B}})_\rho(\mathrm{d}\rho, \mathrm{d}\rho)t^2 + \mathcal{O}(t^3) \tag{8.40}$$

with [132]

$$(g_{\mathrm{B}})_\rho(A, A) = \frac{1}{2}\sum_{k,l=1}^{n} \frac{|\langle k|A|l\rangle|^2}{p_k + p_l} \quad , \quad A \in \mathcal{B}(\mathcal{H})_{\text{s.a.}} \; , \; \rho > 0 \; . \tag{8.41}$$

The last formula defines a scalar product on $\mathcal{B}(\mathcal{H})_{\text{s.a.}}$ by polarization, hence d_{B} is Riemannian with metric g_{B}. One readily obtains from this metric the infinitesimal volume element. The volume of $\mathcal{E}(\mathcal{H})$ and the area of its boundary are determined in [220].

Definition 8.5.1. *Given a state $\rho \in \mathcal{E}(\mathcal{H})$ and an observable $H \in \mathcal{B}(\mathcal{H})_{\text{s.a.}}$, the non-negative number*

$$\mathcal{F}_Q(\rho, H) = 4(g_{\mathrm{B}})_\rho\big(-\mathrm{i}[H, \rho], -\mathrm{i}[H, \rho]\big) = 2\sum_{k,l, p_k + p_l > 0} \frac{(p_k - p_l)^2}{p_k + p_l}|\langle k|H|l\rangle|^2 \tag{8.42}$$

is called the quantum Fisher information of ρ with respect to H.

The quantity $\mathcal{F}_Q(\rho, H)$ has been introduced by Braunstein and Caves [45] as a quantum analog of the Fisher information in statistics. Similarly to the definition of the Bures distance in Sec. 8.2, these authors related it to the metric – called the "distinguishability metric" by Wootters [264] – extending the Fubini-Study metric to mixed states. For a pure state $\rho_\Psi = |\Psi\rangle\langle\Psi|$, the quantum Fisher information reduces to the square quantum fluctuation of H, namely,

$$\mathcal{F}_Q(\rho_\Psi, H) = 4\langle(\Delta H)^2\rangle_\Psi = 4\big(\langle\Psi|H^2|\Psi\rangle - \langle\Psi|H|\Psi\rangle^2\big) \; . \tag{8.43}$$

In general, $\sqrt{\mathcal{F}_Q(\rho,H)}$ gives the speed at which a given state ρ separates from its time-evolved state $\rho(t) = \mathrm{e}^{-\mathrm{i}tH}\rho\mathrm{e}^{\mathrm{i}tH}$ under the dynamics specified by the Hamiltonian H. In fact, by plugging $\mathrm{d}\rho/\mathrm{d}t = -\mathrm{i}[H,\rho]$ into (8.40) one checks that

$$\sqrt{\mathcal{F}_Q(\rho,H)} = \left(2\frac{\mathrm{d}^2}{\mathrm{d}t^2}\, d_{\mathrm{B}}(\rho,\rho(t))^2\Big|_{t=0}\right)^{\frac{1}{2}} \approx \sqrt{2}\frac{\delta d_{\mathrm{B}}}{\delta t}\ . \tag{8.44}$$

We postpone the discussion on the statistical interpretation of $\mathcal{F}_Q(\rho,H)$ to Sec. 9.2 below. It will be argued there that $\mathcal{F}_Q(\rho,H)$ measures the amount of quantum correlations in the state ρ that can be used for improving precision in quantum metrology.

Let us now turn to the quantum Hellinger distance (8.27). We proceed to determine the metric g_α associated to the normal-ordered relative Rényi entropy (7.29), from which the quantum Hellinger metric g_{H} is obtained by setting $\alpha = 1/2$. We demonstrate that the largest metric g_α for $\alpha \in (0,1)$ is achieved for $\alpha = 1/2$ and is equal to $g_{\mathrm{H}}/2$, a result that will be needed later on (Sec. 9.1). The metric g_α is defined by

$$\begin{aligned} S_\alpha^{(\mathrm{n})}(\rho+t\mathrm{d}\rho||\rho) &= (1-\alpha)^{-1}\big(1-F_\alpha^{(\mathrm{n})}(\rho+t\mathrm{d}\rho||\rho)^\alpha\big) + \mathcal{O}(t^3) \\ &= t^2(1-\alpha)^{-1}(g_\alpha)_\rho(\mathrm{d}\rho,\mathrm{d}\rho) + \mathcal{O}(t^3)\ , \end{aligned} \tag{8.45}$$

where $F_\alpha^{(\mathrm{n})}$ is the α-fidelity, see (8.28). To determine g_α for all $\alpha \in (0,1)$, we use (1) in Appendix A to write

$$\begin{aligned} B_\alpha(t) &= \rho^\alpha - (\rho+t\mathrm{d}\rho)^\alpha = \frac{\sin(\alpha\pi)}{\pi}\int_0^\infty \mathrm{d}x\, x^\alpha\left(\frac{1}{x+\rho+t\mathrm{d}\rho} - \frac{1}{x+\rho}\right) \\ &= \frac{\sin(\alpha\pi)}{\pi}\int_0^\infty \mathrm{d}x\ x^\alpha\left(-\frac{t}{x+\rho}\mathrm{d}\rho\frac{1}{x+\rho} + \frac{t^2}{x+\rho}\mathrm{d}\rho\frac{1}{x+\rho}\mathrm{d}\rho\frac{1}{x+\rho}\right) + \mathcal{O}(t^3)\ . \end{aligned} \tag{8.46}$$

Introducing as before the spectral decomposition $\rho = \sum_k p_k|k\rangle\langle k|$ and using known integrals, one finds

$$\begin{aligned} 1 - F_\alpha^{(\mathrm{n})}(\rho+t\mathrm{d}\rho||\rho)^\alpha &= \mathrm{tr}[B_\alpha(t)\rho^{1-\alpha}] \\ &= -t\alpha\sum_{k=1}^n\langle k|\mathrm{d}\rho|k\rangle + t^2\sum_{k,l=1}^n \frac{p_k^{1-\alpha}(p_k^\alpha - p_l^\alpha)}{(p_k-p_l)^2}\big|\langle k|\mathrm{d}\rho|l\rangle\big|^2 + \mathcal{O}(t^3)\ . \end{aligned} \tag{8.47}$$

Because $\mathrm{tr}(\mathrm{d}\rho) = 0$, the linear term in t vanishes as it should be. Plugging (8.47) into (8.45) one gets

$$(g_\alpha)_\rho(A,A) = \sum_{k,l=1}^n c_\alpha(p_k,p_l)|\langle k|A|l\rangle|^2 \quad , \quad c_\alpha(p,q) = \frac{(p^{1-\alpha}-q^{1-\alpha})(p^\alpha - q^\alpha)}{2(p-q)^2}\ . \tag{8.48}$$

It is easy to show that $c_\alpha(p,q) \leq c_{1/2}(p,q)$ for any $p,q>0$, hence

$$\max_{\alpha\in(0,1)}(g_\alpha)_\rho(A,A) = \big(g_{\frac{1}{2}}\big)_\rho(A,A) = \sum_{k,l=1}^n \frac{|\langle k|A|l\rangle|^2}{2(\sqrt{p_k}+\sqrt{p_l})^2} \quad , \quad A\in\mathcal{B}(\mathcal{H})_{\mathrm{s.a.}}\ , \tag{8.49}$$

as claimed above. Furthermore, in view of (8.27) we deduce that the quantum Hellinger distance d_{H} is Riemannian and has a metric $g_{\mathrm{H}} = 2g_{1/2}$.

8.6 Characterization of the Riemannian contractive distances

The complete characterization of Riemannian contractive distances on $\mathcal{E}(\mathcal{H})$ for finite Hilbert spaces $\mathcal{H}$ has been given by Petz [195], following a work by Morozova and Chentsov [174]. Such distances are induced by metrics g satisfying

$$g_{\mathcal{M}(\rho)}\big(\mathcal{M}(A),\mathcal{M}(A)\big) \leq g_\rho(A,A) \quad , \quad A\in\mathcal{B}(\mathcal{H})_{\mathrm{s.a.}}\ , \tag{8.50}$$

for any $\rho\in\mathcal{E}(\mathcal{H})$ and any quantum operation $\mathcal{M}:\mathcal{B}(\mathcal{H})\to\mathcal{B}(\mathcal{H}')$.

In the classical setting, it is remarkable that the contractivity condition leads to a unique metric (up to a multiplicative constant). Quantum operations correspond classically to Markov mappings $\mathbf{p}\mapsto\mathcal{M}^{\mathrm{clas}}\mathbf{p}$ on the probability simplex $\mathcal{E}_{\mathrm{clas}} = \{\mathbf{p}\in\mathbb{R}_+^n; \sum_i p_i = 1\}$, see (5.6), with stochastic matrices $\mathcal{M}^{\mathrm{clas}}$ having non-negative elements $\mathcal{M}_{ij}^{\mathrm{clas}}$ such that $\sum_i \mathcal{M}_{ij}^{\mathrm{clas}} = 1$ for any $j = 1,\ldots,n$. The contractive distances d^{clas} on $\mathcal{E}_{\mathrm{clas}}$ satisfy $d^{\mathrm{clas}}(\mathcal{M}^{\mathrm{clas}}\mathbf{p},\mathcal{M}^{\mathrm{clas}}\mathbf{q}) \leq d^{\mathrm{clas}}(\mathbf{p},\mathbf{q})$ for any such matrices. According to a result of Cencov [55], a Riemannian distance on $\mathcal{E}_{\mathrm{clas}}$ with metric g^{clas} is contractive if and only if $g_{\mathbf{p}}^{\mathrm{clas}}(\mathbf{a},\mathbf{a}) = c\sum_k a_k^2/p_k$ for any $\mathbf{a}\in\mathbb{R}^n$ and some

$c > 0$, that is, the infinitesimal distance between a probability vector $\mathbf{p}$ and a neighboring vector $\mathbf{p} + \mathrm{d}\mathbf{p}$ is proportional to

$$\mathrm{d}s^2_{\mathrm{Fisher}} = \sum_{k=1}^{n} \frac{\mathrm{d}p_k^2}{p_k} \ . \tag{8.51}$$

The associated metric is known as the Fisher metric and plays an important role in statistics. It induces the Hellinger distance (8.23) up to a factor of one fourth.

Let us come back to the quantum case. Although g_ρ is in principle defined on the real vector space $\mathcal{B}(\mathcal{H})_{\mathrm{s.a.}}$ (the tangent space of $\mathcal{E}(\mathcal{H})$), one can extend it as a scalar product on the complex Hilbert space $\mathcal{B}(\mathcal{H})$. Without loss of generality, one may require that this scalar product satisfies

$$g_\rho(A, B) = g_\rho(B^*, A^*) = \overline{g_\rho(A^*, B^*)} \quad , \quad A, B \in \mathcal{B}(\mathcal{H}) \ . \tag{8.52}$$

(for instance, this is the case for the Hilbert-Schmidt product (4.1)). We first note that one can associate to g a family $\{\mathcal{K}_\rho ; \rho \in \mathcal{E}(\mathcal{H})\}$ of positive operators on the Hilbert space $\mathcal{B}(\mathcal{H})$ endowed with the scalar product (4.1), by setting

$$g_\rho(A, B) = \left\langle A, \mathcal{K}_\rho^{-1}(B) \right\rangle \quad , \quad A, B \in \mathcal{B}(\mathcal{H}) \ . \tag{8.53}$$

Let us write $\rho_{\mathcal{M}} = \mathcal{M}(\rho)$. The monotonicity condition (8.50) reads $\mathcal{M}^* \mathcal{K}_{\rho_{\mathcal{M}}}^{-1} \mathcal{M} \leq \mathcal{K}_\rho^{-1}$, which means that $\mathcal{K}_\rho^{1/2} \mathcal{M}^* \mathcal{K}_{\rho_{\mathcal{M}}}^{-1} \mathcal{M} \mathcal{K}_\rho^{1/2}$ is a contraction. This is equivalent to $K_{\rho_{\mathcal{M}}}^{-1/2} \mathcal{M} \mathcal{K}_\rho \mathcal{M}^* \mathcal{K}_{\rho_{\mathcal{M}}}^{-1/2}$ being a contraction. Therefore g is contractive if and only if

$$\mathcal{M} \mathcal{K}_\rho \mathcal{M}^* \leq \mathcal{K}_{\mathcal{M}(\rho)} \tag{8.54}$$

for any ρ and $\mathcal{M}$.

Lemma 8.6.1. [195] *The contractivity condition (8.54) is fulfilled by the positive operators*

$$\mathcal{K}_\rho = \mathcal{R}_\rho^{\frac{1}{2}} f(\Delta_\rho) \mathcal{R}_\rho^{\frac{1}{2}} \ , \tag{8.55}$$

where $\mathcal{R}_\rho$ stands for the right multiplication by ρ (see (4.7)), $\Delta_\rho = \Delta_{\rho|\rho}$ is the modular operator defined in (4.8), and $f : \mathbb{R}_+ \to \mathbb{R}$ is an operator monotone-increasing function with values in $\mathbb{R}_+$.

Proof. Let us recall that the modular operators Δ_ρ and $\Delta_{\rho_{\mathcal{M}}}$ on $\mathcal{B}(\mathcal{H})$ are (self-adjoint and) positive. In analogy with the proof of Theorem 7.2.1, we introduce the contraction $\mathcal{C}_{\mathcal{M}}$ defined by (7.14). It has been observed in this proof that $\mathcal{C}_{\mathcal{M}}^* \Delta_\rho \mathcal{C}_{\mathcal{M}} \leq \Delta_{\rho_{\mathcal{M}}}$. Since asking that a continuous function $f : \mathbb{R}_+ \to \mathbb{R}$ be operator monotone-increasing and non-negative is the same as asking that f be operator concave (see Appendix A and [38], Theorem V.2.5), it follows from the Jensen-type inequality (4) and the monotonicity of f that

$$\mathcal{C}_{\mathcal{M}}^* f(\Delta_\rho) \mathcal{C}_{\mathcal{M}} \leq f(\Delta_{\rho_{\mathcal{M}}}) \ . \tag{8.56}$$

Multiplying both sides by $B' \rho_{\mathcal{M}}^{\frac{1}{2}}$ and taking the scalar product by the same vector, this is equivalent to

$$\left\langle B' \, , \, \mathcal{M} \mathcal{R}_\rho^{\frac{1}{2}} f(\Delta_\rho) \mathcal{R}_\rho^{\frac{1}{2}} \mathcal{M}^*(B') \right\rangle \leq \left\langle B' \, , \, \mathcal{R}_{\rho_{\mathcal{M}}}^{\frac{1}{2}} f(\Delta_{\rho_{\mathcal{M}}}) \mathcal{R}_{\rho_{\mathcal{M}}}^{\frac{1}{2}}(B') \right\rangle \tag{8.57}$$

for any $B' \in \mathcal{B}(\mathcal{H}')$. Thus the operator $\mathcal{K}_\rho$ defined in (8.55) satisfies the contractivity condition (8.54). □

Formulas (8.53) and (8.55) yield a family of monotonous metrics, in one-to-one correspondence with non-negative operator monotone functions f, given by $g_\rho(A, B) = \langle A \rho^{-\frac{1}{2}} \, , \, f(\Delta_\rho)^{-1}(B \rho^{-\frac{1}{2}}) \rangle$ for any $A, B \in \mathcal{B}(\mathcal{H})$. More explicitly, for any ρ with spectral decomposition $\rho = \sum_k p_k |k\rangle\langle k|$ one finds

$$g_\rho(A, A) = \sum_{k,l=1}^{n} c(p_k, p_l) |\langle k | A | l \rangle|^2 \quad , \quad A \in \mathcal{B}(\mathcal{H})_{\mathrm{s.a.}} \ , \tag{8.58}$$

where $c(p, q)$ is given by

$$c(p, q) = \frac{p f(q/p) + q f(p/q)}{2pq f(p/q) f(q/p)} \tag{8.59}$$

and satisfies $c(tp,tq)=t^{-1}c(p,q)$ for any $t\in\mathbb{R}$, $t\neq 0$, and $c(p,p)=f(1)^{-1}p^{-1}$. By using $\Delta_\rho(B^*)=(\Delta_\rho^{-1}(B))^*$, it is easy to see that the condition (8.52) is satisfied if and only if $f(x)=xf(x^{-1})$. In particular, by choosing the following operator monotone functions (see Appendix A) :

$$f_{\rm Harm}(x)=\frac{2x}{x+1}\ \ \leq\ \ f_{\rm KM}(x)=\frac{x-1}{\ln x}\ \ \leq\ \ f_{\rm H}=\frac{(1+\sqrt{x})^2}{4}\ \ \leq\ \ f_{\rm B}(x)=\frac{x+1}{2} \tag{8.60}$$

one is led to

$$c_{\rm Harm}(p,q)=\frac{p+q}{2pq}\ \ \geq\ \ c_{\rm KM}(p,q)=\frac{\ln p-\ln q}{p-q}\ \ \geq\ \ c_{\rm H}(p,q)=\frac{4}{(\sqrt{p}+\sqrt{q})^2}\ \ \geq\ \ c_{\rm B}(p,q)=\frac{2}{p+q}\,. \tag{8.61}$$

In view of (8.41) and (8.49), the last choice $f_{\rm B}$ gives the Bures metrics and $f_{\rm H}$ gives the Hellinger metric up to a factor of one fourth. The second choice corresponds to the so-called Kubo-Mori (or Bogoliubov) metric, which is associated to the relative von Neumann entropy. Actually, by substituting (8.48) into (8.45) and taking $\alpha\to 1$ one obtains

$$S(\rho+{\rm d}\rho||\rho)=\frac{1}{2}\sum_{k,l=1}^{n}c_{\rm KM}(p_k,p_l)\big|\langle k|{\rm d}\rho|l\rangle\big|^2=\frac{1}{2}g_{\rm KM}({\rm d}\rho,{\rm d}\rho)\,. \tag{8.62}$$

According to the formula $S(\rho+t{\rm d}\rho)=S(\rho)-t\,{\rm tr}({\rm d}\rho\ln\rho)-S(\rho+t{\rm d}\rho||\rho)$, one also gets

$$g_{\rm KM}({\rm d}\rho,{\rm d}\rho)=-\frac{{\rm d}^2S(\rho+t{\rm d}\rho)}{{\rm d}t^2}\bigg|_{t=0}\,, \tag{8.63}$$

S being the von Neumann entropy (since S is concave, the second derivative in the right-hand side is non-positive and defines a scalar product on $\mathcal{B}(\mathcal{H})$). As stressed by Balian, Alhassid and Reinhardt [20], this makes the Kubo-Mori metric quite natural from a physical viewpoint. Its properties have been investigated in [23].

A result due to Kubo and Ando [151] states that there is a one-to-one correspondence between operator monotone functions f and operator means, that is, maps $m:(R,L)\in\mathcal{B}(\mathcal{H})_+\times\mathcal{B}(\mathcal{H})_+\mapsto m(R,L)\in\mathcal{B}(\mathcal{H})$ satisfying

(a) if $0\leq R\leq T$ and $0\leq L\leq N$ then $m(R,L)\leq m(T,N)$ (monotonicity);

(b) $C^*m(R,L)C\leq m(C^*RC,C^*LC)$.

This correspondence is given by the formula

$$m_f(R,L)=R^{\frac{1}{2}}f(R^{-\frac{1}{2}}LR^{-\frac{1}{2}})R^{\frac{1}{2}}\,. \tag{8.64}$$

By taking $f_{\rm Harm}$ and $f_{\rm B}$ as in (8.60) one obtains the harmonic mean $m_{\rm Harm}(R,L)=(R/2)^{-1}+(L/2)^{-1}$ and the arithmetic mean $m_{\rm B}(R,L)=(R+L)/2$, respectively, and for $f(x)=\sqrt{x}$ one gets the so-called geometric mean (for more detail see e.g. [50]). The positive operators (8.55) can be written as

$$\mathcal{K}_\rho=m_f(\mathcal{R}_\rho,\mathcal{L}_\rho)\,. \tag{8.65}$$

The theory of Kubo and Ando shows that the harmonic mean $m_{\rm Harm}$ and arithmetic mean $m_{\rm B}$ are respectively the smallest and largest symmetric operator means. Thus the Bures metric $g_{\rm B}$ is the smallest monotone metric among the family of metrics given by (8.53) and (8.55) with the normalization $g_\rho(1,1)={\rm tr}(\rho^{-1})$. It turns out that this family contains all contractive metrics, that is, all such metrics have the form (8.58).

Theorem 8.6.2. (Petz [195]) *The distances with metrics g given by (8.58) are contractive for any non-negative operator monotone-increasing function $f(x)$ satisfying $f(x)=xf(x^{-1})$. Conversely, any continuous metric $g:\rho\mapsto g_\rho$ on $\mathcal{E}(\mathcal{H})$ may be obtained from (8.58) by a choice of a suitable function f with these properties. In particular, there is a one-to-one correspondence between continuous contractive metrics satisfying $g_\rho(1,1)={\rm tr}(\rho^{-1})$ and operator means. The Bures distance is the smallest of all contractive Riemannian distances with metrics satisfying this normalization condition.*

This theorem is of fundamental importance in geometrical approaches to quantum information.

Proof. The first statement has been proven above. Conversely, let g be a continuous contractive metric on $\mathcal{E}(\mathcal{H})$ and let us show that there exists an operator monotone function $f : \mathbb{R}_+ \to \mathbb{R}_+$ such that for any $\rho \in \mathcal{E}(\mathcal{H})$, g_ρ is given by (8.53) and (8.55) or, equivalently, by (8.58) and (8.59). We first note that g being contractive it is in particular unitary invariant, i.e., $g_{U^*\rho U}(U^*AU, U^*BU) = g_\rho(A, B)$ for any unitary U (see Sec.8.1). More generally, if the quantum operations $\mathcal{M}$ and $\mathcal{T}$ are such that ρ, A, and B are invariant under $\mathcal{T} \circ \mathcal{M}$, then $g_{\mathcal{M}(\rho)}(\mathcal{M}(A), \mathcal{M}(B)) = g_\rho(A, B)$. The main idea of the proof is to combine this invariance property with the uniqueness of the contractive classical distance. Denoting by $(g_\rho)_{ij,kl} = g_\rho(|i\rangle\langle j|, |k\rangle\langle l|)$ the matrix elements of the scalar product g_ρ in an orthonormal eigenbasis $\{|k\rangle\}$ of ρ, we need to prove that

$$(g_\rho)_{ij,kl} = \delta_{ik}\delta_{jl}\, c(p_i, p_j) \tag{8.66}$$

where δ_{ik} is the Kronecker symbol. To show that the matrix elements of g_ρ vanish for $i \neq j$ and $(k, l) \neq (i, j)$, it suffices to establish that

$$g_\rho\big(|i\rangle\langle j| + s|k\rangle\langle l|, |i\rangle\langle j| + s|k\rangle\langle l|\big) = g_\rho\big(|i\rangle\langle j| - s|k\rangle\langle l|, |i\rangle\langle j| - s|k\rangle\langle l|\big) \tag{8.67}$$

for $s = 1$ and $s = \mathrm{i}$ (the result then follows by polarization). If one of the indices i, j, k, and l is different from the three others, say $i \notin \{j, k, l\}$, this comes from the invariance of g under the unitary $U^{(i)} = \sum_k u_k^{(i)}|k\rangle\langle k|$ with $u_k^{(i)} = -1$ if $k = i$ and 1 otherwise. Hence $(g_\rho)_{ij,kl} = 0$ when $i \neq j$ and $(i, j) \neq (k, l)$, (l, k). Similarly, by choosing $u_k^{(i)} = \mathrm{i}$ if $i = k$ and 1 otherwise, this is also true for $i \neq j$ and $(i, j) = (l, k)$. The only non-vanishing matrix elements of g_ρ are thus $(g_\rho)_{ii,kk}$ and $(g_\rho)_{ij,ij}$ for $i \neq j$.

To determine $(g_\rho)_{ii,kk}$ we observe that the restriction of g_ρ to the space of matrices commuting with ρ induces a contractive metric on the probability simplex $\mathcal{E}_{\mathrm{clas}}$, defined by $g^{\mathrm{clas}}_{\mathbf{p}}(\mathbf{a}, \mathbf{b}) = g_\rho(\sum_k a_k|k\rangle\langle k|, \sum_k b_k|k\rangle\langle k|)$ for any $\mathbf{a}, \mathbf{b} \in \mathcal{E}_{\mathrm{clas}}$. Indeed, one can associate a quantum operation $\mathcal{M}$ to a stochastic matrix $\mathcal{M}^{\mathrm{clas}}$ by defining $\mathcal{M}(|k\rangle\langle l|) = \delta_{kl}\sum_j \mathcal{M}^{\mathrm{clas}}_{jk}|j\rangle\langle j|$ ($\mathcal{M}$ has the Kraus form (5.10) as $\mathcal{M}^{\mathrm{clas}}_{jk} \geq 0$ and $\sum_j \mathcal{M}^{\mathrm{clas}}_{jk} = 1$ for any k). Then $\mathcal{M}(\rho) = \sum_j (\mathcal{M}^{\mathrm{clas}}\mathbf{p})_j|j\rangle\langle j|$ where $\mathbf{p}$ is the vector of eigenvalues of ρ, and (8.50) implies that g^{clas} is contractive under $\mathcal{M}^{\mathrm{clas}}$. According to the uniqueness of the contractive classical metrics, one has

$$(g_\rho)_{ii,kk} = g^{\mathrm{clas}}_{\mathbf{p}}(\boldsymbol{\delta}_i, \boldsymbol{\delta}_k) = c\,\frac{\delta_{ik}}{p_k}\,, \tag{8.68}$$

with $c > 0$ and $\boldsymbol{\delta}_i = (\delta_{il})_{l=1}^n$.

We now turn to the matrix elements $(g_\rho)_{ij,ij}$ for $i \neq j$. By unitary invariance, it is enough to determine $(g_\rho)_{12,12}$. To this end, we consider the quantum operations $\mathcal{M}$ from the space $\mathcal{B}(\mathcal{H})$ of $n \times n$ matrices to the space $\mathcal{B}(\mathbb{C}^3)$ of 3×3 matrices and $\mathcal{T} : \mathcal{B}(\mathbb{C}^3) \to \mathcal{B}(\mathcal{H})$ with Kraus operators $\{A_i\}_{i=2}^n$ and $\{B_i\}_{i=2}^n$, respectively, given by

$$A_2 = B_2 = |1\rangle\langle 1| + |2\rangle\langle 2| \quad , \quad A_i = |3\rangle\langle i| \quad , \quad B_i = \frac{\sqrt{p_i}}{\sqrt{1 - p_1 - p_2}}|i\rangle\langle 3| \quad , \quad i = 3, \ldots, n\,. \tag{8.69}$$

A simple calculation yields $\mathcal{T} \circ \mathcal{M}(\rho) = \rho$. As stressed above, one can deduce from the contractivity of g_ρ that $(g_\rho)_{12,12} = (g_{\mathcal{M}(\rho)})_{12,12}$, thereby showing that this matrix element depends on p_1 and p_2 only. By unitary invariance, $(g_\rho)_{ij,ij}$ only depends on p_i and p_j and one can set $(g_\rho)_{ij,ij} = c(p_i, p_j)$ for $i \neq j$, $c(p, q)$ being independent of ρ. This complete the proof of (8.66), excepted that it remains to justify that $c(p, p) = c/p$.

We proceed by showing that $c(q, p)$ is given by (8.59) with f having the desired properties. Thanks to (8.52), we know that $c(p, q)$ is real and symmetric. One verifies that $c(p, p) = c/p$ by the following argument. Let us assume that ρ has a degenerate eigenvalue, say $p_1 = p_2$. Then $\rho = U\rho U^*$ for any unitary U acting trivially on $\mathrm{span}\{|3\rangle, \ldots, |n\rangle\}$. By unitary invariance, $g_\rho(|\psi\rangle\langle\psi|, |\psi\rangle\langle\psi|) = (g_\rho)_{11,11} = c/p_1$ for any $|\psi\rangle \in \mathrm{span}\{|1\rangle, |2\rangle\}$. Taking e.g. $|\psi\rangle = (|1\rangle + |2\rangle)/\sqrt{2}$ and using (8.66), we get $(g_\rho)_{12,12} = c(p_1, p_1) = c/p_1$. In order to establish that $c(p, q)$ is homogeneous we consider the quantum operations $\mathcal{M} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}\otimes\mathcal{H}_{\mathsf{E}})$ and $\mathcal{T} : \mathcal{B}(\mathcal{H}\otimes\mathcal{H}_{\mathsf{E}}) \to \mathcal{B}(\mathcal{H})$ defined by $\mathcal{M}(\rho) = \rho \otimes 1/n_{\mathsf{E}}$ and $\mathcal{T}(\widehat{\rho}) = \mathrm{tr}_{\mathsf{E}}(\widehat{\rho})$ (here n_{E} is the dimension of $\mathcal{H}_{\mathsf{E}}$). Clearly, $\mathcal{T} \circ \mathcal{M} = 1$, thus by similar arguments as above and by taking advantage of (8.66), one finds

$$c(p_i, p_j) = (g_\rho)_{ij,ij} = g_{\mathcal{M}(\rho)}(\mathcal{M}(|i\rangle\langle j|), \mathcal{M}(|i\rangle\langle j|)) = n_{\mathsf{E}}^{-1} c\Big(\frac{p_i}{n_{\mathsf{E}}}, \frac{p_j}{n_{\mathsf{E}}}\Big)\,. \tag{8.70}$$

As this is true for any positive integer n_{E} and any state ρ, one concludes that $c(tp, tq) = t^{-1}c(p, q)$ for all $p, q \in [0, 1]$ and all rationals t with tp, $tq \in [0, 1]$. This is the point where we need the continuity of the metric to make sure that $c(p, q)$ is continuous. Then the equality holds for all real t. Setting $f(x) = 1/c(x, 1)$ and

using the symmetry of $c(p,q)$, one easily derives the identities (8.59) and $f(x^{-1}) = x^{-1}f(x)$. Furthermore, $f(1)^{-1} = c(1,1) = c$.

To complete the proof, we have to show that f is operator concave. With this aim, let us consider the inequality (8.54) which is equivalent to g_ρ being contractive. We choose $\mathcal{M}$ in this inequality to be the partial trace operation $\mathcal{T} : \widehat{\rho} \mapsto \mathrm{tr}_{\mathbb{C}^2}(\widehat{\rho}) \otimes 1/2$ on $\mathcal{B}(\mathcal{H} \otimes \mathbb{C}^2)$ and $\widehat{\rho} = (\rho_0 \otimes |0\rangle\langle 0| + \rho_1 \otimes |1\rangle\langle 1|)/2$. From (8.54) we find that for any $A \in \mathcal{B}(\mathcal{H})$,

$$\big\langle \mathcal{T}^*(A \otimes 1)\,,\, \mathcal{K}_{\widehat{\rho}}\mathcal{T}^*(A \otimes 1)\big\rangle \leq \big\langle A \otimes 1\,,\, \mathcal{K}_{\mathcal{T}(\widehat{\rho})}(A \otimes 1)\big\rangle\,. \tag{8.71}$$

But $\mathcal{K}_{\widehat{\rho}}(A \otimes 1) = (\mathcal{K}_{\rho_0}(A) \otimes |0\rangle\langle 0| + \mathcal{K}_{\rho_1}(A) \otimes |1\rangle\langle 1|)/2$. Accordingly, (8.71) reduces to

$$\frac{1}{2}\big\langle A\,,\, (\mathcal{K}_{\rho_0} + \mathcal{K}_{\rho_1})A\big\rangle \leq \big\langle A\,,\, \mathcal{K}_{(\rho_0+\rho_1)/2}A\big\rangle\,, \tag{8.72}$$

thereby showing that the map

$$\rho \mapsto \mathcal{K}_\rho = f(\mathcal{L}_\rho\mathcal{R}_\rho^{-1})\mathcal{R}_\rho \tag{8.73}$$

is mid-point concave. By a standard argument based on a dyadic decomposition, it follows that this map is concave [50]. Using the $*$-isomorphism between the C^*-algebras $\mathcal{B}(\mathcal{B}(\mathcal{H}))$ and $\mathcal{B}(\mathcal{H} \otimes \mathcal{H})$ (Sec.4.1), this is equivalent to say that the map

$$A \mapsto f\big(A \otimes (A^T)^{-1}\big)\, 1 \otimes A^T \tag{8.74}$$

is concave. One easily deduces from this that the map $(A,B) \mapsto f(A \otimes (B^T)^{-1})\, 1 \otimes B^T$ is jointly concave. In particular, $A \mapsto f(A)$ is concave. This shows that f is operator concave. □

Chapter 9

State discrimination and parameter estimation in large systems

A force de savoir tant sur si peu de choses, il finira par savoir tout sur rien (anonymous).

In this chapter we examine two problems related to the state discrimination task discussed in chapter 6, namely, the quantum hypothesis testing and parameter estimation. In the first problem, one wants to determine asymptotically the probability of error in discriminating two states when one has N independent copies of those states, for $N \to \infty$. In the second problem, the goal is to estimate as precisely as possible a real parameter from measurements performed on a large number of particles in a state depending smoothly on this parameter.

9.1 Quantum hypothesis testing: discriminating two states from many identical copies

An important issue in classical information theory is to discriminate two probability measures $\mathbf{p}_1$ and $\mathbf{p}_2$ on a measurable space $(\Omega, \mathcal{F})$, given the outcomes of N independent identically distributed (i.i.d.) random variables, whose law is either $\mathbf{p}_1$ or $\mathbf{p}_2$. Since one has to decide among two hypothesis – the first (second) one being that the observed data is distributed according to $\mathbf{p}_1$ ($\mathbf{p}_2$) – this discrimination task bears the name of "hypothesis testing". For a given test function, i.e., a random variable $M_{\rm clas}$ with values in $[0,1]$, the probability of error is $P_{{\rm err},N} = \eta_1 \mathbf{p}_1^{(N)}(M_{\rm clas}) + \eta_2 \mathbf{p}_2^{(N)}(1 - M_{\rm clas})$, where $\mathbf{p}_i^{(N)} = \mathbf{p}_i^{\otimes N}$ is the N-fold product measure and η_i the prior probability attached to $\mathbf{p}_i$. It is easy to convince oneself that the minimal error is achieved for the maximum likelihood test function defined by[1]

$$M_{\rm clas}^{\rm opt} = 1_{\{\eta_2 \rho_2^{(N)} - \eta_1 \rho_1^{(N)} \geq 0\}} \, , \tag{9.1}$$

$\rho_i^{(N)} = {\rm d}\mathbf{p}_i^{(N)} / {\rm d}\boldsymbol{\mu}^{(N)}$ being the density of $\mathbf{p}_i^{(N)}$ with respect to the measure $\boldsymbol{\mu}^{(N)} = \mathbf{p}_1^{(N)} + \mathbf{p}_2^{(N)} = \boldsymbol{\mu}^{\otimes N}$. The corresponding error is

$$\begin{aligned} P_{{\rm err},N}^{\rm opt}(\{\mathbf{p}_i^{(N)}, \eta_i\}) &= \min_{0 \leq M_{\rm clas} \leq 1} \left\{ \int_{\Omega^N} {\rm d}\boldsymbol{\mu}^{(N)} \left(\eta_1 \rho_1^{(N)} M_{\rm clas} + \eta_2 \rho_2^{(N)} (1 - M_{\rm clas}) \right) \right\} \\ &= \int_{\Omega^N} {\rm d}\boldsymbol{\mu}^{(N)} \min\left\{ \eta_1 \rho_1^{(N)}, \eta_2 \rho_2^{(N)} \right\} . \end{aligned} \tag{9.2}$$

One is typically interested in the limit of a large number of tests, i.e., $N \to \infty$. One can show that the error probability decays exponentially like $P_{{\rm err},N}^{\rm opt} \sim {\rm e}^{-N\xi(\mathbf{p}_1, \mathbf{p}_2)}$, with an exponent given by the Chernoff bound [60]

$$\xi(\mathbf{p}_1, \mathbf{p}_2) = - \lim_{N \to \infty} \frac{1}{N} \ln P_{{\rm err},N}^{\rm opt}(\{\mathbf{p}_i^{(N)}, \eta_i\}) = - \inf_{\alpha \in (0,1)} \left\{ \ln \left(\int_\Omega {\rm d}\boldsymbol{\mu} \, \rho_1^\alpha \rho_2^{1-\alpha} \right) \right\} , \tag{9.3}$$

where we have set $\rho_i = \rho_i^{(1)}$. One recognizes in the infimum in the right-hand side the classical Rényi divergence (7.35) multiplied by $(\alpha - 1)$.

[1] Here 1_A stands for the indicator function on $A \subset \Omega$, i.e., $1_A(\omega) = 1$ if $\omega \in A$ and 0 otherwise.

In quantum mechanics, the hypothesis testing can be rephrased as the discrimination of two N-fold tensor product states $\rho_1^{\otimes N}$ and $\rho_2^{\otimes N}$. The corresponding minimal error probability is given by the Helstrom formula (6.6),

$$P_{\mathrm{err},N}^{\mathrm{opt}}(\{\rho_i^{\otimes N},\eta_i\}) = \frac{1}{2}\big(1-\mathrm{tr}\,|\Lambda_N|\big) \quad , \quad \Lambda_N = \eta_1\rho_1^{\otimes N} - \eta_2\rho_2^{\otimes N}\,, \tag{9.4}$$

and the optimal measurement consists of the orthogonal projectors $M_{\pm}^{\mathrm{opt}}$ on the supports of the positive and negative parts of Λ_N. Note that if ρ_1 and ρ_2 commute then M_-^{opt} can be identified with the maximum likelihood test function and one recovers the classical formula (9.2) from (9.4). Surprisingly, the generalization of the Chernoff bound (9.3) to the quantum setting has been settled out only recently. It has been highlighted in Sec. 7.3 that the Rényi divergences appearing in this bound have several natural quantum extensions, according to the choice of operator ordering. It was proven by Audenaert *et al.* [19] and by Nussbaum and Szkola [181] that the right extension is the normal-ordered relative Rényi entropy $S_\alpha^{(\mathrm{n})}(\rho||\sigma)$ defined in (7.29).

Proposition 9.1.1. (Quantum Chernoff bound [19, 181]) *One has*

$$-\lim_{N\to\infty}\frac{1}{N}\ln P_{\mathrm{err},N}^{\mathrm{opt}}(\{\rho_i^{\otimes N},\eta_i\}) = -\inf_{\alpha\in(0,1)}\Big\{\ln\big(\mathrm{tr}[\rho_1^\alpha\rho_2^{1-\alpha}]\big)\Big\} = \sup_{\alpha\in(0,1)}\Big\{(1-\alpha)S_\alpha^{(\mathrm{n})}(\rho_1||\rho_2)\Big\}\,. \tag{9.5}$$

This limit defines a jointly convex function $\xi_Q(\rho_1,\rho_2)$ with values in $\mathbb{R}_+\cup\{+\infty\}$, which is contractive under quantum operations. Moreover, ξ_Q induces the quantum Hellinger metric up to a factor of one half, that is, if ρ and $\rho+\mathrm{d}\rho$ are infinitesimally close then $\xi_Q(\rho+\mathrm{d}\rho,\rho) = g_{\mathrm{H}}(\mathrm{d}\rho,\mathrm{d}\rho)/2$ is given by (8.49).

The infimum in (9.5) is attained for a unique $\alpha\in(0,1)$ satisfying $\mathrm{tr}(\rho_1^\alpha\rho_2^{1-\alpha}(\ln\rho_1-\ln\rho_2))=0$ [19]. Actually, for any fixed ρ and σ, the function $\alpha\mapsto F_\alpha^{(\mathrm{n})}(\rho||\sigma)^\alpha = \mathrm{tr}[\rho^\alpha\sigma^{1-\alpha}]$ is convex (this is a simple consequence of the convexity of $\alpha\mapsto p^\alpha q^{1-\alpha}$ for $p,q>0$) and $F_\alpha^{(\mathrm{n})}(\rho||\sigma)\le F_{0,1}^{(\mathrm{n})}(\rho||\sigma)=1$ by the Hölder inequality (4.3). Before entering into the proof, let us also mention that $\xi_Q(\rho,\sigma)<\infty$ whenever ρ and σ do not have orthogonal supports. If $\rho=|\psi\rangle\langle\psi|$ is pure, the quantum Chernoff bound is related to the fidelity by $\xi_Q(\rho,\sigma) = -\ln F(\rho,\sigma) = -\ln\langle\psi|\sigma|\psi\rangle$ (in fact, then $F_\alpha^{(\mathrm{n})}(\rho||\sigma)^\alpha=\langle\psi|\sigma^{1-\alpha}|\psi\rangle$ is minimum for $\alpha=0$).

Proof. To shorten notation we write $P_{\mathrm{eff},N}^{\mathrm{opt}}$ when referring to $P_{\mathrm{err},N}^{\mathrm{opt}}(\{\rho^{\otimes N},\eta,\sigma^{\otimes N},1-\eta\})$. The fact that

$$\limsup_{N\to\infty}\frac{1}{N}\ln P_{\mathrm{eff},N}^{\mathrm{opt}} \le -\xi_Q(\rho,\sigma) = \inf_{\alpha\in(0,1)}\big\{\ln(\mathrm{tr}[\rho^\alpha\sigma^{1-\alpha}])\big\} \tag{9.6}$$

follows from (9.4) and the trace inequality

$$\frac{1}{2}\Big(\mathrm{tr}(A)+\mathrm{tr}(B)-\mathrm{tr}\,|A-B|\Big)\le\mathrm{tr}(A^\alpha B^{1-\alpha})\,, \tag{9.7}$$

where A and B are non-negative operators and $\alpha\in[0,1]$. This inequality has been first established in [19]. A simple proof due to N. Ozawa is reported in Appendix B. The reverse inequality to (9.6) is a consequence of the classical Chernoff bound. This can be justified as follows [181]. Let us observe that the optimal measurement is a von Neumann measurement $\{\Pi^{\mathrm{opt}},1-\Pi^{\mathrm{opt}}\}$ with Π^{opt} a projector, so that

$$\begin{aligned} P_{\mathrm{err},N}^{\mathrm{opt}} = 1-P_{\mathrm{S},N}^{\mathrm{opt}} &= \eta\,\mathrm{tr}\big((1-\Pi^{\mathrm{opt}})\rho^{\otimes N}\big)+(1-\eta)\,\mathrm{tr}\big(\Pi^{\mathrm{opt}}\sigma^{\otimes N}\big) \\ &= \sum_{\underline{k},\underline{l}}\Big(\eta p_{\underline{k}}\big|\langle\Phi_{\underline{l}}|(1-\Pi^{\mathrm{opt}})|\Psi_{\underline{k}}\rangle\big|^2+(1-\eta)q_{\underline{l}}\big|\langle\Psi_{\underline{k}}|\Pi^{\mathrm{opt}}|\Phi_{\underline{l}}\rangle\big|^2\Big)\,, \end{aligned} \tag{9.8}$$

where $\{|\Psi_{\underline{k}}\rangle\}$ and $\{|\Phi_{\underline{l}}\rangle\}$ are orthonormal eigenbases of $\rho^{\otimes N}$ and $\sigma^{\otimes N}$, respectively, and $p_{\underline{k}}$ and $q_{\underline{l}}$ are the corresponding eigenvalues. We may without loss of generality assume that $\eta\le 1/2$. By using the inequality $|a|^2+|b|^2\ge|a+b|^2/2$ one gets

$$P_{\mathrm{err},N}^{\mathrm{opt}}\ge\eta\sum_{\underline{k},\underline{l}}\frac{1}{2}\min\{p_{\underline{k}},q_{\underline{l}}\}\big|\langle\Phi_{\underline{l}}|\Psi_{\underline{k}}\rangle\big|^2\,. \tag{9.9}$$

But $\rho^{\otimes N}$ corresponds to N independent copies of the state $\rho=\sum_k p_k|\psi_k\rangle\langle\psi_k|$, hence its eigenvalues $p_{\underline{k}}$ and eigenvectors $|\Psi_{\underline{k}}\rangle$ are products of N eigenvalues p_k and N eigenvectors $|\psi_k\rangle$ of ρ, respectively, and similarly for $\sigma^{\otimes N}$ with the eigenvalues q_l and eigenvectors $|\phi_l\rangle$ of σ. This means that $p_{\underline{k}}|\langle\Phi_{\underline{l}}|\Psi_{\underline{k}}\rangle|^2$ can be viewed as the N-fold product of the probability $\boldsymbol{\pi}_1$ on $\{1,\dots,n\}^2$ defined by $(\boldsymbol{\pi}_1)_{kl}=p_k|\langle\phi_l|\psi_k\rangle|^2$. Analogously, $q_{\underline{l}}|\langle\Phi_{\underline{l}}|\Psi_{\underline{k}}\rangle|^2$

is the N-fold product of $\boldsymbol{\pi}_2$ with $(\boldsymbol{\pi}_2)_{kl} = q_l|\langle\phi_l|\psi_k\rangle|^2$. Consequently, the sum in (9.9) is the minimal error probability $P_{\mathrm{err},N}^{\mathrm{opt}}(\{\boldsymbol{\pi}_i^{(N)}, 1/2\})$ for discriminating $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ with equal prior probabilities (see (9.2)). One then deduces from the classical Chernoff bound (9.3) that

$$\liminf_{N\to\infty} \frac{1}{N} \ln P_{\mathrm{err}}^{\mathrm{opt}} \geq \inf_{\alpha\in(0,1)} \left\{ \ln\Big(\sum_{k,l=1}^{n} (\boldsymbol{\pi}_1)_{kl}^{\alpha} (\boldsymbol{\pi}_2)_{kl}^{1-\alpha} \Big) \right\} = -\xi_Q(\rho,\sigma) \; . \tag{9.10}$$

Together with (9.6) this proves the quantum Chernoff bound.

It is nevertheless instructive to show (9.10) directly from (9.9), without relying on the classical result, by using the theory of large deviations for sums of i.i.d. random variables and the relative modular operator $\Delta_{\sigma|\rho}$ (see chapter 4), which appears here quite naturally [140]. Indeed, let us set $\xi = \rho^{\frac{1}{2}}$ and note that for any real function $f : (0,\infty) \to \mathbb{R}$, according to (4.8) and by the functional calculus, it holds

$$\langle \xi \, , \, f(\Delta_{\sigma|\rho})\, \xi \rangle = \sum_{k,l=1}^{n} p_k f\Big(\frac{q_l}{p_k}\Big) |\langle\phi_l|\psi_k\rangle|^2 \; . \tag{9.11}$$

In particular, $\langle \xi \, , \, \ln(\Delta_{\sigma|\rho})\, \xi \rangle = \mathrm{tr}[\rho(\ln\sigma - \ln\rho)] = -S(\rho||\sigma)$, as already observed in Sec. 7.2. Let $\mathbf{m}_{\sigma|\rho}$ be the spectral measure of $-\ln\Delta_{\sigma|\rho}$ with respect to the vector ξ. This is a probability measure (ξ is normalized), which is related to the relative entropy by $S(\rho||\sigma) = \int \mathrm{d}\mathbf{m}_{\sigma|\rho}(t)\, t$. Taking $f(x) = \min\{x,1\} = g(-\ln x)$ with $g(t) = \min\{e^{-t},1\}$ in (9.11), one finds

$$\sum_{k,l=1}^{n} \min\{p_k,q_l\} |\langle\phi_l|\psi_k\rangle|^2 = \big\langle \xi \, , \, g(-\ln\Delta_{\sigma|\rho})\, \xi \big\rangle = \int_{\mathbb{R}} \mathrm{d}\mathbf{m}_{\sigma|\rho}(t)\, g(t) \;\geq\; \mathbf{m}_{\sigma|\rho}(\mathbb{R}_-) \; . \tag{9.12}$$

A similar inequality holds for the sum in the right-hand side of (9.9): it suffices to substitute $\Delta_{\sigma|\rho}$ by $\Delta_{\sigma^{\otimes N}|\rho^{\otimes N}} = \Delta_{\sigma|\rho}^{\otimes N}$. The spectral measure of $-\ln\Delta_{\sigma|\rho}^{\otimes N}$ is a product measure $\mathbf{m}_{\sigma|\rho}^{(N)}$ and thus $-\ln\Delta_{\sigma|\rho}^{\otimes N}$ can be interpreted as a sum of i.i.d. random variables $-\ln\Delta_{\sigma|\rho}^{(\nu)}$ with law $\mathbf{m}_{\sigma|\rho}$. The large deviation principle ensures that if $e'_{\sigma|\rho}(0) < \theta < e'_{\sigma|\rho}(1)$ then [82]

$$\lim_{N\to\infty} \frac{1}{N} \ln\Big(\mathbf{m}_{\sigma|\rho}^{(N)} \Big(-\sum_{\nu=1}^{N} \ln\Delta_{\sigma|\rho}^{(\nu)} \leq -\theta N \Big) \Big) = -\sup_{\alpha\in[0,1]} \big\{ \alpha\theta - e_{\sigma|\rho}(\alpha) \big\} \tag{9.13}$$

is up to a minus sign the Legendre transform of

$$e_{\sigma|\rho}(\alpha) = \ln\Big(\int_{\mathbb{R}} \mathrm{d}\mathbf{m}_{\sigma|\rho}(t) e^{-t\alpha} \Big) = \ln\big(\langle \xi \, , \, \Delta_{\sigma|\rho}^{\alpha}\, \xi \rangle \big) = \ln\big(\mathrm{tr}[\rho^{1-\alpha}\sigma^{\alpha}] \big) \; . \tag{9.14}$$

If $\rho \neq \sigma$ then $e'_{\sigma|\rho}(0) = -S(\rho||\sigma) < 0$ and $e'_{\sigma|\rho}(1) = S(\sigma||\rho) > 0$ (the second identity follows from the first one by symmetry $e_{\sigma|\rho}(1-\alpha) = e_{\rho|\sigma}(\alpha)$). Thus the large deviation bound (9.13) holds for $\theta = 0$. Taking advantage of (9.9) and (9.12) one is led to

$$\begin{aligned} \liminf_{N\to\infty} \frac{1}{N} \ln P_{\mathrm{err},N}^{\mathrm{opt}} \;&\geq\; \liminf_{N\to\infty} \frac{1}{N} \ln\Big(\sum_{\underline{k},\underline{l}} \min\{p_{\underline{k}}, q_{\underline{l}}\} \big|\langle \Phi_{\underline{l}} | \Psi_{\underline{k}} \rangle\big|^2 \Big) \\ &\geq\; \lim_{N\to\infty} \frac{1}{N} \ln\Big(\mathbf{m}_{\sigma|\rho}^{(N)} \Big(-\sum_{\nu=1}^{N} \ln\Delta_{\sigma|\rho}^{(\nu)} \leq 0 \Big) \Big) = \inf_{\alpha\in[0,1]} \big\{ e_{\sigma|\rho}(\alpha) \big\} = -\xi_Q(\rho,\sigma) \; , \end{aligned} \tag{9.15}$$

in agreement with (9.10). Note that these arguments justify in particular that the second member in the classical Chernoff bound (9.3) is bounded from above by the third one, as a consequence of the large deviation principle. Applying (9.6) for commuting matrices ρ and σ, this gives a full proof of this classical bound.

The joint convexity of $\xi_Q(\rho,\sigma)$ mentioned in the proposition results from the joint convexity of the relative entropies $S_{\alpha}^{(\mathrm{n})}(\rho||\sigma)$ for $\alpha \in (0,1)$, which follows from the Lieb concavity theorem, see Sec. 7.3. One then gets the contractivity of ξ_Q with respect to quantum operations from Proposition 7.2.2. This concludes the proof. □

Remark 9.1.2. *The quantum Chernoff bound (9.5) can be generalized to the case where the two states $\rho_{i,N} \in \mathcal{E}(\mathcal{H}^{\otimes N})$ to discriminate are not product states (i.e., for dependent copies).*

Actually, the large deviation principle used in the proof is not restricted to sums of i.i.d. random variables. It must be assumed that the limit $e(\alpha) = \lim_{N\to\infty} N^{-1} \ln \mathrm{tr}[\rho_{1,N}^{\alpha} \rho_{2,N}^{1-\alpha}]$ exists, is continuous in α on $[0,1]$ and differentiable on $(0,1)$, and its right derivative $e'(0)$ is smaller than its left derivative $e'(1)$ (see [140]).

Remark 9.1.3. *In asymmetric hypothesis testing one is interested by the minimal error probability of identifying the second state under the constraint that the error on the identification of the first state is smaller than ε,*

$$P_{\mathrm{err},N,\varepsilon}^{\mathrm{asym}} = \min_{0\leq M\leq 1} \left\{ \mathrm{tr}[M\rho_2^{\otimes N}] \, ; \, \mathrm{tr}[(1-M)\rho_1^{\otimes N}] \leq \varepsilon \right\} . \tag{9.16}$$

The quantum Stein's lemma [122, 182] shows that this probability decays exponentially with a rate given by the relative von Neumann entropy, i.e.

$$- \lim_{N\to\infty} \frac{1}{N} \ln P_{\mathrm{err},N,\varepsilon}^{\mathrm{asym}} = S(\rho_1 || \rho_2) . \tag{9.17}$$

The limit one gets by replacing the fixed parameter $\varepsilon > 0$ by e^{-rN} (that is, asking for an exponentially decaying error on the identification of ρ_1) is, in turn, given by the Hoeffding bound (see e.g. [140] for more detail).

An interesting link between the quantum hypothesis testing and fluctuation theorems in quantum statistical physics has been found by Jakšić *et al.* [140]. They have shown that the quantum Chernoff bound for discriminating the forward and backward time-evolved states $\rho_{\pm T/2}$ as $T \to \infty$ appears in the large deviation principle for the full counting statistics of measurements of the energy/entropy flow over the time interval $[0,T]$.

9.2 Parameter estimation in quantum metrology

The parameter estimation problem is a kind of continuous version of quantum state discrimination, in which the system state $\rho(\theta)$ depends on a continuous parameter θ. One aims at estimating this unknown parameter with the highest possible precision $\Delta\theta$ by performing measurements on $\rho(\theta)$. This precision is limited by our ability to distinguish the states $\rho(\theta)$ for values of θ differing by $\Delta\theta$.

9.2.1 Phase estimation in Mach-Zehnder interferometers

An important example is phase estimation in the Mach-Zehnder interferometer represented in Fig. 9.1. An input photon passes through a beam splitter [41] which transforms its state into a superposition of two modes propagating along different paths. These two modes acquire distinct phases θ_1 and θ_2 during the propagation and are finally recombined in a second beam splitter to read out interference fringes, from which the phase difference $\theta = \theta_1 - \theta_2$ is inferred. The interferometric sequence can be described by means of rotation matrices acting on the two-mode photon state. We shall assume at this point that the reader is familiar with second quantization[2]. The generators of the aforementioned rotations are the angular momentum operators J_x, J_y, and J_z related to the bosonic annihilation and creation operators b_j and b_j^* of a photon in mode $j = 1, 2$ by $J_x = (b_1^* b_2 + b_2^* b_1)/2$, $J_y = -\mathrm{i}(b_1^* b_2 - b_2^* b_1)/2$, and $J_z = (b_1^* b_1 - b_2^* b_2)/2$ (Schwinger representation). These operators act on the bosonic Fock space $\mathcal{F}_{\mathrm{b}}(\mathbb{C}^2)$ associated to the single photon space $\mathcal{H} \simeq \mathbb{C}^2$. The output state of the interferometer is given in terms of the input state ρ_{in} by [273]

$$\rho_{\mathrm{out}}(\theta) = e^{-\mathrm{i}\theta J_{\mathbf{n}}} \rho_{\mathrm{in}} e^{\mathrm{i}\theta J_{\mathbf{n}}} , \tag{9.18}$$

where θ is the phase to be estimated and $J_{\mathbf{n}} = n_x J_x + n_y J_y + n_z J_z$ the angular momentum in the direction specified by the unit vector $\mathbf{n} \in \mathbb{R}^3$.

One can also realize a Mach-Zehnder interferometer with ultracold atoms forming a Bose-Einstein condensate in an optical trap, instead of photons. Then the two modes correspond to two distinct atomic energy levels and the total number of atoms $N_{\mathrm{p}} = N_1 + N_2$ in these modes is fixed. In such a case the Hilbert space of the system has finite dimension $N_{\mathrm{p}} + 1$ (one deals here with indistinguishable particles). Atom interferometry in Bose-Einstein condensates is very promising due to the tunable interactions between atoms, which make it

[2] A good mathematical introduction to this formalism can be found in [42].

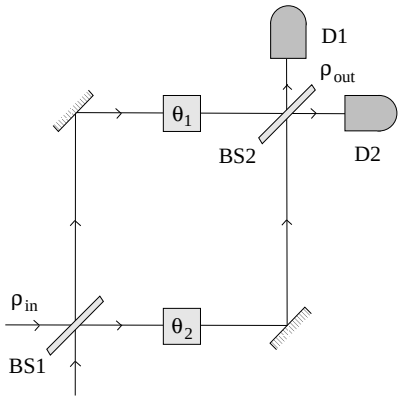

Figure 9.1: In a Mach-Zehnder interferometer, the light entering in one of the two input modes is split into two beams by a beam splitter (represented by the rectangle BS1 inclined by 45°). The photons in the first and second beams acquire some phase shifts θ_1 and θ_2, respectively, before going through a second beam splitter (rectangle BS2) and into the detectors D1 and D2, which count the number of photons in the two output modes.

possible to generate dynamically entangled states involving a large number of particles[3]. We will see below that using such entangled states as inputs leads to smaller errors $\Delta\theta$ in the phase estimation than for separable inputs. For independent (i.e., separable) particles the precision is of the order of $(\Delta\theta)_{\rm SN} \approx 1/\sqrt{N_{\rm p}}$ (shot noise limit). Higher precisions than $(\Delta\theta)_{\rm SN}$ have been reported experimentally [106, 202]. Important potential applications of these ultra-precise interferometers include atomic clocks and magnetic sensors with enhanced sensitivities [254, 209].

9.2.2 Quantum Cramér-Rao bound

In the more general setting, the problem of estimating an unknown parameter θ from a θ-dependent state evolution and measurements on the output states can be described as follows. For simplicity we assume that the evolution is given by a self-adjoint operator H (equal to $J_{\mathbf{n}}$ in the above Mach-Zehnder interferometer), i.e.

$$\rho(\theta) = \mathrm{e}^{-\mathrm{i}\theta H}\rho\,\mathrm{e}^{\mathrm{i}\theta H}\,, \tag{9.19}$$

where $\rho = \rho(0) = \rho_{\rm in}$ is the input state. One performs generalized measurements given by a POVM $\{M_i\}_{i=1}^m$ on the output state $\rho(\theta) = \rho_{\rm out}$. The probability to get the outcome i is $p_{i|\theta} = {\rm tr}[M_i\rho(\theta)]$ (Sec. 5.3). After N independent measurements[4] on copies of $\rho(\theta)$ yielding the outcomes $i_1, i_2, \ldots, i_N$, the parameter θ is estimated by using a statistical estimator depending on these outcomes, that is, a function $\theta_{\rm est}(i_1, i_2, \ldots, i_N)$. The precision of the estimation is defined by the variance

$$\Delta\theta = \left\langle \left(\Big| \frac{\partial\langle\theta_{\rm est}\rangle_\theta}{\partial\theta} \Big|^{-1} \theta_{\rm est} - \theta \right)^2 \right\rangle_\theta^{\frac{1}{2}}, \tag{9.20}$$

where $\langle\,\cdot\,\rangle_\theta$ denotes the average for the product probability measure $\{p_{i_1|\theta}\ldots p_{i_N|\theta}\}_{i_1,\ldots,i_N=1}^m$ of the independent outcomes. The factor $|\partial\langle\theta_{\rm est}\rangle_\theta/\partial\theta|^{-1}$ is put in front of $\theta_{\rm est}$ to remove some possible differences in physical units between θ and its estimator $\theta_{\rm est}$ (see [45]). We restrict our attention to unbiased estimators satisfying $|\partial\langle\theta_{\rm est}\rangle_\theta/\partial\theta|^{-1}\langle\theta_{\rm est}\rangle_\theta = \theta$. For a given input state ρ, one looks for the smallest error $\Delta\theta$ that can be achieved. This involves two different optimization steps, associated to the optimization over (i) all possible estimators $\theta_{\rm est}$ and (ii) all possible measurements. The step (i) relies on a classical result in statistics known as the Cramér-Rao bound,

$$\big\langle (\Delta\theta_{\rm est})^2 \big\rangle_\theta \geq \frac{1}{N\mathcal{F}(\{p_{i|\theta}\})} \Big(\frac{\partial\langle\theta_{\rm est}\rangle_\theta}{\partial\theta} \Big)^2, \tag{9.21}$$

where $\Delta\theta_{\rm est} = \theta_{\rm est} - \langle\theta_{\rm est}\rangle_\theta$ and

$$\mathcal{F}(\{p_{i|\theta}\}) = \sum_{i=1}^m \frac{1}{p_{i|\theta}} \Big(\frac{\partial p_{i|\theta}}{\partial\theta} \Big)^2 \tag{9.22}$$

[3] In contrast, because of the absence of direct interactions between photons it is difficult to generate large numbers of photons having multipartite entanglement.

[4] In practice the experiment is repeated N times, starting from the same initial state ρ and in similar conditions, so that the quantum evolution can be considered to be the same at each run.

is the Fisher information. The inequality (9.21) is saturated asymptotically for $N \to \infty$ by the maximum-likelihood estimator. The second optimization step (ii) has been solved in Ref. [45], leading to the following important statement. Recall that the quantum Fisher information is defined as (see Sec. 8.5)

$$\mathcal{F}_Q(\rho, H) = 4(g_{\rm B})_\rho(-{\rm i}[H,\rho], -{\rm i}[H,\rho]) = 4d_{\rm B}(\rho, \rho + {\rm d}\rho)^2 \,, \tag{9.23}$$

where $g_{\rm B}$ is the Bures metric and ${\rm d}\rho = (\partial\rho/\partial\theta){\rm d}\theta = -{\rm i}[H,\rho]{\rm d}\theta$.

Proposition 9.2.1. (Braunstein and Caves [45]) *The smallest error $\Delta\theta$ that can be achieved in the parameter estimation is*

$$(\Delta\theta)_{\rm best} = \frac{1}{\sqrt{N}\sqrt{\mathcal{F}_Q(\rho, H)}} \,, \tag{9.24}$$

where N is the number of measurements and $\mathcal{F}_Q(\rho, H)$ is the quantum Fisher information. Thus $\Delta\theta \geq (\Delta\theta)_{\rm best}$ and the equality $\Delta\theta = (\Delta\theta)_{\rm best}$ can be reached asymptotically as $N \to \infty$.

It is worth noting that (9.24) can be interpreted as a generalized uncertainty principle [45]. In fact, if $\rho = |\Psi\rangle\langle\Psi|$ is a pure state, in view of the relation (8.43) between $\mathcal{F}_Q(\rho, H)$ and the square fluctuation $\langle(\Delta H)^2\rangle_\Psi$ of H, the bound $\Delta\theta \geq (\Delta\theta)_{\rm best}$ can be written as

$$\Delta\theta \, \langle(\Delta H)^2\rangle_\Psi^{\frac{1}{2}} \geq \frac{1}{2\sqrt{N}} \,. \tag{9.25}$$

In this uncertainty relation H plays the role of the variable conjugated to the parameter θ.

Proof. We present here a direct proof of (9.24) based on the results of chapter 8 (see [45] for an independent proof). Before that, let us explain how the classical Cramér-Rao bound is derived. By differentiating with respect to θ the identity

$$0 = \langle \Delta\theta_{\rm est}\rangle_\theta = \sum_{i_1,\ldots,i_N} p_{i_1|\theta} \ldots p_{i_N|\theta} \Delta\theta_{\rm est}(i_1, \ldots, i_N) \tag{9.26}$$

one obtains

$$0 = \sum_{i_1,\ldots,i_N} p_{i_1|\theta} \ldots p_{i_N|\theta} \sum_{\nu=1}^{N} \frac{\partial \ln p_{i_\nu|\theta}}{\partial\theta} \Delta\theta_{\rm est}(i_1, \ldots, i_N) - \frac{\partial\langle\theta_{\rm est}\rangle_\theta}{\partial\theta} \,. \tag{9.27}$$

Then the Cramér-Rao bound (9.21) readily follows from the Cauchy-Schwarz inequality. Of course, the interesting point is that equality can be achieved in the limit $N \to \infty$, but we will not dwell into that. Going back to the quantum problem, we rearrange (9.21) as

$$\frac{({\rm d}\theta)^2}{N} \leq (\Delta\theta)^2 \sum_{i=1}^{m} \frac{({\rm tr}[M_i {\rm d}\rho(\theta)])^2}{{\rm tr}[M_i\rho(\theta)]} \tag{9.28}$$

with ${\rm d}\rho(\theta) = (\partial\rho/\partial\theta){\rm d}\theta$. Now, by using Proposition 8.3.1 and performing an expansion up to the second order in ${\rm d}\rho$, one finds

$$\mathcal{F}_Q(\rho(\theta), H)({\rm d}\theta)^2 = \sup_{\{M_i\}} \left\{ \sum_{i=1}^{m} \frac{({\rm tr}[M_i {\rm d}\rho(\theta)])^2}{{\rm tr}[M_i\rho(\theta)]} \right\} . \tag{9.29}$$

Here, the supremum is over all POVMs $\{M_i\}$ and we have used $\sum_i {\rm tr}[M_i {\rm d}\rho(\theta)] = {\rm tr}[{\rm d}\rho(\theta)] = 0$. But $\mathcal{F}_Q(\rho(\theta), H) = \mathcal{F}_Q(\rho, H)$ as a consequence of (8.42), since $\rho(\theta)$ and ρ are related by a unitary evolution generated by H. Comparing (9.28) and (9.29), we conclude that $\inf_{\{M_i\}} \Delta\theta \geq (\Delta\theta)_{\rm best}$, with equality as $N \to \infty$ for the maximum likelihood estimator, as stated in the proposition. □

Before proceeding to derive upper bounds on $\mathcal{F}_Q(\rho, H)$, let us observe that the monotonicity of the Bures metric $g_{\rm B}$ implies [96]:

Corollary 9.2.2. *The quantum Fisher information $\mathcal{F}_Q(\rho, H)$ is convex in ρ.*

Proof. Given two states ρ_0 and ρ_1 on $\mathcal{H}$ and $\eta_0, \eta_1 \geq 0$, $\eta_0 + \eta_1 = 1$, we introduce the state $\widehat{\rho} = \eta_0\rho_0 \otimes |0\rangle\langle 0| + \eta_1\rho_1 \otimes |1\rangle\langle 1|$ on $\mathcal{H} \otimes \mathbb{C}^2$ as in the proof of Theorem 8.6.2. From the expression of $\mathcal{F}_Q$ in the right-hand side of (8.42) one deduces that

$$\mathcal{F}_Q(\widehat{\rho}, H \otimes 1) = \eta_0 \mathcal{F}_Q(\rho_0, H) + \eta_1 \mathcal{F}_Q(\rho_1, H) \,. \tag{9.30}$$

Let $\mathcal{T} : \widehat{\sigma} \mapsto \mathrm{tr}_{\mathbb{C}^2}(\widehat{\sigma})$ denote the partial trace on $\mathbb{C}^2$. Then $\mathcal{T}(\widehat{\rho}) = \rho = \eta_0\rho_0 + \eta_1\rho_1$ and $\mathcal{T}([H \otimes 1, \widehat{\rho}]) = [H, \rho]$. As $\mathcal{T}$ is a quantum operation, it results from the contractivity of the Bures metric that

$$(g_{\mathrm{B}})_{\widehat{\rho}}\big(-\mathrm{i}[H \otimes 1, \widehat{\rho}], -\mathrm{i}[H \otimes 1, \widehat{\rho}]\big) \geq (g_{\mathrm{B}})_{\rho}\big(-\mathrm{i}[H, \rho], -\mathrm{i}[H, \rho]\big) \; . \tag{9.31}$$

Collecting together (9.30) and (9.31) yields $\mathcal{F}_Q(\rho, H) \leq \eta_0\mathcal{F}_Q(\rho_0, H) + \eta_1\mathcal{F}_Q(\rho_1, H)$. □

9.2.3 Interferometer precision and inter-particle entanglement

We now show by relying on Proposition 9.2.1 that if the input state has N_{p} particles in a maximally entangled state, the precision $(\Delta\theta)_{\mathrm{best}}$ is smaller by a factor $1/\sqrt{N_{\mathrm{p}}}$ with respect to the precision obtained with separable input states. The Hilbert space of the particles is $\mathcal{H}^{(N_{\mathrm{p}})} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_{N_{\mathrm{p}}}$, $\mathcal{H}_\nu$ being the Hilbert space of the νth particle. Assuming that the particles do not interact between themselves, the Hamiltonian reads

$$H = \sum_{\nu=1}^{N_{\mathrm{p}}} 1 \otimes \cdots \otimes H_\nu \otimes \cdots \otimes 1 \; , \tag{9.32}$$

where H_ν acts on $\mathcal{H}_\nu$. To simplify the discussion we suppose that the single particle Hamiltonians H_ν have the same highest eigenvalue $\lambda_{\max}$ and the same lowest eigenvalue $\lambda_{\min}$. This is the case for instance if H is the angular momentum $J_{\mathbf{n}}$ in the interferometer of Sec. 9.2.1 (then $H_\nu = (n_x\sigma_{x\nu} + n_y\sigma_{y\nu} + n_z\sigma_{z\nu})/2$ with $|\mathbf{n}| = 1$ and $\sigma_{x\nu}$, $\sigma_{y\nu}$, and $\sigma_{z\nu}$ the three Pauli matrices acting on $\mathcal{H}_\nu \simeq \mathbb{C}^2$, so that $\lambda_{\max} = -\lambda_{\min} = 1/2$). Let us recall that the quantum Fisher information $\mathcal{F}_Q(|\Psi\rangle, H)$ of a pure state $|\Psi\rangle$ is given by the square fluctuation $\langle(\Delta H)^2\rangle_\Psi = \langle\Psi|H^2|\Psi\rangle - \langle\Psi|H|\Psi\rangle^2$ up to a factor of four (see Sec. 8.5). We first observe that the maximum of $\langle(\Delta H_\nu)^2\rangle_{\psi_\nu}$ over all pure states $|\psi_\nu\rangle \in \mathcal{H}_\nu$ is equal to $(\Delta h)^2 = (\lambda_{\max} - \lambda_{\min})^2/4$, the maximum being attained when $|\psi_\nu\rangle = (|\phi_{\nu,\max}\rangle + |\phi_{\nu,\min}\rangle)/\sqrt{2}$, where $|\phi_{\nu,\max}\rangle$ and $|\phi_{\nu,\min}\rangle$ are the eigenvectors of H_ν with eigenvalues $\lambda_{\max}$ and $\lambda_{\min}$, respectively. Let the N_{p} particles be in a separable state ρ_{sep} and let $\{|\Psi_i\rangle, \eta_i\}$ be a decomposition of ρ_{sep} into pure product states $|\Psi_i\rangle = |\psi_{i1}\rangle \otimes \cdots \otimes |\psi_{iN_{\mathrm{p}}}\rangle \in \mathcal{H}^{(N_{\mathrm{p}})}$. A simple calculation gives [100]

$$\mathcal{F}_Q(|\Psi_i\rangle, H) = 4\big\langle(\Delta H)^2\big\rangle_{\Psi_i} = 4\sum_{\nu=1}^{N_{\mathrm{p}}}\langle(\Delta H_\nu)^2\rangle_{\psi_{i\nu}} \; \leq \; 4(\Delta h)^2 N_{\mathrm{p}} \; . \tag{9.33}$$

By applying Corollary 9.2.2 we get

$$\rho_{\mathrm{sep}} \;\; \text{separable} \;\; \Rightarrow \;\; \mathcal{F}_Q(\rho_{\mathrm{sep}}, H) \leq 4(\Delta h)^2 N_{\mathrm{p}} \; . \tag{9.34}$$

According to Proposition 9.2.1 the phase precision of the interferometer satisfies for separable inputs

$$\Delta\theta \geq (\Delta\theta)_{\mathrm{SN}} = \frac{1}{2\Delta h\sqrt{N N_{\mathrm{p}}}} \; . \tag{9.35}$$

This means that separable input states cannot do better than N_{p} independent particles sent one-by-one through the interferometer, henceforth producing an error of the order of $1/\sqrt{N_{\mathrm{p}}}$. Note that (9.34) provides a sufficient condition $\mathcal{F}_Q(\rho, H) > 4(\Delta h)^2 N_{\mathrm{p}}$ for entanglement of ρ [197]. There are, however, entangled states which do not satisfy this criterion [197]. Such entangled states are not useful for interferometry, in the sense that they produce phase errors larger than the shot noise value $(\Delta\theta)_{\mathrm{SN}}$.

We now argue that much higher Fisher informations, of the order of N_{p}^2, can be achieved for entangled states. By the same observation as above, $\langle(\Delta H)^2\rangle_\Psi$ has a maximum given by the square of the half difference of the maximal and minimal eigenvalues of H. For the Hamiltonian (9.32), one immediately finds

$$\mathcal{F}_Q(|\Psi\rangle, H) \leq 4(\Delta h)^2 N_{\mathrm{p}}^2 \; . \tag{9.36}$$

This upper bound is often called the *Heisenberg bound* in the literature. It is saturated for the entangled states [100]

$$|\Psi_{\mathrm{ent}}^{\pm}\rangle = \frac{1}{\sqrt{2}}\Big(|\phi_{1,\max}\rangle|\phi_{2,\max}\rangle \ldots |\phi_{N_{\mathrm{p}},\max}\rangle \pm |\phi_{1,\min}\rangle|\phi_{2,\min}\rangle \ldots |\phi_{N_{\mathrm{p}},\min}\rangle\Big) \; . \tag{9.37}$$

For large N_{p} such states deserve the name of *macroscopic superpositions*, as they are formed by a superposition of two macroscopically distinct states in which each particle is in the highest energy eigenstate of the single particle Hamiltonian (for the first component of the superposition) or in the lowest energy eigenstate (for the second component). If one uses these superpositions as input states of the interferometer, an error of $\Delta\theta = 1/(2\Delta h\sqrt{N}N_{\mathrm{p}}) = (\Delta\theta)_{\mathrm{SN}}/\sqrt{N_{\mathrm{p}}}$ can be achieved asymptotically for $N \to \infty$ on the unknown phase. According to (9.24) and (9.36), this is the best possible precision.

Chapter 10

Measures of entanglement in bipartite systems

Maximale Kenntnis von einem Gesamtsystem schließt nicht notwendig maximale Kenntnis aller seiner Teile ein, auch dann nicht, wenn dieselben völlig voneinander abgetrennt sind und einander zur Zeit gar nicht beeinflussen.[1]
(E. Schrödinger 1935) [211]

Even if it would be better for many computational and communication tasks to work with maximally entangled pure states, in practice the coupling of the system with its environment transforms such states into non-maximally entangled mixed states because of the induced decoherence processes [46, 102, 113]. It is thus important to quantify the amount of entanglement in an arbitrary quantum state. Unfortunately, this amount of entanglement is not a directly measurable quantity. It is quantified by an entanglement measure, which vanishes if and only if the state is separable and cannot increase under local operations on each subsystems and classical communication (entanglement monotonicity). All measures satisfying these two requirements are not equivalent, i.e., a state ρ can be more entangled than a state σ for one measure and less entangled for the other. In this chapter, we investigate the properties of entanglement measures, give their general form for pure states, and study more especially two of the most popular ones, the entanglement of formation and the concurrence. We restrict our attention to bipartite entanglement (see [110, 130] for generalizations to entanglement in systems with more than two parties).

10.1 Entanglement as correlations between local measurements

Let $|\Psi\rangle$ be a pure state of a bipartite system AB. In view of the discussion in Sec. 4.4, it seems natural physically to characterize the entanglement in $|\Psi\rangle$ by maximizing the correlator $G_{AB}(|\Psi\rangle)$ in (4.17) over all local observables $A \in \mathcal{B}(\mathcal{H}_{\mathsf{A}})_{\text{s.a.}}$ and $B \in \mathcal{B}(\mathcal{H}_{\mathsf{B}})_{\text{s.a.}}$ and to define

$$G(|\Psi\rangle) = \max_{A=A^*, \|\Delta A\|_{\infty,\Psi} \leq 1} \; \max_{B=B^*, \|\Delta B\|_{\infty,\Psi} \leq 1} \big\{ \big| G_{AB}(|\Psi\rangle) \big| \big\} \; . \tag{10.1}$$

One must face with some arbitrariness on the choice of the norm used to bound $\Delta A = A - \langle A \otimes 1 \rangle_\Psi$ and $\Delta B = B - \langle 1 \otimes B \rangle_\Psi$. In order to obtain an entanglement measure with the required properties, we take the Ψ-dependent norm $\|\Delta A\|_{\infty,\Psi} = \max_{i,j} |\langle \alpha_i | \Delta A | \alpha_j \rangle|$, where $\{|\alpha_i\rangle\}$ is an orthonormal eigenbasis of the reduced state $[\rho_\Psi]_{\mathsf{A}}$, and similarly for $\|\Delta B\|_{\infty,\Psi}$ with the eigenbasis $\{|\beta_k\rangle\}$ of $[\rho_\Psi]_{\mathsf{B}}$. These norms correspond to the infinity norms of the vectors in $\mathcal{H}_{\mathsf{AA}}$ and $\mathcal{H}_{\mathsf{BB}}$ associated to ΔA and ΔB via the isometry (4.5). By using the Schmidt decomposition (4.9) and setting $A_{ij} = \langle \alpha_i | A | \alpha_j \rangle$ and $B_{ij} = \langle \beta_i | B | \beta_j \rangle$, one finds

$$G_{AB}(|\Psi\rangle) = \langle \Delta A \otimes \Delta B \rangle_\Psi = \sum_{i=1}^n \mu_i (\Delta A)_{ii} (\Delta B)_{ii} + \sum_{i \neq j}^n \sqrt{\mu_i \mu_j} A_{ij} B_{ij} \; . \tag{10.2}$$

The Cauchy-Schwarz inequality immediately yields

$$G(|\Psi\rangle) = \max_{\|\Delta \mathbf{a}\|_\infty \leq 1} \big\{ \sqrt{(\Delta \mathbf{a})^2} \big\} + C(|\Psi\rangle) \; , \tag{10.3}$$

[1] Maximal knowledge on a total system does not necessarily include maximal knowledge on all its parts, even if these are completely separated from each other and for now cannot affect each other.

where the overline stands for the average with respect to the Schmidt coefficients μ_i (e.g. $\overline{\mathbf{a}} = \sum_i \mu_i a_i$), $\Delta\mathbf{a} = \mathbf{a} - \overline{\mathbf{a}}$ with $\mathbf{a} = (A_{11}, \ldots, A_{nn})$, $\|\Delta\mathbf{a}\|_\infty = \max_i |(\Delta\mathbf{a})_i|$, and

$$C(|\Psi\rangle) = \sum_{i\neq j}^{n} \sqrt{\mu_i \mu_j} = \Big(\mathrm{tr}\big(\sqrt{[\rho_\Psi]_{\mathsf{A}}}\big)\Big)^2 - 1 \; . \tag{10.4}$$

Thus $C(|\Psi\rangle) = 0$ (similarly, $G(|\Psi\rangle) = 0$) is equivalent to $\mu_i = 0$ save for one index i, that is, to $|\Psi\rangle$ being separable. Furthermore, $C(|\Psi\rangle) \leq n-1$ with equality if and only if $\mu_i = 1/n$ for all i, that is, if and only if $|\Psi\rangle$ is maximally entangled (Sec. 4.4)[2]. Finally, we note that G and C are invariant under local unitaries, i.e., $G(U_A \otimes U_B|\Psi\rangle) = G(|\Psi\rangle)$ for any unitaries U_A and U_B on $\mathcal{H}_{\mathsf{A}}$ and $\mathcal{H}_{\mathsf{B}}$. For two qubits one obtains

$$G(|\Psi\rangle) = \mu_{\max}^{-1} - 1 + C(|\Psi\rangle) \quad , \quad C(|\Psi\rangle) = 2\sqrt{\mu_0 \mu_1} \tag{10.5}$$

with $\mu_{\max} = \max\{\mu_0, \mu_1\}$. It is easy to show that $C(|\Psi\rangle) = |\langle\Psi|\sigma_y \otimes \sigma_y J|\Psi\rangle|$, where $\sigma_y = \mathrm{i}(|0\rangle\langle 1| - |1\rangle\langle 0|)$ is the y-Pauli matrix and J the complex conjugation in the canonical basis. This quantity has been first introduced by Wootters [265] and is known as the *concurrence*.

One may wonder how the correlator G_{AB} could be generalized for mixed states. The first guess would be to replace the expectation value $\langle\cdot\rangle_\Psi$ by $\langle\cdot\rangle_\rho = \mathrm{tr}(\rho\,\cdot\,)$, but one easily sees that then $G(\rho)$ can be non-zero even for separable mixed states, because this correlator contains both the quantum and classical (i.e., statistical) correlations in the density matrix ρ. Noting that

$$G_{AB}(|\Psi\rangle) = \frac{1}{2}\big\langle \big(\Delta(A\otimes 1 + 1\otimes B)\big)^2\big\rangle_\Psi - \frac{1}{2}\big\langle \big(\Delta(A\otimes 1)\big)^2\big\rangle_\Psi - \frac{1}{2}\big\langle \big(\Delta(1\otimes B)\big)^2\big\rangle_\Psi \; , \tag{10.6}$$

it is tempting to define a correlator for ρ in terms of the quantum Fisher information (8.42), i.e., of the Bures metric g_{B},

$$\begin{aligned} G_{AB}(\rho) &= \frac{1}{8}\Big(\mathcal{F}_Q(\rho, A\otimes 1 + 1\otimes B) - \mathcal{F}_Q(\rho, A\otimes 1) - \mathcal{F}_Q(\rho, 1\otimes B)\Big) \\ &= \mathrm{Re}\,\big\{(g_{\mathrm{B}})_\rho\big(-\mathrm{i}[A\otimes 1, \rho], -\mathrm{i}[1\otimes B, \rho]\big)\big\} \; . \end{aligned}$$

By inspection on (8.43), $G_{AB}(\rho)$ reduces for pure states to the previous correlator. However, the maximum of $|G_{AB}(\rho)|$ over all A and B does not fulfill the axioms of an entanglement measure. We will see in Sec. 10.4 another way to define the concurrence C for mixed states, by using on a convex roof construction.

10.2 LOCC operations

The main physical postulate on entanglement measures is that they must be monotonous with respect to certain state transformations. Such transformations that cannot increase entanglement are called *Local Operations and Classical Communication* (LOCC) and can be described as follows [34, 130]. Let us consider an entangled state ρ shared by two observers Alice and Bob. Alice and Bob can perform any quantum operations $\mathcal{M}_{\mathsf{A}} : \mathcal{B}(\mathcal{H}_{\mathsf{A}}) \to \mathcal{B}(\mathcal{H}'_{\mathsf{A}})$ and $\mathcal{M}_{\mathsf{B}} : \mathcal{B}(\mathcal{H}_{\mathsf{B}}) \to \mathcal{B}(\mathcal{H}'_{\mathsf{B}})$ on their respective subsystems A and B. Here, the final spaces $\mathcal{H}'_{\mathsf{A}}$ and $\mathcal{H}'_{\mathsf{B}}$ may include local ancillae, or may be some subspaces of $\mathcal{H}_{\mathsf{A}}$ and $\mathcal{H}_{\mathsf{B}}$, respectively. The corresponding transformations on the system AB are called *local quantum operations*. They are of the form $\mathcal{M}_{\mathrm{loc}} = \mathcal{M}_{\mathsf{A}} \otimes \mathcal{M}_{\mathsf{B}}$ and are given by families $\{A_i \otimes B_j\}$ of Kraus operators, where A_i and B_j are local observables on A and B. Local operations are performed physically by coupling each subsystem to a local ancilla and by making joint unitary evolutions and von Neumann measurements on the subsystem and its ancilla (see Sec. 5.2). Such processes can clearly not increase the amount of entanglement between A and B. In addition to performing local generalized measurements, Alice and Bob can communicate their measurement outcomes to each other via a classical communication channel (two-way communication). No transfer of quantum systems between them is allowed. Thanks to classical communication, the observers can increase the classical correlations between A and B, but not the AB-entanglement. A *LOCC operation* is a quantum operation on $\mathcal{B}(\mathcal{H}_{\mathsf{AB}})$ obtained through a succession of the aforementioned actions of Alice and Bob, taken in arbitrary order. For example, if Alice performs a measurement on A and Bob a measurement on B depending on Alice's outcome i (one way communication), the post-measurement state in the absence of readout is

$$\mathcal{M}_{1-\mathrm{way}}(\rho) = \sum_i 1 \otimes \mathcal{M}_{\mathsf{B}}^{(i)}(A_i \otimes 1 \rho A_i^* \otimes 1) \; . \tag{10.7}$$

[2]This last property is not true if one uses the operator norm instead of $\|\cdot\|_{\infty,\Psi}$ in (10.1), except in the two-qubit case $n = 2$.

This defines a LOCC operation with Kraus operators $A_i \otimes B_j^{(i)}$, where $\sum_i A_i^* A_i = \sum_j (B_j^{(i)})^* B_j^{(i)} = 1$.

Any LOCC operation can be obtained by composing local operations $\mathcal{M}_{\rm loc}$ with the maps

$$\mathcal{M}^{\mathsf{A}}_{\rm LOCC}(\rho) = \sum_i \big(A_i \otimes 1\, \rho\, A_i^* \otimes 1\big) \otimes |\kappa_i\rangle\langle\kappa_i| \quad , \quad \mathcal{M}^{\mathsf{B}}_{\rm LOCC}(\rho) = \sum_j \big(1 \otimes B_j\, \rho\, 1 \otimes B_j^*\big) \otimes |\epsilon_j\rangle\langle\epsilon_j| \,, \tag{10.8}$$

where $\sum_i A_i^* A_i = \sum_j B_j^* B_j = 1$ and $\{|\kappa_i\rangle\}$ (respectively $\{|\epsilon_j\rangle\}$) is an orthonormal basis for Bob's ancilla (respectively Alice's ancilla) [130]. A strictly larger but much simpler class of transformations, known as the *separable quantum operations* [246], is the set of all operations with Kraus operators $A_i \otimes B_i$, i.e.

$$\mathcal{M}_{\rm sep}(\rho) = \sum_i A_i \otimes B_i \rho A_i^* \otimes B_i^* \tag{10.9}$$

with $A_i \in \mathcal{B}(\mathcal{H}_{\mathsf{A}}, \mathcal{H}'_{\mathsf{A}})$, $B_i \in \mathcal{B}(\mathcal{H}_{\mathsf{B}}, \mathcal{H}'_{\mathsf{B}})$, and $\sum_i A_i^* A_i \otimes B_i^* B_i = 1$. The local operations and maps (10.8) being separable, any LOCC operation is separable. A result from Ref. [35] shows, however, that certain separable operations are not LOCCs.

It is clear that the set $\mathcal{S}_{\mathsf{AB}}$ of separable states is invariant under separable operations. It is also true that every separable state can be converted into any other separable state by a separable operation. Actually, any separable state can be obtained from the classical state $\rho_{\rm clas} = \sum_{jk} p_{jk} |j\rangle\langle j| \otimes |k\rangle\langle k|$ by such an operation (take $A_{ijk} = \sqrt{\eta_i}|\psi_i\rangle\langle j|$ and $B_{ijk} = |\phi_i\rangle\langle k|$ with η_i, $|\psi_i\rangle$, and $|\phi_i\rangle$ as in (4.19)). Furthermore, an arbitrary state ρ can be transformed into a classical state $\rho_{\rm clas}$ by a measurement in the product basis $\{|j\rangle|k\rangle\}$, which is a local operation.

When one restricts LOCC transformations to pure states, a great simplification comes from the following observation. If the space dimensions of A and B are such that $n_{\mathsf{A}} \geq n_{\mathsf{B}}$, any measurement by Bob can be simulated by a measurement by Alice followed by a unitary transformation by Bob conditioned to Alice's outcome (such a conditioning is allowed as Alice and Bob can communicate classically). In fact, let $\{|\alpha_i\rangle\}_{i=1}^{n_{\mathsf{A}}}$ and $\{|\beta_i\rangle\}_{i=1}^{n_{\mathsf{B}}}$ be orthonormal eigenbasis of the reduced states $[\rho_\Psi]_{\mathsf{A}}$ and $[\rho_\Psi]_{\mathsf{B}}$, and let B_i be the Kraus operators describing Bob's measurement. Consider the measurement done by Alice with Kraus operators $A_i = \sum_{j,l} (B_i)_{lj} |\alpha_l\rangle\langle\alpha_j|$, where $(B_i)_{lj} = \langle\beta_l|B_i|\beta_j\rangle$. The unnormalized post-measurement states

$$|\widetilde{\Phi}_i\rangle = 1 \otimes B_i|\Psi\rangle = \sum_{j,l} \sqrt{\mu_j} (B_i)_{lj} |\alpha_j\rangle|\beta_l\rangle \quad , \quad |\widetilde{\Phi}'_i\rangle = A_i \otimes 1|\Psi\rangle = \sum_{j,l} \sqrt{\mu_j} (B_i)_{lj} |\alpha_l\rangle|\beta_j\rangle \tag{10.10}$$

have the same Schmidt coefficients because $\mathrm{tr}_{\mathsf{B}}(|\widetilde{\Phi}_i\rangle\langle\widetilde{\Phi}_i|)$ and $\mathrm{tr}_{\mathsf{A}}(|\widetilde{\Phi}'_i\rangle\langle\widetilde{\Phi}'_i|)$ are related by an isometry $\mathcal{H}_{\mathsf{A}} \to \mathcal{H}_{\mathsf{B}}$. Thus $|\widetilde{\Phi}'_i\rangle = U_i \otimes V_i |\widetilde{\Phi}_i\rangle$ for some local unitaries U_i on $\mathcal{H}_{\mathsf{A}}$ and V_i on $\mathcal{H}_{\mathsf{B}}$. Consequently, Bob performing the measurement $\{B_i\}$ is equivalent to Alice performing the measurement $\{U_i^* A_i\}$ and Bob performing the unitary transformation V_i^* when Alice gets the outcome i. Applying this result to all Bob's measurements, we conclude that a LOCC acting on a pure state $|\Psi\rangle$ may always be simulated by a one-way communication protocol involving only three steps: (1) Alice first performs a generalized measurement on subsystem A; (2) she sends her measurement result to Bob; (3) Bob performs a unitary evolution on B conditional to Alice's result.

Based on this observation, we say that a pure state $|\Psi\rangle \in \mathcal{H}_{\mathsf{AB}}$ can be transformed by a LOCC into the pure state $|\Phi\rangle \in \mathcal{H}_{\mathsf{AB}}$ if there are families of Kraus operators $\{A_i\}$ on $\mathcal{H}_{\mathsf{A}}$ and unitaries $\{V_i\}$ on $\mathcal{H}_{\mathsf{B}}$ such that all unnormalized conditional states $A_i \otimes V_i|\Psi\rangle$ are proportional to $|\Phi\rangle$, irrespective of the measurement outcome i. Note that this is equivalent to $\mathcal{M}_{\rm LOCC}(|\Psi\rangle\langle\Psi|)$ being equal to $|\Phi\rangle\langle\Phi|$, with $\mathcal{M}_{\rm LOCC}$ the LOCC operation with Kraus family $\{A_i \otimes V_i\}$. One defines in this way an order relation on the set of pure states. Nielsen [179] discovered a nice relation between this order and the theory of majorization for n-dimensional vectors [38]. Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ be two vectors in $\mathbb{R}^n$. We denote by $\mathbf{x}^\downarrow$ the vector formed by the components of $\mathbf{x}$ in decreasing order, and similarly for $\mathbf{y}^\downarrow$. One says that $\mathbf{x}$ is majorized by $\mathbf{y}$ and write $\mathbf{x} \prec \mathbf{y}$ if $\sum_{i=1}^k x_i^\downarrow \leq \sum_{i=1}^k y_i^\downarrow$ for any $k = 1, \ldots, n$, with equality instead of inequality for $k = n$.

Proposition 10.2.1. (Nielsen [179]) *A pure state $|\Psi\rangle$ of the bipartite system AB can be transformed into another pure state $|\Phi\rangle$ of AB by a LOCC if and only if $\boldsymbol{\mu}_\Psi \prec \boldsymbol{\mu}_\Phi$, where $\boldsymbol{\mu}_\Psi$ and $\boldsymbol{\mu}_\Phi$ are the vectors formed by the Schmidt coefficients of $|\Psi\rangle$ and $|\Phi\rangle$, respectively.*

A detailed proof of this result can be found in [180] (Sect. 12.5), so we omit it here. This proof relies on the following theorem: if $\boldsymbol{\lambda}_H$ and $\boldsymbol{\lambda}_K$ are vectors formed by the eigenvalues of two Hermitian matrices H and K, respectively, then $\boldsymbol{\lambda}_H \prec \boldsymbol{\lambda}_K$ if and only if $H = \sum_i \eta_i U_i K U_i^*$ with $\{\eta_i\}$ a set of probabilities and U_i some unitary matrices.

Remark 10.2.2. *Even if $|\Psi\rangle$ cannot be transformed into $|\Phi\rangle$ by a LOCC, it may still happen that $|\Psi\rangle\otimes|\kappa\rangle$ can be transformed into $|\Phi\rangle\otimes|\kappa\rangle$ by a LOCC (here the state of the ancilla does not change during the transformation, i.e., it acts as catalysts in chemical reactions) [143].*

10.3 Axioms on entanglement measures

We are now in position to formulate the physical postulates on entanglement measures [34, 246, 248].

Definition 10.3.1. *An entanglement measure of a bipartite system* AB *is a function* $E : \mathcal{E}(\mathcal{H}_{\mathsf{AB}}) \to \mathbb{R}$ *such that*

(i) $E(\rho) = 0$ *if and only if* ρ *is separable;*

(ii) E *is convex;*

(iii) E *cannot increase under LOCCs, i.e., if* $\mathcal{M}_{\mathrm{LOCC}}$ *is a LOCC operation then* $E(\mathcal{M}_{\mathrm{LOCC}}(\rho)) \leq E(\rho)$.

As any two separable states can be transformed one into each other by means of a LOCC operation, the monotonicity (iii) implies that E is constant on the set of separable states $\mathcal{S}_{\mathsf{AB}}$. Taking this constant equal to zero yields $\rho \in \mathcal{S}_{\mathsf{AB}} \Rightarrow E(\rho) = 0$, so that only the reverse implication is needed in (i). Furthermore, any state ρ can be converted into a separable state by a LOCC, thus $E(\rho)$ is minimum for separable states and $E(\rho) \geq 0$. The convexity condition (ii) is motivated by the following observation [248]. Assume that Alice and Bob share m pairs of particles in the states $\rho_1, \ldots, \rho_m$. By classical communication, they can agree to keep the ith pair with probability η_i, thus preparing the ensemble $\{\rho_i, \eta_i\}_{i=1}^m$. By erasing the information about which state ρ_i was kept, the state becomes $\rho = \sum \eta_i \rho_i$ (see Sec. 4.3). The inequality $E(\rho) \leq \sum_i \eta_i E(\rho_i)$ means that this local loss of information does not increase the average entanglement.

It results from the monotonicity (iii) that entanglement measures are invariant under conjugations by local unitaries, i.e., $E(U_{\mathsf{A}} \otimes U_{\mathsf{B}}\, \rho\, U_{\mathsf{A}}^* \otimes U_{\mathsf{B}}^*) = E(\rho)$. For pure states $|\Psi\rangle$, this implies that $E(|\Psi\rangle)$ only depends on the Schmidt coefficients μ_i of $|\Psi\rangle$. Consequently, $E(|\Psi\rangle) = f([\rho_\Psi]_{\mathsf{A}})$ is a unitary-invariant function of the reduced state $[\rho_\Psi]_{\mathsf{A}} = \mathrm{tr}_{\mathsf{B}}(|\Psi\rangle\langle\Psi|)$ (or, equivalently, of $[\rho_\Psi]_{\mathsf{B}} = \mathrm{tr}_{\mathsf{A}}(|\Psi\rangle\langle\Psi|)$). Given that a pure state is separable if and only if it has a single non-vanishing Schmidt coefficient, one deduces from axiom (i) that $f(\rho_{\mathsf{A}})$ vanishes if and only if ρ_{A} is of rank one. The result below due to Vidal [248] characterizes all entanglement measures on pure states satisfying a slightly stronger condition than (iii). This shows in particular that there are many measures of entanglement fulfilling the three physical requirements (i-iii) of Definition 10.3.1, given by concave functions f.

Proposition 10.3.2. (Vidal [248]) *Let* $f : \mathcal{E}(\mathcal{H}_{\mathsf{A}}) \to \mathbb{R}$ *be concave, unitary invariant, and such that* $f(\rho_{\mathsf{A}}) = 0$ *if and only if* ρ_{A} *is a pure state. Then*

$$E_f(|\Psi\rangle) = f([\rho_\Psi]_{\mathsf{A}}) \tag{10.11}$$

defines an entanglement measure on the set of pure states of AB*, which satisfies the monotonicity condition*

(iii') $\sum_i p_i E(|\Phi_i\rangle) \leq E(|\Psi\rangle)$*, where* $p_i = \|A_i \otimes B_i|\Psi\rangle\|^2$ *and* $|\Phi_i\rangle = p_i^{-1/2} A_i \otimes B_i|\Psi\rangle$ *and the probabilities and conditional states of a separable measurement with Kraus operators* $A_i \otimes B_i$.

Conversely, any entanglement measure on pure states fulfilling (iii') is given by (10.11) for some function f *satisfying the above assumptions.*

It should be noted that asking $E(|\Phi_i\rangle) \leq E(|\Psi\rangle)$ for all outcomes i would put a too strong condition on E. Indeed, local measurements can in principle create entanglement on some conditional states, but not on average (see below).

Proof. Let f be like in the proposition. We have already argued above that E_f fulfills axiom (i), and (ii) is empty because of the restriction to pure states. Recall that for such states any measurement on B can be simulated by a measurement on A followed by a unitary operation on B conditioned to the measurement result. Hence it suffices to show the monotonicity (iii') for $B_i = V_i$ unitary. Let us set $\rho_{\mathsf{B}|i} = \mathrm{tr}_{\mathsf{A}}(|\Phi_i\rangle\langle\Phi_i|)$. Then $\{V_i^* \rho_{\mathsf{B}|i} V_i, p_i\}$ is a pure state decomposition of $[\rho_\Psi]_{\mathsf{B}}$, i.e., $\sum_i p_i V_i^* \rho_{\mathsf{B}|i} V_i = [\rho_\Psi]_{\mathsf{B}}$. This can be interpreted

by saying that a local measurement on A does not modify the state of B when B has no information on the measurement outcomes[3]. The concavity and unitary invariance of f imply

$$\sum_i p_i E_f(|\Phi_i\rangle) = \sum_i p_i f(V_i^* \rho_{\mathsf{B}|i} V_i) \leq f([\rho_\Psi]_\mathsf{B}) = E_f(|\Psi\rangle) \; . \tag{10.12}$$

This shows (iii'). Thus E_f is an entanglement measure.

Reciprocally, let E be an entanglement measure on pure states satisfying (iii'). From the discussion before the proposition we know that $E(|\Psi\rangle) = f([\rho_\Psi]_\mathsf{A}) = f([\rho_\Psi]_\mathsf{B})$ for some unitary-invariant function f vanishing on pure states only. It remains to show that f is concave. We may assume that the space dimensions of A and B are such that $n_\mathsf{A} \leq n_\mathsf{B}$ (otherwise one can exchange the role of A and B in the arguments below). Let ρ_A be an arbitrary state of A and $\sigma_\mathsf{A}^{(1)}$, $\sigma_\mathsf{A}^{(2)}$ be such that $\rho_\mathsf{A} = p_1\sigma_\mathsf{A}^{(1)} + p_2\sigma_\mathsf{A}^{(2)}$ with $p_1 + p_2 = 1$. As $n_\mathsf{A} \leq n_\mathsf{B}$, one may find a purification $|\Psi\rangle$ of ρ_A on $\mathcal{H}_{\mathsf{AB}}$ (Sec. 4.3). If one can exhibit a measurement on B with outcome probabilities p_i and conditional states $|\Phi_i\rangle$ having marginals $\mathrm{tr}_\mathsf{B}(|\Phi_i\rangle\langle\Phi_i|) = \sigma_\mathsf{A}^{(i)}$ for $i = 1, 2$, then the concavity of f can be deduced from (iii') thanks to the bound

$$f(\rho_\mathsf{A}) = E(|\Psi\rangle) \geq p_1 E(|\Phi_1\rangle) + p_2 E(|\Phi_2\rangle) = p_1 f(\sigma_\mathsf{A}^{(1)}) + p_2 f(\sigma_\mathsf{A}^{(2)}) \; . \tag{10.13}$$

The measurement we are looking for is just the square root measurement associated to $\{\sigma_\mathsf{A}^{(i)}, p_i\}$ (Sec. 5.5.3). Indeed, let $\{|\alpha_j\rangle\}_{j=1}^{n_\mathsf{A}}$ and $\{|\beta_k\rangle\}_{k=1}^{n_\mathsf{B}}$ be eigenbases of $[\rho_\Psi]_\mathsf{A}$ and $[\rho_\Psi]_\mathsf{B}$ and M_i^{lsm}, $i = 1, 2$, be the operators on $\mathcal{H}_\mathsf{B}$ with matrix elements given by (compare with (5.36))

$$\langle\beta_j|M_i^{\mathrm{lsm}}|\beta_l\rangle = \begin{cases} p_i \langle\alpha_l|\rho_\mathsf{A}^{-\frac{1}{2}} \sigma_\mathsf{A}^{(i)} \rho_\mathsf{A}^{-\frac{1}{2}}|\alpha_j\rangle & \text{if } j, l = 1, \ldots, n_\mathsf{A} \\ 0 & \text{otherwise.} \end{cases} \tag{10.14}$$

If $n_\mathsf{B} > n_\mathsf{A}$ we add a third measurement operator, equal to the projector onto $\mathrm{span}\{|\beta_k\rangle; n_\mathsf{A} < k \leq n_\mathsf{B}\}$. Then $M_1^{\mathrm{lsm}} + M_2^{\mathrm{lsm}} + M_3^{\mathrm{lsm}} = 1$. With the help of the Schmidt decomposition (4.9) one finds that $\langle\Psi|1 \otimes M_i^{\mathrm{lsm}}|\Psi\rangle$ equals p_i for $i = 1, 2$ and zero for $i = 3$, and the conditional state $|\Phi_i\rangle = p_i^{-1/2} 1 \otimes \sqrt{M_i^{\mathrm{lsm}}}|\Psi\rangle$ has marginal $\mathrm{tr}_\mathsf{B}(|\Phi_i\rangle\langle\Phi_i|) = \sigma_\mathsf{A}^{(i)}$ for $i = 1, 2$. This concludes the proof. □

Proposition 10.3.2 can be partially justified with the help of Proposition 10.2.1. More precisely, the latter implies that $E_f(|\Psi\rangle) \geq E_f(|\Phi\rangle)$ if $|\Phi\rangle\langle\Phi| = \mathcal{M}_{\mathrm{LOCC}}(|\Psi\rangle\langle\Psi|)$, that is, if there exists a LOCC measurement on $|\Psi\rangle$ with all conditional states $|\Phi_i\rangle$ equal to $|\Phi\rangle$. This comes from the fact that, by unitary invariance, $f([\rho_\Psi]_\mathsf{A})$ is a symmetric function of the eigenvalues $(\mu_\Psi)_1, \ldots, (\mu_\Psi)_n$ of $[\rho_\Psi]_\mathsf{A}$. But concave symmetric functions $f : \mathbb{R}^n \to \mathbb{R}$ are Schur-concave, i.e., $\mathbf{x} \prec \mathbf{y} \Rightarrow f(\mathbf{x}) \geq f(\mathbf{y})$ (see [38], Theorem II.3.3).

Many entanglement measures satisfying the axioms (i-iii) of Definition 10.3.1 have been defined in the literature. Their restrictions to pure states are all given by (10.11) for specific concave functions f. We present in the next section a few of these measures, namely, the entanglement of formation, the concurrence, and the Schmidt number. An integer-valued entanglement measure has been introduced in [208] by using a symplectic geometry approach, but this goes beyond the scope of this article.

10.4 Entanglement of formation

10.4.1 Entanglement of formation for pure states

A natural choice for the function f is the von Neumann entropy. We set

$$E_{\mathrm{EoF}}(|\Psi\rangle) = S\big([\rho_\Psi]_\mathsf{A}\big) = S\big([\rho_\Psi]_\mathsf{B}\big) = -\sum_i \mu_i \ln \mu_i \; . \tag{10.15}$$

Then $E_{\mathrm{EoF}}(|\Psi\rangle) = 0$ if and only if $|\Psi\rangle$ is separable and $E_{\mathrm{EoF}}(|\Psi\rangle)$ is maximum (and equal to $\ln n$ with $n = \min\{n_\mathsf{A}, n_\mathsf{B}\}$) if and only if $|\Psi\rangle$ is maximally entangled. Since the von Neumann entropy is concave, Proposition 10.3.2 ensures that E_{EoF} is an entanglement measure on pure states.

An important result due to Bennett *et al.* [33] relates $E_{\mathrm{EoF}}(|\Psi\rangle)$ to entanglement distillation and entanglement cost, which consist in the following problems. The EPR two-qubit state $|\Phi_+\rangle = (|0\rangle|0\rangle + |1\rangle|1\rangle)/\sqrt{2} \in \mathbb{C}^4$ corresponds to an *e*-bit of information shared by Alice and Bob. One such *e*-bit is required, for instance, if Alice

[3] If this would not be true, information could be sent faster than light in contradiction with Einstein's principle of relativity [194].

wants to teleport an unknown quantum state to Bob [180]. Entanglement distillation is the transformation of N copies of $|\Psi\rangle$ onto $M < N$ copies of $|\Phi_+\rangle$. It was demonstrated by Bennett *et al.* that in the large N limit, $E_{\rm EoF}(|\Psi\rangle)$ is equal to the maximal rate of distillation M/N, the maximum being over all LOCC operations. Stated differently, $E_{\rm EoF}(|\Psi\rangle)$ is the highest number of e-bits per input copy of $|\Psi\rangle$ that can be distilled from $|\Psi\rangle$ via LOCCs. Conversely, $E_{\rm EoF}(|\Psi\rangle)$ is the smallest number of e-bits per unit copy of $|\Psi\rangle$ from which $|\Psi\rangle$ may be obtained via LOCCs. The precise mathematical statement is given in the proposition below.

Proposition 10.4.1. (Bennett *et al.* [33])

$$\frac{E_{\rm EoF}(|\Psi\rangle)}{\ln 2} = \sup\Big\{r\,;\,\lim_{N\to\infty}\Big(\inf_{\rm LOCC}\big\|\mathcal{M}^{(N)}_{\rm LOCC}(|\Psi^{\otimes N}\rangle\langle\Psi^{\otimes N}|)-|\Phi_+^{\otimes rN}\rangle\langle\Phi_+^{\otimes rN}|\big\|_1\Big)=0\Big\} \tag{10.16}$$

$$= \inf\Big\{r\,;\,\lim_{N\to\infty}\Big(\inf_{\rm LOCC}\big\||\Psi^{\otimes N}\rangle\langle\Psi^{\otimes N}|-\mathcal{M}^{(N)}_{\rm LOCC}(|\Phi_+^{\otimes rN}\rangle\langle\Phi_+^{\otimes rN}|)\big\|_1\Big)=0\Big\}\,. \tag{10.17}$$

Let us stress that these identities are no longer valid for mixed states: then the right-hand sides of (10.16) and (10.17) are, in general, not equal. They define two measures of entanglement called the distillable entanglement and the entanglement cost (see [130] and references therein). The fact that these quantities coincide with $E_{\rm EoF}(|\Psi\rangle)$ for pure states basically indicates that, among all the possible entanglement measures, only one (namely $E_{\rm EoF}(|\Psi\rangle)$) becomes relevant asymptotically when dealing with many copies of $|\Psi\rangle$.

Proof. A simple and illuminating proof due to Nielsen [179] is based on Proposition 10.2.1 and the Shannon equipartition theorem. It runs as follows. Let μ_i be the Schmidt coefficients of $|\Psi\rangle$. Consider N i.i.d. random variables with distribution $\{\mu_i\}$ and values in $I=\{1,\dots,n\}$. The joint probabilities of these random variables are $p(\underline{i}) = \mu_{i_1}\dots\mu_{i_N}$ with $\underline{i}=(i_1,\dots,i_N)\in I^N$. Given $\varepsilon>0$, the "most likely set" $\mathcal{A}_{N,\varepsilon}\subset I^N$ is by definition the set of all $\underline{i}\in I^N$ such that $2^{-N(H+\varepsilon)}\le p(\underline{i})\le 2^{-N(H-\varepsilon)}$, H being the Shannon entropy of $\{\mu_i\}$, which is defined here by using the binary logarithm (in our case, $H = E_{\rm EoF}(|\Psi\rangle)/\ln 2$). The Shannon equipartition theorem [215] tells us that $\mathcal{A}_{N,\varepsilon}$ has probability $P_{N,\varepsilon}>1-\varepsilon$ and cardinality $|\mathcal{A}_{N,\varepsilon}|$ satisfying $(1-\varepsilon)2^{N(H-\varepsilon)}\le|\mathcal{A}_{N,\varepsilon}|\le 2^{N(H+\varepsilon)}$ for sufficiently large N. The idea of Nielsen's proof is to approximate

$$|\Psi^{\otimes N}\rangle = \sum_{\underline{i}\in I^N}\sqrt{p(\underline{i})}\,|\alpha_{i_1}\rangle\dots|\alpha_{i_N}\rangle\otimes|\beta_{i_1}\rangle\dots|\beta_{i_N}\rangle$$

$$\simeq |\Phi_{N,\varepsilon}\rangle = \sum_{\underline{i}\in\mathcal{A}_{N,\varepsilon}}\sqrt{q(\underline{i})}\,|\alpha_{i_1}\rangle\dots|\alpha_{i_N}\rangle\otimes|\beta_{i_1}\rangle\dots|\beta_{i_N}\rangle \tag{10.18}$$

with $q(\underline{i})=p(\underline{i})/P_{N,\varepsilon}$ and $|\alpha_i\rangle$, $|\beta_i\rangle$ as in Theorem 4.2.1. Observe that the fidelity $|\langle\Psi^{\otimes N}|\Phi_{N,\varepsilon}\rangle|^2 = P_{N,\varepsilon}$ is almost one for small ε. For any $\mathcal{A}\subset|\mathcal{A}_{N,\varepsilon}|$, one has

$$\frac{(1-\varepsilon)|\mathcal{A}|\,2^{-2N\varepsilon}}{|\mathcal{A}_{N,\varepsilon}|}\le\sum_{\underline{i}\in\mathcal{A}}q(\underline{i})\le\frac{|\mathcal{A}|\,2^{2N\varepsilon}}{(1-\varepsilon)|\mathcal{A}_{N,\varepsilon}|}\,. \tag{10.19}$$

The second inequality implies that $\mathbf{q}=(q(\underline{i}))_{i\in\mathcal{A}_{N,\varepsilon}}\prec(2^{-M},\dots,2^{-M},0,\dots,0)$ with

$$M=\ln_2(|\mathcal{A}_{N,\varepsilon}|(1-\varepsilon))-2N\varepsilon\,. \tag{10.20}$$

By Proposition 10.2.1, this means that $|\Phi_{N,\varepsilon}\rangle$ can be transformed by a LOCC into the M-qubit state

$$|\Phi_+^{\otimes M}\rangle=\sum_{\underline{j}\in\{0,1\}^M}2^{-\frac{M}{2}}|j_1\rangle\dots|j_M\rangle\otimes|j_1\rangle\dots|j_M\rangle\,. \tag{10.21}$$

We conclude that for N sufficiently large there exists a LOCC operation $\mathcal{M}^{(N,\varepsilon)}_{\rm LOCC}$ from $\mathcal{B}(\mathcal{H}_{\sf AB}^{\otimes N})$ into $\mathcal{B}(\mathbb{C}^{\otimes 2M})$ such that

$$\big\|\mathcal{M}^{(N,\varepsilon)}_{\rm LOCC}(|\Psi^{\otimes N}\rangle\langle\Psi^{\otimes N}|)-|\Phi_+^{\otimes M}\rangle\langle\Phi_+^{\otimes M}|\big\|_1 \le \big\||\Psi^{\otimes N}\rangle\langle\Psi^{\otimes N}|-|\Phi_{N,\varepsilon}\rangle\langle\Phi_{N,\varepsilon}|\big\|_1$$

$$\le 2\big(1-\big|\langle\Psi^{\otimes N}|\Phi_{N,\varepsilon}\rangle\big|^2\big)^{\frac12}\le 2\sqrt{\varepsilon} \tag{10.22}$$

(we have used Propositions 8.1.2 and 8.4.1 to get the first and second inequalities, respectively). In addition, the distillation rate M/N is bounded from below by $H-3\varepsilon+2N^{-1}\ln(1-\varepsilon)$. Taking e.g. $\varepsilon=1/\sqrt{N}$, this proves that $E_{\rm EoF}(|\Psi\rangle)\le E_D(|\Psi\rangle)$, where $E_D(|\Psi\rangle)$ denotes the right-hand side of (10.16).

Similarly, the first inequality in (10.19) implies that $|\Phi_{N,\varepsilon}\rangle$ can be obtained asymptotically by transforming M' copies of $|\Phi_+\rangle$ with LOCCs, more precisely it shows the existence of a LOCC operation $\mathcal{M}_{\rm LOCC}^{(N,\varepsilon)'}$ such that

$$\big\||\Psi^{\otimes N}\rangle\langle\Psi^{\otimes N}| - \mathcal{M}_{\rm LOCC}^{(N,\varepsilon)'}(|\Phi_+^{\otimes M'}\rangle\langle\Phi_+^{\otimes M'}|\big\|_1 \leq 2\sqrt{\varepsilon} \tag{10.23}$$

for N large enough, with

$$M' = \ln_2(|\mathcal{A}_{N,\varepsilon}|/(1-\varepsilon)) + 2N\varepsilon\,. \tag{10.24}$$

The production rate M'/N is bounded from above by $H + 3\varepsilon - N^{-1}\ln(1-\varepsilon)$. This establishes that $E_{\rm EoF}(|\Psi\rangle) \geq E_C(|\Psi\rangle)$, where $E_C(|\Psi\rangle)$ denotes the right-hand side of (10.17). But $E_D(|\Psi\rangle) \leq E_C(|\Psi\rangle)$, as otherwise one could transform asymptotically by a LOCC $r'N$ e-bits into rN e-bits with $r' < r$, which is impossible. Hence $E_{\rm EoF}(|\Psi\rangle = E_D(|\Psi\rangle) = E_C(|\Psi\rangle)$. □

10.4.2 Convex roof constructions

The extension of $E_{\rm EoF}$ to mixed states is done via a convex roof construction [34].

Definition 10.4.2. *The entanglement of formation of a mixed state $\rho \in \mathcal{E}(\mathcal{H}_{\mathsf{AB}})$ is*

$$E_{\rm EoF}(\rho) = \min_{\{|\Psi_i\rangle,\eta_i\}} \Big\{ \sum_i \eta_i E_{\rm EoF}(|\Psi_i\rangle) \Big\}\,, \tag{10.25}$$

where the minimum is over all pure state decompositions $\rho = \sum_i \eta_i |\Psi_i\rangle\langle\Psi_i|$ of ρ.

Proposition 10.4.3. (Vidal [248]) *$E_{\rm EoF}(\rho)$ is an entanglement measure with values in the interval $[0, \ln n]$. It satisfies the monotonicity condition (which is stronger than (iii))*

(iii") $\sum_i p_i E_{\rm EoF}(p_i^{-1}\mathcal{M}_{\rm loc}^{(i)}(\rho)) \leq E_{\rm EoF}(\rho)$ *with* $p_i = {\rm tr}[\mathcal{M}_{\rm loc}^{(i)}(\rho)]$, *for any family of CP local maps $\mathcal{M}_{\rm loc}^{(i)}$ with Kraus operators $\{A_{ij} \otimes B_{ik}\}_{j,k}$ such that $\sum_{i,j,k} A_{ij}^* A_{ij} \otimes B_{ik}^* B_{ik} = 1$.*

Note that the maps $\mathcal{M}_{\rm loc}^{(i)}$ are not required to be trace preserving (but ${\rm tr}[\mathcal{M}_{\rm loc}^{(i)}(\rho)] \leq 1$). Modulo a state normalization, they describe wavepacket reduction processes, see (5.16).

Proof. One has clearly $0 \leq E_{\rm EoF}(\rho) \leq \ln n$. We now argue that $E_{\rm EoF}$ satisfies all the axioms (i-iii) of an entanglement measure. In fact, $E_{\rm EoF}$ is convex by construction. Moreover, it follows from the aforementioned properties of $E_{\rm EoF}(|\Psi\rangle)$ and the definition of mixed state entanglement (Sec. 4.4) that $E_{\rm EoF}(\rho) = 0$ if and only if $\rho \in \mathcal{S}_{\mathsf{AB}}$. Finally, the monotonicity with respect to LOCC operations is a consequence of the convexity and can be shown as follows. Let $\rho = \sum_i \eta_i |\Psi_i\rangle\langle\Psi_i|$ be the pure state decomposition minimizing the average entanglement in the right-hand side of (10.25). Let $\mathcal{M}$ be a separable operation with Kraus operators $A_j \otimes B_j$. We denote by $\eta_{j|i} = \|A_j \otimes B_j|\Psi_i\rangle\|^2$ the probability of outcome j given that the state is $|\Psi_i\rangle$. From the convexity of $E_{\rm EoF}$ and its monotonicity (iii') for pure states (which holds by Proposition 10.3.2) one finds

$$\begin{aligned} E_{\rm EoF}\big(\mathcal{M}(\rho)\big) &\leq \sum_i \eta_i E_{\rm EoF}\big(\mathcal{M}(|\Psi_i\rangle\langle\Psi_i|)\big) \leq \sum_{ij} \eta_i \eta_{j|i} E_{\rm EoF}\big(\eta_{j|i}^{-\frac{1}{2}} A_j \otimes B_j |\Psi_i\rangle\big) \\ &\leq \sum_i \eta_i E_{\rm EoF}(|\Psi_i\rangle\langle\Psi_i|) = E_{\rm EoF}(\rho)\,. \end{aligned} \tag{10.26}$$

Thus $E_{\rm EoF}$ is an entanglement measure. A similar reasoning shows that $E_{\rm EoF}$ satisfies (iii"). □

More generally, one can construct entanglement measures by extending to mixed states any entanglement measure on pure states via a convex roof construction analog to (10.25). One gets in this way a family of measures E_f depending on the choice of the function f in Proposition 10.3.2. Conversely, any entanglement measure E satisfying the axiom (iii") above coincides with E_f on pure states for some function f fulfilling the assumptions of Proposition 10.3.2 [248]. In particular, this suggests to define the concurrence for mixed states as

$$C(\rho) = \min_{\{|\Psi_i\rangle,\eta_i\}} \Big\{ \sum_i \eta_i C(|\Psi_i\rangle) \Big\}\,, \tag{10.27}$$

where $C(|\Psi_i\rangle)$ is given by (10.4). It is known that $\rho_{\mathsf{A}} \mapsto \|\rho_{\mathsf{A}}\|_{1/2} = (\mathrm{tr}[\rho_{\mathsf{A}}^{1/2}])^2$ is concave (see (1) in Appendix B), whence $C(\rho)$ is an entanglement measure. Another measure of entanglement of common use for pure states is the Schmidt number obtained by choosing $f(\rho_{\mathsf{A}}) = 1/\,\mathrm{tr}(\rho_{\mathsf{A}}^2)$ in Proposition 10.3.2.

As stated above, (iii") means that separable measurements cannot increase the average entanglement, but entanglement can increase if one considers conditional expectations over subgroups of outcomes, i.e., one may have $E_{\mathrm{EoF}}(p_i^{-1}\mathcal{M}_{\mathrm{loc}}^{(i)}(\rho)) \geq E_{\mathrm{EoF}}(\rho)$ for some i. An example is given by the qutrit-qutrit system in the state

$$\rho = \frac{1}{2}|\Phi_+\rangle\langle\Phi_+| + \frac{1}{2}|2\rangle\langle 2| \otimes |2\rangle\langle 2| \quad , \quad |\Phi_+\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle|0\rangle + |1\rangle|1\rangle\big) \, . \tag{10.28}$$

Assume that Alice and Bob perform each a von Neumann measurement with projectors Π_1 onto $\mathrm{span}\{|0\rangle, |1\rangle\}$ and Π_2 onto $\mathbb{C}|2\rangle$. The conditional states $\rho_{\mathsf{AB}|11} = |\Phi_+\rangle\langle\Phi_+|$ and $\rho_{\mathsf{AB}|22} = |2\rangle\langle 2| \otimes |2\rangle\langle 2|$ have entanglement of formations $\ln 2$ and 0, respectively. The first value is larger than $E_{\mathrm{EoF}}(\rho)$, which is equal to $\ln 2/2$ according to the following result.

Corollary 10.4.4. *Let ρ_1 and ρ_2 be two states on $\mathcal{H}_{\mathsf{AB}}$ with bi-orthogonal supports $\mathrm{ran}\,\rho_i \subset \mathcal{V}_i^{\mathsf{A}} \otimes \mathcal{V}_i^{\mathsf{B}}$, where $\mathcal{V}_i^{\mathsf{A}} \subset \mathcal{H}_{\mathsf{A}}$ and $\mathcal{V}_i^{\mathsf{B}} \subset \mathcal{H}_{\mathsf{B}}$ are such that $\mathcal{V}_2^{\mathsf{A}} = (\mathcal{V}_1^{\mathsf{A}})^\perp$ and $\mathcal{V}_2^{\mathsf{B}} = (\mathcal{V}_1^{\mathsf{B}})^\perp$. Let $\rho = \eta_1\rho_1 + \eta_2\rho_2$ with $\eta_i \geq 0$, $\eta_1 + \eta_2 = 1$. Then $E_{\mathrm{EoF}}(\rho) = \eta_1 E_{\mathrm{EoF}}(\rho_1) + \eta_2 E_{\mathrm{EoF}}(\rho_2)$.*

Proof. The inequality $E_{\mathrm{EoF}}(\rho) \leq \eta_1 E_{\mathrm{EoF}}(\rho_1) + \eta_2 E_{\mathrm{EoF}}(\rho_2)$ follows from convexity. The reverse inequality is a consequence of the monotonicity property (iii") applied to the maps

$$\mathcal{M}_{\mathrm{loc}}^{(i)}(\rho) = \pi_i^{\mathsf{A}} \otimes \pi_i^{\mathsf{B}} \, \rho \, \pi_i^{\mathsf{A}} \otimes \pi_i^{\mathsf{B}} \, , \; i = 1, 2 \quad , \quad \mathcal{M}_{\mathrm{loc}}^{(3)}(\rho) = \pi_1^{\mathsf{A}} \otimes \pi_2^{\mathsf{B}} \, \rho \, \pi_1^{\mathsf{A}} \otimes \pi_2^{\mathsf{B}} + \pi_2^{\mathsf{A}} \otimes \pi_1^{\mathsf{B}} \, \rho \, \pi_2^{\mathsf{A}} \otimes \pi_1^{\mathsf{B}} \, , \tag{10.29}$$

where π_i^{A} and π_i^{B} are the projectors onto $\mathcal{V}_i^{\mathsf{A}}$ and $\mathcal{V}_i^{\mathsf{B}}$, respectively. □

It is worth realizing the link between $E_{\mathrm{EoF}}(\rho)$ and the classical mutual information $I_{X:Y}$, where $X = \{\eta_i\}$ is associated to a pure state decomposition $\{|\Psi_i\rangle, \eta_i\}$ of ρ and Y to the outcomes of a local measurement on A (Sec. 6.6). Indeed, the maximum of $I_{X:Y}$ over all pure state decompositions and all POVMs on A is bounded by

$$\max_{\{|\Psi_i\rangle, \eta_i\}, \{M_i^{\mathsf{A}}\}} \big\{ I_{X:Y} \big\} \leq S(\rho_{\mathsf{A}}) - E_{\mathrm{EoF}}(\rho) \, . \tag{10.30}$$

This inequality is a direct consequence of the Holevo bound (6.51) and the definition (10.25) of $E_{\mathrm{EoF}}(\rho)$.

10.4.3 The Wootters formula for two qubits

The main problem with the convex-roof construction (10.25) is that finding the pure state decomposition minimizing the average entanglement is a non-trivial task. Nevertheless, an astonishing formula enabling to evaluate $E_{\mathrm{EoF}}(\rho)$ explicitly for two qubits was found by Wootters [265]. It reads

$$E_{\mathrm{EoF}}(\rho) = h(C(\rho)) \tag{10.31}$$

where $C(\rho)$ is given by (10.27) and $h : [0, 1] \to [0, \ln n]$ is the convex increasing function

$$h(C) = -\frac{1 + \sqrt{1 - C^2}}{2} \ln\Big(\frac{1 + \sqrt{1 - C^2}}{2}\Big) - \frac{1 - \sqrt{1 - C^2}}{2} \ln\Big(\frac{1 - \sqrt{1 - C^2}}{2}\Big) \, . \tag{10.32}$$

The main point is that $C(\rho)$ can be calculated explicitly as follows. Let $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$ be the square roots of the eigenvalues of $\rho\sigma_y \otimes \sigma_y \, \overline{\rho} \, \sigma_y \otimes \sigma_y$ (here σ_y is the y-Pauli matrix and $\overline{\rho} = J\rho J$ the complex conjugate of ρ in the canonical basis). Then

$$C(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\} \, . \tag{10.33}$$

For pure states this yields $C(|\Psi\rangle) = |\langle\Psi|\sigma_y \otimes \sigma_y J|\Psi\rangle|^2$, in agreement with the result of Sec. 10.1. The proof of (10.31) is somehow tricky but relies on simple linear algebra arguments (see [265]).

10.5 Maximally entangled states

One may expect intuitively that the most entangled states are extremal states in $\mathcal{E}(\mathcal{H}_{\mathsf{AB}})$, that is, they are the pure maximally entangled states described in Sec. 4.4. If one uses as a criterion for being mostly entangled the property of having the highest entanglement of formation, this is indeed correct when the dimensions of $\mathcal{H}_{\mathsf{A}}$ and $\mathcal{H}_{\mathsf{B}}$ are such that $n_{\mathsf{A}}/2 < n_{\mathsf{B}} < 2n_{\mathsf{A}}$. When $n_{\mathsf{B}} \geq 2n_{\mathsf{A}}$, convex combinations of pure maximally entangled states with reduced B-states living on orthogonal subspaces of $\mathcal{H}_{\mathsf{B}}$ are also maximally entangled (a similar statement holds of course by exchanging A and B).

Proposition 10.5.1. *Assume that $n = n_{\mathsf{A}} \leq n_{\mathsf{B}}$ and let $r = 1, 2, \ldots$ be such that $rn_{\mathsf{A}} \leq n_{\mathsf{B}} < (r+1)n_{\mathsf{A}}$. Then the states $\rho \in \mathcal{E}(\mathcal{H}_{\mathsf{AB}})$ having a maximal entanglement of formation $E_{\mathrm{EoF}}(\rho) = \ln n$ are convex combinations of the r orthogonal maximally entangled states*

$$|k\rangle = n^{-\frac{1}{2}} \sum_{i=1}^{n} |\alpha_i^{(k)}\rangle \otimes |\beta_i^{(k)}\rangle \quad , \quad k = 1, \ldots, r \; , \tag{10.34}$$

with $\langle \alpha_i^{(k)} | \alpha_j^{(k)} \rangle = \delta_{ij}$ and $\langle \beta_i^{(k)} | \beta_j^{(l)} \rangle = \delta_{kl}\delta_{ij}$.

Proof. Let ρ be a state with $E_{\mathrm{EoF}}(\rho) = \ln n$. According to Definition 10.4.2 and given that $E_{\mathrm{EoF}}(|\Psi\rangle) \leq \ln n$ with equality if and only if $|\Psi\rangle$ is maximally entangled, this means that any pure state decomposition of ρ is made of maximally entangled states. This is the case in particular for the spectral decomposition $\rho = \sum_k p_k |k\rangle\langle k|$, from which one can obtain all other pure state decompositions $\{|\Psi_i\rangle, \eta_i\}$ by the formula $\sqrt{\eta_i}|\Psi_i\rangle = \sum_k u_{ik}\sqrt{p_k}|k\rangle$ with $\eta_i = \sum_k |u_{ik}|^2 p_k$ (see (4.16)). Let us set $D_{kl} = \mathrm{tr}_{\mathsf{B}}(|k\rangle\langle l|)$. We would like to show that $D_{kl} = n^{-1}\delta_{kl}$ if $p_k p_l \neq 0$. We already know that $D_{kk} = 1/n$ if $p_k \neq 0$, since $|k\rangle$ is maximally entangled. By plugging the above expression of $\sqrt{\eta_i}|\Psi_i\rangle$ into $\mathrm{tr}_{\mathsf{B}}(|\Psi_i\rangle\langle\Psi_i|) = 1/n$, one is led to

$$\sum_{k,l,k\neq l} \sqrt{p_k p_l}\, u_{ik}\overline{u}_{il} D_{kl} = 0 \; . \tag{10.35}$$

This equality holds for any i and any unitary matrix (u_{ik}), hence $\sqrt{p_k p_l} D_{kl} = 0$ if $k \neq l$ and the above claim is true. One deduces from $D_{kk} = 1/n$ that the eigenvectors $|k\rangle$ with eigenvalues $p_k > 0$ have Schmidt decompositions given by (10.34). For $k \neq l$, $D_{kl} = 0$ is then equivalent to $\mathcal{V}_{\mathsf{B}}^{(k)} \perp \mathcal{V}_{\mathsf{B}}^{(l)}$ with $\mathcal{V}_{\mathsf{B}}^{(k)} = \mathrm{span}\{|\beta_i^{(k)}\rangle\}_{i=1}^n \subset \mathcal{H}_{\mathsf{B}}$. If $n_{\mathsf{B}} < (r+1)n$ then at most r subspaces $\mathcal{V}_{\mathsf{B}}^{(k)}$ may be pairwise orthogonal. Thus at most r eigenvalues p_k are non-zero. □

Chapter 11

The quantum discord

Wer Kràch säiht, schaft fir deifels Schiir [Qui sème la discorde, travaille pour la grange du diable] (proverbe alsacien).

The quantum discord was introduced by Ollivier and Zurek [184] and Henderson and Vedral [120] as an indicator of the "degree of quantumness" of mixed states. For pure states it coincides with the entanglement of formation. Certain separable mixed states have, however, a non-zero discord. These states are obtained by preparing locally mixtures of non-orthogonal states, which cannot be perfectly discriminated by local measurements. Such separable states cannot be classified as "classical" and actually contain quantum correlations that are not captured by the entanglement measures reviewed in chapter 10. Apart from this observation, a motivation for the quantum discord came out in the last decade from the claim that it could play the role of a resource in certain quantum algorithms and quantum communication protocols [72, 152, 190, 163, 108, 68]. In particular, it has been suggested [72, 152, 190] that the discord might capture the quantum correlations at the origin of the quantum speedup in the deterministic quantum computation with one qubit (DQC1) of Knill and Laflamme [147]. The DQC1 algorithm computes the trace of a $2^N \times 2^N$ unitary matrix exponentially faster than all known classical algorithms. The entanglement produced during the computation with $(N+1)$ qubits is bounded independently of N, for any bipartition of the $(N+1)$ qubits [71]. This means that the total amount of bipartite entanglement is a negligible fraction of the maximal entanglement possible. However, a non-vanishing quantum discord between the control qubit and the N target qubits appears during the computation [72], save for particular unitaries [67]. The DCQ1 algorithm is singled out by the fact that it uses mixed states, the N target qubits being initially in a Gibbs state at infinite temperature. In contrast, for quantum computations using pure states, Jozsa and Linden [145] have shown that in order to offer an exponential speedup over classical computers, the computation must produce entanglement which is not restricted to qubit blocks of fixed size as the problem size increases.

The definition of the quantum discord δ_{A} is given in Sec. 11.1. We then characterize the states with vanishing discord in Sec. 11.2 and exhibit some important properties of δ_{A} in Sec. 11.3. The so-called monogamy relation linking the discord and the entanglement of formation in tripartite systems is stated and proven in Sec. 11.4.

11.1 Definition of the quantum discord

Let us first consider some classical discrete random variables A and B with joint probabilities p_{ij} and marginals $p_A(i) = \sum_j p_{ij}$ and $p_B(j) = \sum_i p_{ij}$. The correlations between A and B are measured by the mutual information $I_{A:B} = H(A) + H(B) - H(A,B)$. We recall from Sec. 6.6 that

$$I_{A:B} = H(B) - H(B|A) \;, \tag{11.1}$$

where $H(B|A) = \sum_i p_A(i) H(B|i)$ is the conditional entropy, see (6.48). This conditional entropy gives the amount of information on B left after the value $A = i$ has been measured, averaged over all possible outcomes i.

In the quantum setting, the analog of the random variables A and B is a bipartite quantum system AB in a state ρ. The marginals are the reduced states $\rho_{\mathsf{A}} = \mathrm{tr}_{\mathsf{B}}(\rho)$ and $\rho_{\mathsf{B}} = \mathrm{tr}_{\mathsf{A}}(\rho)$. The generalization of the mutual information reads

$$I_{\mathsf{A}:\mathsf{B}}(\rho) = S(\rho_{\mathsf{A}}) + S(\rho_{\mathsf{B}}) - S(\rho) \;, \tag{11.2}$$

where $S(\cdot)$ is the von Neumann entropy (7.1). Similarly to the classical case, one has $I_{\mathsf{A}:\mathsf{B}}(\rho) \geq 0$ and $I_{\mathsf{A}:\mathsf{B}}(\rho) = 0$ if and only if ρ is a product state, i.e., $\rho = \rho_{\mathsf{A}} \otimes \rho_{\mathsf{B}}$ (this is nothing but the subadditivity property of S, see

Sec. 7.1). It is easy to verify that $I_{\mathsf{A}:\mathsf{B}}(\rho)$ is related to the relative entropy (7.9) by

$$I_{\mathsf{A}:\mathsf{B}}(\rho) = S(\rho||\rho_{\mathsf{A}} \otimes \rho_{\mathsf{B}}) \; . \tag{11.3}$$

By the monotonicity of the relative entropy (Theorem 7.2.1), $I_{\mathsf{A}:\mathsf{B}}(\mathcal{M}_{\rm loc}(\rho)) \leq I_{\mathsf{A}:\mathsf{B}}(\rho)$ for any local operation $\mathcal{M}_{\rm loc} = \mathcal{M}_{\mathsf{A}} \otimes \mathcal{M}_{\mathsf{B}}$, where the operations $\mathcal{M}_{\mathsf{A}} : \mathcal{B}(\mathcal{H}_{\mathsf{A}}) \to \mathcal{B}(\mathcal{H}'_{\mathsf{A}})$ and $\mathcal{M}_{\mathsf{B}} : \mathcal{B}(\mathcal{H}_{\mathsf{B}}) \to \mathcal{B}(\mathcal{H}'_{\mathsf{B}})$ may have different initial and final spaces (for instance, $\mathcal{M}_{\mathsf{A}}$ can be the partial trace over a part of A).

However, there is no quantum analog of the identity (11.1). Let us define a conditional entropy of B given a von Neumann measurement $\{\pi_i^{\mathsf{A}}\}$ on A by $S_{\mathsf{B}|\mathsf{A}}(\rho|\{\pi_i^{\mathsf{A}}\}) = \sum_i \eta_i S(\rho_{\mathsf{B}|i})$, where

$$\rho_{\mathsf{B}|i} = \eta_i^{-1} \, {\rm tr}_{\mathsf{A}}(\pi_i^{\mathsf{A}} \otimes 1 \, \rho) \quad , \quad \eta_i = {\rm tr}(\pi_i^{\mathsf{A}} \otimes 1 \, \rho) \; . \tag{11.4}$$

Here η_i is the probability of the measurement outcome i and $\rho_{\mathsf{B}|i} = {\rm tr}_{\mathsf{A}}(\rho_{\mathsf{A}\mathsf{B}|i})$ is the corresponding conditional state of B (see chapter 5). The ensemble $\{\rho_{\mathsf{B}|i}, \eta_i\}$ defines a convex decomposition of ρ_{B} (i.e., $\rho_{\mathsf{B}} = \sum_i \eta_i \rho_{\mathsf{B}|i}$) describing a state preparation of subsystem B realized by the measurement on A. The quantum version of the right-hand side of (11.1) is the maximal reduction of entropy of B due to a von Neumann measurement on A,

$$J^{\rm v.N.}_{\mathsf{B}|\mathsf{A}}(\rho) = S(\rho_{\mathsf{B}}) - \min_{\{\pi_i^{\mathsf{A}}\}} \left\{ \sum_i \eta_i S(\rho_{\mathsf{B}|i}) \right\} , \tag{11.5}$$

the minimum being over all orthonormal families of projectors on $\mathcal{H}_{\mathsf{A}}$. This quantity represents the classical correlations between A and B (see the discussion after Proposition 11.1.2 below). Note that $J^{\rm v.N.}_{\mathsf{B}|\mathsf{A}}(\rho)$ places an upper bound on the classical mutual information between the ensemble $\{\rho_{\mathsf{B}|i}, \eta_i\}$ and the outcome probabilities when performing measurements on B to discriminate the states $\rho_{\mathsf{B}|i}$ (Sec. 6.6). Actually, $J^{\rm v.N.}_{\mathsf{B}|\mathsf{A}}(\rho)$ coincides with the corresponding Holevo quantity (6.51). By concavity of the von Neumann entropy, one has $J^{\rm v.N.}_{\mathsf{B}|\mathsf{A}}(\rho) \geq 0$. Furthermore, (7.8) entails $J^{\rm v.N.}_{\mathsf{B}|\mathsf{A}}(\rho) \leq \max_{\{\pi_i^{\mathsf{A}}\}} H(\{\eta_i\})$.

It also follows from the concavity of S that the minimum in (11.5) is achieved for rank-one projectors. In fact, by decomposing each projector π_i^{A} of rank r_i as a sum of r_i rank-one projectors π_{ik}^{A}, one finds that $\rho_{\mathsf{B}|i} = \sum_k (\eta_{ik}/\eta_i) \rho_{\mathsf{B}|ik}$ is a convex combination of the states $\rho_{\mathsf{B}|ik} = \eta_{ik}^{-1} \, {\rm tr}_{\mathsf{A}}(\pi_{ik}^{\mathsf{A}} \otimes 1 \, \rho)$ if $\eta_i = \sum_k \eta_{ik} > 0$. Thereby $\sum_i \eta_i S(\rho_{\mathsf{B}|i}) \geq \sum_{ik} \eta_{ik} S(\rho_{\mathsf{B}|ik})$.

Ollivier and Zurek [184] and Henderson and Vedral [120] proposed in two independent works published in 2001 to characterize the amount of non-classicality in the state ρ by forming the difference between the total correlations given by $I_{\mathsf{A}:\mathsf{B}}(\rho)$ and the classical correlations given by $J^{\rm v.N.}_{\mathsf{B}|\mathsf{A}}(\rho)$.

Definition 11.1.1. *The quantum discord of the bipartite system* AB *in state* ρ *is*

$$\delta^{\rm v.N.}_{\mathsf{A}}(\rho) = I_{\mathsf{A}:\mathsf{B}}(\rho) - J^{\rm v.N.}_{\mathsf{B}|\mathsf{A}}(\rho) = S(\rho_{\mathsf{A}}) - S(\rho) + \min_{\{\pi_i^{\mathsf{A}}\}} \left\{ \sum_i \eta_i S(\rho_{\mathsf{B}|i}) \right\} . \tag{11.6}$$

In [120], the minimization is done over generalized measurements given by POVMs $\{M_i^{\mathsf{A}}\}$ on $\mathcal{H}_{\mathsf{A}}$, instead of von Neumann measurements. The conditional states and outcome probabilities are then (chapter 5)

$$\rho_{\mathsf{B}|i} = \eta_i^{-1} \, {\rm tr}_{\mathsf{A}}(M_i^{\mathsf{A}} \otimes 1 \, \rho) \quad , \quad \eta_i = {\rm tr}(M_i^{\mathsf{A}} \otimes 1 \, \rho) \; . \tag{11.7}$$

We denote the corresponding discord by $\delta_{\mathsf{A}}(\rho)$. As in the case of von Neumann measurements, the minimum is achieved for rank-one measurement operators M_i^{A}. In general, the inequality $\delta_{\mathsf{A}}(\rho) < \delta^{\rm v.N.}_{\mathsf{A}}(\rho)$ is strict[1]. Nevertheless, by the Neumark extension theorem, δ_{A} coincides with $\delta^{\rm v.N.}_{\mathsf{A}}$ up to an enlargement of the space $\mathcal{H}_{\mathsf{A}}$. More precisely, by plugging $M_i^{\mathsf{A}} = \langle \epsilon_0 | \Pi^{\mathsf{A}\mathsf{E}} | \epsilon_0 \rangle$ (see Remark 5.3.3) into (11.7) and using the additivity of S under tensor products, a simple calculation gives

$$\delta_{\mathsf{A}}(\rho) = \delta^{\rm v.N.}_{\mathsf{A}\mathsf{E}}(\rho \otimes |\epsilon_0\rangle\langle\epsilon_0|) \; , \tag{11.8}$$

the right-hand side being independent of the ancilla state $|\epsilon_0\rangle \in \mathcal{H}_{\mathsf{E}}$.

The discords $\delta^{\rm v.N.}_{\mathsf{A}}(\rho)$ and $\delta_{\mathsf{A}}(\rho)$ thus measure the amount of total correlations between A and B which cannot be accessed by local measurements on the subsystem A. Note that they are asymmetric under the exchange $\mathsf{A} \leftrightarrow \mathsf{B}$. One can define similarly the discords $\delta^{\rm v.N.}_{\mathsf{B}}(\rho)$ and $\delta_{\mathsf{B}}(\rho)$ by performing the measurements on the subsystem B.

[1] See e.g. [111, 98] for a comparison of the von Neumann and POVM discords for two qubits.

For pure states $\rho_\Psi = |\Psi\rangle\langle\Psi|$, the mutual information $I_{\mathsf{A}:\mathsf{B}}(\rho_\Psi)$ is equal to $2S([\rho_\Psi]_\mathsf{B})$, see (7.4), and the measurement minimizing the conditional entropy of B is the measurement in the eigenbasis $\{|\alpha_i\rangle\}$ of the reduced state $[\rho_\Psi]_\mathsf{A}$. In fact, according to (4.9) the corresponding post-measurement states $\rho_{\mathsf{B}|i} = |\beta_i\rangle\langle\beta_i|$ are pure and thus have zero entropy. Then (11.5) yields $J_{\mathsf{B}|\mathsf{A}}(\rho_\Psi) = S([\rho_\Psi]_\mathsf{B})$. As a result, the discords coincide for pure states with the entanglement of formation,

$$\delta_\mathsf{A}(|\Psi\rangle) = \delta_\mathsf{A}^{\text{v.N.}}(|\Psi\rangle) = \delta_\mathsf{B}(|\Psi\rangle) = \delta_\mathsf{B}^{\text{v.N.}}(|\Psi\rangle) = E_{\text{EoF}}(|\Psi\rangle)\ . \tag{11.9}$$

For mixed states, it was pointed out in [184] that if the measurement operators M_i^A are of rank one then

$$\sum_i \eta_i S(\rho_{\mathsf{B}|i}) = S\big(\mathcal{M}_\mathsf{A}\otimes 1(\rho)\big) - S\big([\mathcal{M}_\mathsf{A}\otimes 1(\rho)]_\mathsf{A}\big) = -I_{\mathsf{A}:\mathsf{B}}\big(\mathcal{M}_\mathsf{A}\otimes 1(\rho)\big) + S(\rho_\mathsf{B})\ , \tag{11.10}$$

where $\mathcal{M}_\mathsf{A}$ is the quantum operation on A associated to the measurement. Actually, consider the family of Kraus operators for $\mathcal{M}_\mathsf{A}$ given by $\{A_i = |i\rangle\langle\widetilde{\mu}_i|\}$, where $|\widetilde{\mu}_i\rangle$ are unnormalized vectors such that $M_i^\mathsf{A} = |\widetilde{\mu}_i\rangle\langle\widetilde{\mu}_i|$ and $\{|i\rangle\}$ is an orthonormal basis of a pointer space $\mathcal{H}_\mathsf{P}$. Then $\mathcal{M}_\mathsf{A}\otimes 1(\rho) = \sum_i \eta_i |i\rangle\langle i|\otimes\rho_{\mathsf{B}|i}$ and the reduced state $[\mathcal{M}_\mathsf{A}\otimes 1(\rho)]_\mathsf{A} = \sum_i \eta_i |i\rangle\langle i|$ has entropy $-\sum_i \eta_i \ln\eta_i$. A simple calculation yields the first equality in (11.10). The second equality is clear once one notices that $[\mathcal{M}_\mathsf{A}\otimes 1(\rho)]_\mathsf{B} = \rho_\mathsf{B}$.

Therefore, by combining (11.5), (11.6), and (11.3) one obtains the following result.

Proposition 11.1.2. [162] *The discord $\delta(\rho) = I_{\mathsf{A}:\mathsf{B}}(\rho) - J_{\mathsf{B}|\mathsf{A}}(\rho)$ is the minimal difference of mutual information of AB before and after a measurement on A, i.e.*

$$J_{\mathsf{B}|\mathsf{A}}(\rho) = \max_{\{M_i^\mathsf{A}\}} \big\{ I_{\mathsf{A}:\mathsf{B}}(\mathcal{M}_\mathsf{A}\otimes 1(\rho)) \big\}\ , \tag{11.11}$$

where the maximum is over all POVMs on A with rank-one operators M_i^A and $\mathcal{M}_\mathsf{A}$ is the associated quantum operation on $\mathcal{B}(\mathcal{H}_\mathsf{A})$. As a result,

$$\delta_\mathsf{A}(\rho) = \min_{\{M_i^\mathsf{A}\}} \Big\{ S\big(\rho||\rho_\mathsf{A}\otimes\rho_\mathsf{B}\big) - S\big(\mathcal{M}_\mathsf{A}\otimes 1(\rho)||\mathcal{M}_\mathsf{A}(\rho_\mathsf{A})\otimes\rho_\mathsf{B}\big)\Big\}\ . \tag{11.12}$$

Similarly, $J_{\mathsf{B}|\mathsf{A}}^{\text{v.N.}}(\rho)$ is given by maximizing $I_{\mathsf{A}:\mathsf{B}}(\mathcal{M}_{\pi^\mathsf{A}}\otimes 1(\rho))$ over all von Neumann measurements $\mathcal{M}_{\pi^\mathsf{A}}$ on A of the form (5.5) with rank-one projectors π_i^A.

Observing that a measurement on A with no readout removes the quantum correlations between A and B, the right-hand side of (11.11) can be interpreted as the amount of classical correlations between the two subsystems. These subsystems are not correlated classically, i.e., $J_{\mathsf{B}|\mathsf{A}}(\rho) = 0$, if and only if $\rho = \rho_\mathsf{A}\otimes\rho_\mathsf{B}$ is a product state. This result holds for $J_{\mathsf{B}|\mathsf{A}}^{\text{v.N.}}(\rho)$ as well. Actually, by (11.11), $J_{\mathsf{B}|\mathsf{A}}(\rho) = 0$ is equivalent to $\mathcal{M}_\mathsf{A}\otimes 1(\rho)$ being a product state for any collection of operators $M_i^\mathsf{A} = |\widetilde{\mu}_i\rangle\langle\widetilde{\mu}_i|$ forming a POVM. This implies $\eta_i\rho_{\mathsf{B}|i} = \langle\widetilde{\mu}_i|\rho|\widetilde{\mu}_i\rangle = \eta_i\rho_\mathsf{B}$ for all i (see the discussion before Proposition 11.1.2). Choosing the $|\widetilde{\mu}_i\rangle$ to be the eigenvectors of the observable A, one obtains that $\langle A\otimes B\rangle_\rho = \langle A\otimes 1\rangle_\rho\langle 1\otimes B\rangle_\rho$ for any $A\in\mathcal{B}(\mathcal{H}_\mathsf{A})_{\text{s.a.}}$ and $B\in\mathcal{B}(\mathcal{H}_\mathsf{B})_{\text{s.a.}}$, with $\langle\cdot\rangle_\rho = \text{tr}(\cdot\rho)$.

Let us emphasize that finding the optimal measurement which maximizes the post-measurement mutual information, and hence calculating the discords $\delta_\mathsf{A}^{\text{v.N.}}(\rho)$ and $\delta_\mathsf{A}(\rho)$, is a difficult problem in general. Even for two qubits, this problem has been solved so far for a restricted family of states only, namely, the states ρ with maximally mixed marginals $\rho_\mathsf{A} = \rho_\mathsf{B} = 1/2$ [161]. In other cases[2] the discords must be evaluated numerically (however, $\delta_\mathsf{A}(\rho)$ can be determined analytically for low-rank density matrices with the help of the monogamy relation, see Sec. 11.4 and [170]).

11.2 The A-classical states

The monotonicity property of the relative entropy and formula (11.12) imply that $\delta_\mathsf{A}(\rho)$ is non-negative. The states with vanishing discord can be determined with the help of Theorem 7.2.1, leading to the following result[3].

[2] An incorrect work [3] claiming to extend the result of Ref. [161] to the larger family of the so-called X-states has generated a profusion of articles. Comparing with numerical evaluations, the result of [3] apparently gives good approximations of the discord for randomly chosen X-states (see the discussion in [170]).

[3] In Ref. [184], the authors argue that the non-negativity of $\delta_\mathsf{A}^{\text{v.N.}}(\rho)$ is a direct consequence of (11.10) and the concavity of $S(\rho) - S(\rho_\mathsf{A})$ with respect to ρ. I do not see how such a claim could be justified and believe that the simplest proof of Proposition 11.2.1 is to rely on Theorem 7.2.1. Alternatively, the non-negativity of the discord can be justified with the help of the strong subadditivity of the von Neumann entropy (which is closely related to Theorem 7.2.1, see Sec. 7.2), as shown in Ref. [164].

Proposition 11.2.1. *The quantum discord is non-negative and $\delta_{\mathsf{A}}(\sigma) = 0$ if and only if*

$$\sigma = \sum_{i=1}^{n_{\mathsf{A}}} q_i |\varphi_i\rangle\langle\varphi_i| \otimes \sigma_{\mathsf{B}|i} \ , \tag{11.13}$$

where $\{|\varphi_i\rangle\}_{i=1}^{n_{\mathsf{A}}}$ is an orthonormal basis of $\mathcal{H}_{\mathsf{A}}$, $\sigma_{\mathsf{B}|i}$ are some (arbitrary) states of B depending on the index i, and $q_i \geq 0$ are some probabilities, $\sum_i q_i = 1$.

The non-negativity of $\delta_{\mathsf{A}}(\rho)$ means that one cannot gain more information on a bipartite system AB by performing a measurement on the subsystem A than the entropy of A, namely, $S(\rho_{\mathsf{AB}}) - \sum \eta_i S(\rho_{\mathsf{AB}|i}) \leq S(\rho_{\mathsf{A}})$ for any $\rho_{\mathsf{AB}} \in \mathcal{E}(\mathcal{H}_{\mathsf{AB}})$. The important point is that if ρ_{AB} is not of the form (11.13), then any measurement on A gives *less* information on AB than $S(\rho_{\mathsf{A}})$. Stated differently, one can not retrieve all the information on A by a local measurement, because of the presence of quantum correlations between A and B.

Proof. It remains to show the second affirmation. It is easy to convince oneself that the states (11.13) have a vanishing discord. In fact, one finds $I_{\mathsf{A}:\mathsf{B}}(\sigma) = S(\sigma_{\mathsf{B}}) - \sum_i q_i S(\sigma_{\mathsf{B}|i}) \leq J^{\mathrm{v.N.}}_{\mathsf{B}|\mathsf{A}}(\sigma)$ (the inequality follows by noting that $\sigma_{\mathsf{B}|i}$ and q_i are the conditional state and outcome probability for a measurement on A in the basis $\{|\varphi_i\rangle\}$). Hence $\delta_{\mathsf{A}}(\sigma) = \delta^{\mathrm{v.N.}}_{\mathsf{A}}(\sigma) = 0$ as a consequence of the non-negativity of δ_{A}. Reciprocally, let $\sigma \in \mathcal{E}(\mathcal{H}_{\mathsf{AB}})$ be such that $\delta^{\mathrm{v.N.}}_{\mathsf{A}}(\sigma) = 0$. As we shall see below it is enough to work with the von Neumann discord, the result for δ_{A} will then follow from (11.8). According to (11.12) and Theorem 7.2.1, $\delta^{\mathrm{v.N.}}_{\mathsf{A}}(\sigma) = 0$ if and only if there exists a von Neumann measurement $\mathcal{M}_{\mathsf{A}}$ on A with rank-one projectors $\pi^{\mathsf{A}}_i = |\varphi_i\rangle\langle\varphi_i|$ such that $\sigma = \mathcal{R}_{\mathsf{A}}\mathcal{M}_{\mathsf{A}} \otimes 1(\sigma)$, where $\mathcal{R}_{\mathsf{A}} = \mathcal{R}_{\mathcal{M}_{\mathsf{A}}\otimes 1, \sigma_0}$ is the transpose operation of $\mathcal{M}_{\mathsf{A}} \otimes 1$ for the state $\sigma_0 = \sigma_{\mathsf{A}} \otimes \sigma_{\mathsf{B}}$. Without loss of generality we may assume $\eta_i = \langle\varphi_i|\sigma_{\mathsf{A}}|\varphi_i\rangle > 0$ for all i. Thanks to (7.10) and to the identity $\mathcal{M}_{\mathsf{A}} \otimes 1(\sigma_0) = \sum_i \eta_i |\varphi_i\rangle\langle\varphi_i| \otimes \sigma_{\mathsf{B}}$, the transpose operation $\mathcal{R}_{\mathsf{A}}$ has Kraus operators $R_i = \eta_i^{-1/2}\sqrt{\sigma_{\mathsf{A}}}|\varphi_i\rangle\langle\varphi_i| \otimes 1$. We now argue that this implies that $\sigma = \widehat{\mathcal{M}}_{\mathsf{A}} \otimes 1(\sigma)$ with $\widehat{\mathcal{M}}_{\mathsf{A}}$ the von Neumann measurement with projectors $\widehat{\pi}^{\mathsf{A}}_k$ onto the subspaces $\mathrm{span}\{|\varphi_i\rangle; i \in I_k\}$, where $\{I_1, \ldots, I_d\}$ is a partition of $\{1, \ldots, n_{\mathsf{A}}\}$. Actually, the condition $\sigma = \mathcal{R}_{\mathsf{A}}\mathcal{M}_{\mathsf{A}} \otimes 1(\sigma)$ reads

$$\langle\varphi_i|\sigma|\varphi_j\rangle = \sum_{l=1}^{n_{\mathsf{A}}} \eta_l^{-1} (\sqrt{\sigma_{\mathsf{A}}})_{il} (\sqrt{\sigma_{\mathsf{A}}})_{lj} \langle\varphi_l|\sigma|\varphi_l\rangle \quad , \quad i,j = 1, \ldots, n_{\mathsf{A}} \tag{11.14}$$

with $(\sqrt{\sigma_{\mathsf{A}}})_{ij} = \langle\varphi_i|\sqrt{\sigma_{\mathsf{A}}}|\varphi_j\rangle \in \mathbb{R}$. Let us set $\sigma_{\mathsf{B}|i} = \eta_i^{-1}\langle\varphi_i|\sigma|\varphi_i\rangle$ and $\eta_{l|i} = |(\sqrt{\sigma_{\mathsf{A}}})_{il}|^2/\eta_i$. This defines respectively a state on $\mathcal{H}_{\mathsf{B}}$ and a probability distribution for any fixed i. With this notation, (11.14) can be rewritten for $i = j$ as

$$\sigma_{\mathsf{B}|i} = \sum_{l=1}^{n_{\mathsf{A}}} \eta_{l|i} \sigma_{\mathsf{B}|l} \quad , \quad i = 1, \ldots, n_{\mathsf{A}} \ . \tag{11.15}$$

Let $I_i = \{j; \sigma_{\mathsf{B}|j} = \sigma_{\mathsf{B}|i}\} \subset \{1, \ldots, n_{\mathsf{A}}\}$. Clearly, the sets I_i are either equal or disjoint. Hence one can extract from them a partition $\{I_{i_1}, I_{i_2}, \ldots, I_{i_d}\}$ of $\{1, \ldots, n_{\mathsf{A}}\}$. We claim that (11.15) implies $\eta_{l|i} = 0$ for $l \notin I_i$. This is a consequence of the following lemma.

Lemma 11.2.2. *Let $\mathbf{x} = (x_1, \ldots, x_d)$ be a vector of $\mathcal{X}^d$ with distinct components x_k, where $\mathcal{X}$ is a real vector space, and $\{\xi_{k|m}\}_{k=1}^d$ be some probability distributions such that $\xi_{k|m} = 0 \Leftrightarrow \xi_{m|k} = 0$ and the components of $\mathbf{x}$ have convex decompositions*

$$x_m = \sum_{k=1}^{d} \xi_{k|m} x_k \quad \forall \ m = 1, \ldots, d \ . \tag{11.16}$$

Then $\xi_{k|m} = \delta_{km}$ for any $k, m = 1, \ldots, d$.

We postpone the proof of this result to the next paragraph. By rewriting (11.15) as

$$\sigma_{\mathsf{B}|i_m} = \sum_{k=1}^{d} \xi_{k|m} \sigma_{\mathsf{B}|i_k} \quad \text{with} \quad \xi_{k|m} = |I_{i_m}|^{-1} \sum_{(l,i) \in I_{i_k} \times I_{i_m}} \eta_{l|i} \ , \tag{11.17}$$

one concludes from Lemma 11.2.2 that $\xi_{k|m} = 0$ for $k \neq m$, i.e., $\eta_{l|i} = (\sqrt{\sigma_{\mathsf{A}}})_{il} = 0$ for any (i,l) such that $l \notin I_i$. One then obtains from (11.14)

$$\sigma = \sum_{i,j=1}^{n_{\mathsf{A}}} \sum_{l \in I_i \cap I_j} (\sqrt{\sigma_{\mathsf{A}}})_{il} (\sqrt{\sigma_{\mathsf{A}}})_{lj} |\varphi_i\rangle\langle\varphi_j| \otimes \sigma_{\mathsf{B}|l} = \sum_{k=1}^{d} \sum_{i,j \in I_{i_k}} (\sigma_{\mathsf{A}})_{ij} |\varphi_i\rangle\langle\varphi_j| \otimes \sigma_{\mathsf{B}|i_k} \ . \tag{11.18}$$

This gives

$$\sigma = \sum_{k=1}^{d} \widehat{\pi}_k^{\mathsf{A}} \sigma_{\mathsf{A}} \widehat{\pi}_k^{\mathsf{A}} \otimes \sigma_{\mathsf{B}|i_k} \quad , \quad \widehat{\pi}_k^{\mathsf{A}} = \sum_{i \in I_{i_k}} |\varphi_i\rangle\langle\varphi_i| \, . \tag{11.19}$$

The last expression is of the form (11.13) (note that the vectors $|\varphi_i\rangle$ in the latter formula are the eigenvectors of $\widehat{\pi}_k^{\mathsf{A}} \sigma_{\mathsf{A}} \widehat{\pi}_k^{\mathsf{A}}$, so that they are in general linear combinations of the vectors $|\varphi_i\rangle$ defined above). To get the result for the discord δ_{A} we take advantage of (11.8). From the foregoing result, $\delta_{\mathsf{A}}(\sigma) = 0$ is equivalent to $\sigma \otimes |\epsilon_0\rangle\langle\epsilon_0|$ being of the form (11.13) for some orthonormal basis $\{|\varphi_i^{\mathsf{AE}}\rangle\}$ of $\mathcal{H}_{\mathsf{AE}}$. This straightforwardly implies $|\varphi_i^{\mathsf{AE}}\rangle = |\varphi_i\rangle|\epsilon_0\rangle$ with $\{|\varphi_i\rangle\}$ an orthonormal basis of $\mathcal{H}_{\mathsf{A}}$. □

Proof of Lemma 11.2.2. One proceeds by induction on d. The result is trivial for $d = 2$. Let us assume that it holds true for $d \geq 2$ and that one can find a vector $\mathbf{x} \in \mathcal{X}^{d+1}$ and some probabilities $\{\xi_{k|m}\}_{k=1}^{d+1}$ like in the lemma such that $\xi_{k_0|k_0} < 1$ for some $k_0 \in \{1, \ldots, d+1\}$. We are going to show that this leads to a contradiction. By plugging $x_{k_0} = (1 - \xi_{k_0|k_0})^{-1} \sum_{k \neq k_0} \xi_{k|k_0} x_k$ into the p other convex decompositions, one gets $x_m = \sum_{k \neq k_0} \zeta_{k|m} x_k$ for $k \neq k_0$, with $\zeta_{k|m} = \xi_{k|m} + (1 - \xi_{k_0|k_0})^{-1} \xi_{k_0|m} \xi_{k|k_0}$. As $\{\zeta_{k|m}\}_{k \neq k_0}$ is a probability distribution satisfying $\zeta_{k|m} = 0 \Leftrightarrow \zeta_{m|k} = 0$, by the induction hypothesis one has $\zeta_{k|m} = \delta_{km}$ for any $k, m \in \{1, \ldots, d+1\} \setminus \{k_0\}$. Now $\xi_{m_0|k_0} > 0$ for some index $m_0 \neq k_0$ (because $\xi_{k_0|k_0} < 1$). One deduces from the above identities and the hypothesis on $\xi_{k|m}$ that the only non-vanishing probabilities are $\xi_{k_0|m_0}$, $\xi_{m_0|k_0}$, and $\xi_{k|k}$, $k = 1, \ldots, p+1$. The problem then reduces to the case $p = 2$. Thus $\xi_{k_0|k_0} = \xi_{m_0|m_0} = 1$, in contradiction with our assumption. □

Definition 11.2.3. *The zero-discord states of the form (11.13) are called the* A*-classical states. We denote by* $\mathcal{C}_{\mathsf{A}}$ *the set of all* A*-classical states. Similarly,* $\mathcal{C}_{\mathsf{B}}$ *is the set of all* B*-classical states, namely, the states with vanishing* B*-discord. A classical state is a state which is both* A*- and* B*-classical. We write* $\mathcal{C}_{\mathsf{AB}} = \mathcal{C}_{\mathsf{A}} \cap \mathcal{C}_{\mathsf{B}}$.

Our terminology can be justified by noting that if AB is in a state of the form (11.13) then the subsystem A is in one of the orthogonal states $|\varphi_i\rangle$ with probability q_i, whence A behaves as a classical system being in state i with probability q_i. Alternatively, a state σ is A-classical if and only if there exists a von Neumann measurement on A with rank-one projectors $\pi_i^{\mathsf{A}} = |\varphi_i\rangle\langle\varphi_i|$ which does not perturb it in the absence of readout, i.e., $\sigma = \mathcal{M}_{\{\pi_i^{\mathsf{A}}\}} \otimes 1(\sigma)$. The unfortunate name "classical-quantum states" has become popular in the literature to refer to the A-classical states, the B-classical states being called "quantum-classical". Using the spectral decompositions of the $\sigma_{\mathsf{B}|i}$'s, any A-classical state $\sigma_{\mathsf{A-cl}} \in \mathcal{C}_{\mathsf{A}}$ can be decomposed as

$$\sigma_{\mathsf{A-cl}} = \sum_{i=1}^{n_{\mathsf{A}}} \sum_{j=1}^{n_{\mathsf{B}}} q_{ij} |\varphi_i\rangle\langle\varphi_i| \otimes |\chi_{j|i}\rangle\langle\chi_{j|i}| \, , \tag{11.20}$$

where $q_{ij} \geq 0$, $\sum_{i,j} q_{ij} = 1$ and, for any i, $\{|\chi_{j|i}\rangle\}_{j=1}^{n_{\mathsf{B}}}$ is an orthonormal basis of $\mathcal{H}_{\mathsf{B}}$ (note that the $|\chi_{j|i}\rangle$ need not be orthogonal for distinct i's). A classical state $\sigma_{\mathrm{clas}} \in \mathcal{C}_{\mathsf{A}} \cap \mathcal{C}_{\mathsf{B}}$ possesses an eigenbasis $\{|\varphi_i\rangle \otimes |\chi_j\rangle\}_{i=1,j=1}^{n_{\mathsf{A}},n_{\mathsf{B}}}$ of product vectors. It is fully classical, in the sense that any quantum system in this state can be "simulated" by a classical apparatus being in the state (i, j) with probability q_{ij}.

Let us point out that $\mathcal{C}_{\mathsf{A}}$, $\mathcal{C}_{\mathsf{B}}$, and $\mathcal{C}_{\mathsf{AB}}$ are not convex. Their convex hull is the set $\mathcal{S}_{\mathsf{AB}}$ of separable states. It is also important to realize that for pure states, A-classical, B-classical, classical, and separable states all coincide. Actually, according to (11.20) the pure A-classical (and, similarly, the pure B-classical) states are product states. In contrast, one can find mixed separable states which are not A-classical. An example for two qubits is

$$\rho = \frac{1}{4}\big(|+\rangle\langle+| \otimes |0\rangle\langle 0| + |-\rangle\langle-| \otimes |1\rangle\langle 1| + |0\rangle\langle 0| \otimes |-\rangle\langle-| + |1\rangle\langle 1| \otimes |+\rangle\langle+|\big) \tag{11.21}$$

with $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. It is clear that $\rho \in \mathcal{S}_{\mathsf{AB}}$, but ρ is neither A-classical nor B-classical. A schematic picture of the sets $\mathcal{S}_{\mathsf{AB}}$, $\mathcal{C}_{\mathsf{A}}$, $\mathcal{C}_{\mathsf{B}}$, and $\mathcal{C}_{\mathsf{AB}}$ for a general bipartite system AB is displayed in Fig. 11.1.

11.3 Properties of the quantum discord

11.3.1 Invariance and monotonicity properties

Unlike entanglement measures, the quantum discord is *not* monotonous with respect to LOCCs. In particular, local operations on the measured subsystem A can create discord. For instance, consider the classical state

$$\sigma = \frac{1}{2}\Big(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1|\Big) \tag{11.22}$$

of two qubits. One can transform this state by a local operation $\mathcal{M}_{\mathsf{A}}$ on A into

$$\rho = \mathcal{M}_{\mathsf{A}} \otimes 1(\sigma) = \frac{1}{2}\Big(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |+\rangle\langle +| \otimes |1\rangle\langle 1|\Big) \,, \tag{11.23}$$

where $\mathcal{M}_{\mathsf{A}}$ has Kraus operators $A_0 = |0\rangle\langle 0|$ and $A_1 = |+\rangle\langle 1|$. The final state ρ has less total correlations than σ, its mutual information $I_{\mathsf{A}:\mathsf{B}}(\rho) = -p\ln p - (1-p)\ln(1-p)$ being smaller than $I_{\mathsf{A}:\mathsf{B}}(\sigma) = \ln 2$ (here $p = 1/2 + \sqrt{2}/4$). However, it has a positive discord $\delta_{\mathsf{A}}(\rho) > \delta_{\mathsf{A}}(\sigma) = 0$. This means that the loss of classical correlations $J_{\mathsf{B}|\mathsf{A}}(\sigma) - J_{\mathsf{B}|\mathsf{A}}(\rho)$ is larger than the loss of total correlations $I_{\mathsf{A}:\mathsf{B}}(\sigma) - I_{\mathsf{A}:\mathsf{B}}(\rho)$.

In contrast, as far as local operations on B are concerned everything goes as expected, as shown by the following result.

Proposition 11.3.1. *The quantum discord δ_{A} and classical correlations $J_{\mathsf{B}|\mathsf{A}}(\rho)$ are invariant with respect to unitary conjugations $\mathcal{U}_{\mathsf{A}} : \rho_{\mathsf{A}} \mapsto U_{\mathsf{A}}\rho_{\mathsf{A}}U_{\mathsf{A}}^*$ on A and monotonous with respect to quantum operations $\mathcal{M}_{\mathsf{B}}$ on B, namely,*

$$\begin{array}{ccccccc} \delta_{\mathsf{A}}(\mathcal{U}_{\mathsf{A}} \otimes 1(\rho)) & = & \delta_{\mathsf{A}}(\rho) & , & \delta_{\mathsf{A}}(1 \otimes \mathcal{M}_{\mathsf{B}}(\rho)) & \leq & \delta_{\mathsf{A}}(\rho) \\ J_{\mathsf{B}|\mathsf{A}}(\mathcal{U}_{\mathsf{A}} \otimes 1(\rho)) & = & J_{\mathsf{B}|\mathsf{A}}(\rho) & , & J_{\mathsf{B}|\mathsf{A}}(1 \otimes \mathcal{M}_{\mathsf{B}}(\rho)) & \leq & J_{\mathsf{B}|\mathsf{A}}(\rho) \end{array} \tag{11.24}$$

and similarly for $\delta_{\mathsf{A}}^{\text{v.N.}}$ and $J_{\mathsf{B}|\mathsf{A}}^{\text{v.N.}}$.

Proof. The unitary invariance is trivial. The monotonicity of $J_{\mathsf{B}|\mathsf{A}}(\rho)$ with respect to operations on B comes from the monotonicity of the relative entropy and the formula

$$J_{\mathsf{B}|\mathsf{A}}(\rho) = \max_{\{M_i^{\mathsf{A}}\}} \Big\{ \sum_i \eta_i S(\rho_{\mathsf{B}|i} || \rho_{\mathsf{B}}) \Big\} \,, \tag{11.25}$$

which is a consequence of the definition (11.5) and of $\rho_{\mathsf{B}} = \sum_i \eta_i \rho_{\mathsf{B}|i}$. A simple justification of the monotonicity of δ_{A} with respect to operations on B uses the following reasoning [198]. Let us consider a generalized measurement $\{M_i^{\mathsf{A}}\}$ on A with associated quantum operation $\mathcal{M}_{\mathsf{A}}$. By invoking the Stinespring theorem, one can represent $\mathcal{M}_{\mathsf{A}}$ as $\mathcal{M}_{\mathsf{A}} \otimes 1(\rho) = \text{tr}_{\mathsf{E}}(\sigma_{\mathsf{ABE}})$ with $\sigma_{\mathsf{ABE}} = U_{\mathsf{AE}}\rho \otimes |\epsilon_0\rangle\langle\epsilon_0| U_{\mathsf{AE}}^*$ pertaining to an enlarged space $\mathcal{H}_{\mathsf{ABE}}$ and U_{AE} a unitary on $\mathcal{H}_{\mathsf{AE}}$. Thanks to the additivity and unitary invariance of the von Neumann entropy and to the relation $\text{tr}_{\mathsf{AE}}(\sigma_{\mathsf{ABE}}) = \rho_{\mathsf{B}}$, one finds

$$I_{\mathsf{A}:\mathsf{B}}(\rho) = I_{\mathsf{AE}:\mathsf{B}}(\sigma_{\mathsf{ABE}}) \quad , \quad I_{\mathsf{A}:\mathsf{B}}(\mathcal{M}_{\mathsf{A}} \otimes 1(\rho)) = I_{\mathsf{A}:\mathsf{B}}(\sigma_{\mathsf{AB}}) \,. \tag{11.26}$$

Plugging these expressions into (11.12) gives the following expression of $\delta_{\mathsf{A}}(\rho)$ in terms of the conditional mutual informations

$$\delta_{\mathsf{A}}(\rho) = \min_{\{M_i^{\mathsf{A}}\}} \big\{ I_{\mathsf{AE}:\mathsf{B}}(\sigma_{\mathsf{ABE}}) - I_{\mathsf{A}:\mathsf{B}}(\sigma_{\mathsf{AB}}) \big\} = \min_{\{M_i^{\mathsf{A}}\}} \big\{ I_{\mathsf{AB}:\mathsf{E}}(\sigma_{\mathsf{ABE}}) - I_{\mathsf{A}:\mathsf{E}}(\sigma_{\mathsf{AE}}) \big\} \,. \tag{11.27}$$

The monotonicity of δ_{A} then follows from the monotonicity of the mutual information with respect to local operations (Sec. 11.1). □

11.3.2 States with the highest discord

As stated at the beginning of this chapter, the quantum discord $\delta_{\mathsf{A}}(\rho)$ is an indicator of the degree of quantumness of ρ. It is thus natural to ask whether the " most quantum" states having the highest discord are the maximally entangled states characterized in Proposition 10.5.1. The answer is affirmative when $n_{\mathsf{A}} \leq n_{\mathsf{B}}$.

Proposition 11.3.2. *For any state ρ of the bipartite system AB, one has*

$$\delta_{\mathsf{A}}(\rho) \leq \delta_{\mathsf{A}}^{\text{v.N.}}(\rho) \leq S(\rho_{\mathsf{A}}) \leq \ln n_{\mathsf{A}} \,. \tag{11.28}$$

If $n_{\mathsf{A}} \leq n_{\mathsf{B}}$ then the maximal value of $\delta_{\mathsf{A}}(\rho)$ over all states $\rho \in \mathcal{E}(\mathcal{H}_{\mathsf{AB}})$ is equal to $\ln n_{\mathsf{A}}$ and $\delta_{\mathsf{A}}(\rho) = \ln n_{\mathsf{A}}$ if and only if ρ has highest entanglement of formation. Thus, the states ρ_{ent} with highest discord are the maximally entangled states given by Proposition 10.5.1, which satisfy

$$\delta_{\mathsf{A}}(\rho_{\text{ent}}) = \delta_{\mathsf{A}}^{\text{v.N.}}(\rho_{\text{ent}}) = E_{\text{EoF}}(\rho_{\text{ent}}) = \ln n_{\mathsf{A}} \,. \tag{11.29}$$

The statements in this proposition are probably well known in the literature, although I have not found an explicit reference.

Proof. Let $\rho = \sum_k p_k |k\rangle\langle k|$ be the spectral decomposition of ρ and $r = \mathrm{rank}(\rho)$. As mentioned earlier, the von Neumann measurement minimizing the conditional entropy $\sum_i \eta_i S(\rho_{\mathsf{B}|i})$ consists of rank-one projectors $\pi_i^{\mathsf{A}} = |\varphi_i\rangle\langle\varphi_i|$. The conditional states (11.4) take the form

$$\rho_{\mathsf{B}|i} = \sum_{k=1}^{r} p_{k|i} |\phi_{ki}\rangle\langle\phi_{ki}| \quad \text{with} \quad p_{k|i} = \frac{p_k \eta_{i|k}}{\eta_i} \quad \text{and} \quad \sqrt{\eta_{i|k}}|\phi_{ki}\rangle = \langle\varphi_i|k\rangle \in \mathcal{H}_{\mathsf{B}} \,, \tag{11.30}$$

where $\eta_{i|k} = \|\langle\varphi_i|k\rangle\|^2$ is the probability of outcome i given the state $|k\rangle$ and $p_{k|i}$ is the "a posteriori" probability that the state is $|k\rangle$ given the measurement outcome i (Bayes rules). Since $\{|\phi_{ki}\rangle, p_{k|i}\}$ is a pure state decomposition of $\rho_{\mathsf{B}|i}$, the formula (7.5) yields

$$\sum_i \eta_i S(\rho_{\mathsf{B}|i}) \leq \sum_i \eta_i H(\{p_{k|i}\}) \,. \tag{11.31}$$

The right-hand side is the classical conditional entropy given the measurement outcomes, see (6.48). By the non-negativity of the classical mutual information, it is bounded from above by the Shannon entropy $H(\{p_k\}) = -\sum_k p_k \ln p_k = S(\rho)$. Hence $\delta_{\mathsf{A}}^{\mathrm{v.N.}}(\rho) \leq S(\rho_{\mathsf{A}})$ by (11.6). But $S(\rho_{\mathsf{A}}) \leq \ln n_{\mathsf{A}}$, thus we have proven (11.28).

Let us assume that $\delta_{\mathsf{A}}(\rho) = S(\rho_{\mathsf{A}})$. We know from Sec. 7.1 that a necessary and sufficient condition for (11.31) to be an equality is that $\{|\phi_{ki}\rangle, p_{k|i}\}$ be a spectral decomposition of $\rho_{\mathsf{B}|i}$, for any i. Setting $D_{kl} = \mathrm{tr}_{\mathsf{B}}(|k\rangle\langle l|)$ as in the proof of Proposition 10.5.1, one gets $\sqrt{\eta_{i|l}\eta_{i|k}}\langle\phi_{li}|\phi_{ki}\rangle = \langle\varphi_i|D_{kl}|\varphi_i\rangle = 0$ if $k \neq l$ and $p_{k|i}p_{l|i} > 0$. Since $\delta_{\mathsf{A}}(\rho) = S(\rho_{\mathsf{A}})$, (11.31) holds with equality for any orthonormal basis $\{|\varphi_i\rangle\}$ and thus $D_{kl} = 0$ for such k and l. In addition, the conditional entropy in the right-hand side of (11.31) is equal to its upper bound $H(\{p_k\}) = S(\rho)$. This can happen only if $p_{k|i} = p_k$, i.e., $\eta_{i|k} = \langle\varphi_i|D_{kk}|\varphi_i\rangle = \eta_i$, for all i and k (indeed, the mutual information vanishes for independent random variables only). Hence $\delta_{\mathsf{A}}(\rho) = S(\rho_{\mathsf{A}})$ if and only if D_{kk} is independent of k and $D_{kl} = 0$ when $k \neq l$ and $p_k p_l > 0$. Suppose now that $\delta_{\mathsf{A}}(\rho) = \ln n_{\mathsf{A}}$. Then $\delta_{\mathsf{A}}(\rho) = S(\rho_{\mathsf{A}}) = \ln n_{\mathsf{A}}$ and the foregoing conditions on D_{kl} are fulfilled. In addition, $\rho_{\mathsf{A}} = \sum p_k D_{kk} = 1/n_{\mathsf{A}}$, whence $D_{kk} = 1/n_{\mathsf{A}}$ for all k with $p_k > 0$. One concludes that the eigenvectors $|k\rangle$ are as in Proposition 10.5.1 by following the same steps as in the proof of this proposition. □

Note that when $n_{\mathsf{A}} > n_{\mathsf{B}}$, $\delta_{\mathsf{A}}(\rho)$ is strictly smaller than $\ln n_{\mathsf{A}}$ for any $\rho \in \mathcal{E}(\mathcal{H}_{\mathsf{AB}})$. In fact, in that case $\mathrm{rank}(D_{kk}) \leq n_{\mathsf{B}} < n_{\mathsf{A}}$ by the Schmidt decomposition (4.9), and the necessary condition $D_{kk} = 1/n_{\mathsf{A}}$ for having $\delta_{\mathsf{A}}(\rho) = \ln n_{\mathsf{A}}$ cannot be fulfilled.

11.3.3 Monotonicity when disregarding a part of the measured subsystem

We close this review of the properties of the discord by a simple remark concerning tripartite systems ABC. If such a system is in the state ρ_{ABC}, it is easy to show that

$$J_{\mathsf{B}|\mathsf{AC}}(\rho_{\mathsf{ABC}}) \geq J_{\mathsf{B}|\mathsf{A}}(\rho_{\mathsf{AB}}) \,. \tag{11.32}$$

This means that if B is coupled to both A and C, the gain of information on B from joint measurements on A and C is larger than the gain of information by measuring A only and ignoring C, as this sounds reasonable. A similar bound exists for the total correlations: by (11.3) and the monotonicity of the relative entropy (or, equivalently, the strong subadditivity of S),

$$I_{\mathsf{AC}:\mathsf{B}}(\rho_{\mathsf{ABC}}) \geq I_{\mathsf{A}:\mathsf{B}}(\rho_{\mathsf{AB}}) \,. \tag{11.33}$$

Remark 11.3.3. *The Holevo bound (6.51) can be derived by using the monotonicity of the quantum mutual information under operations acting on one subsystem (Sec. 11.1) and the property (11.33).*

Sketch of the proof [180]. Given an ensemble $\{\rho_i, \eta_i\}_{i=1}^m$ of states on $\mathcal{H}_{\mathsf{A}}$ and a family $\{A_j\}_{j=1}^p$ of Kraus operators describing the measurement on A, consider the state $\rho_{\mathsf{ARP}} = \sum_i \eta_i \rho_i \otimes |\nu_i\rangle\langle\nu_i| \otimes |0\rangle\langle 0|$ on $\mathcal{H}_{\mathsf{ARP}}$, where R and P are auxiliary systems with orthonormal bases $\{|\nu_i\rangle\}_{i=1}^m$ and $\{|j\rangle\}_{j=0}^{p-1}$. These systems represent a register of the state preparation and a pointer for the measurement, respectively. Let $\mathcal{M}_{\mathsf{AP}}$ be the quantum operation on $\mathcal{B}(\mathcal{H}_{\mathsf{AP}})$ with Kraus operators $A_j \otimes U_j$, U_j being the unitary on $\mathcal{H}_{\mathsf{P}}$ defined by $U_j|l\rangle = |l + j\rangle$ for any $l = 0, \ldots, p-1$ (the addition is modulo p). It is an easy exercise to show that the Holevo bound (6.51) is equivalent to $I_{\mathsf{R}:\mathsf{P}}([\mathcal{M}_{\mathsf{AP}} \otimes 1(\rho_{\mathsf{ARP}})]_{\mathsf{RP}}) \leq I_{\mathsf{AP}:\mathsf{R}}(\rho_{\mathsf{ARP}})$.

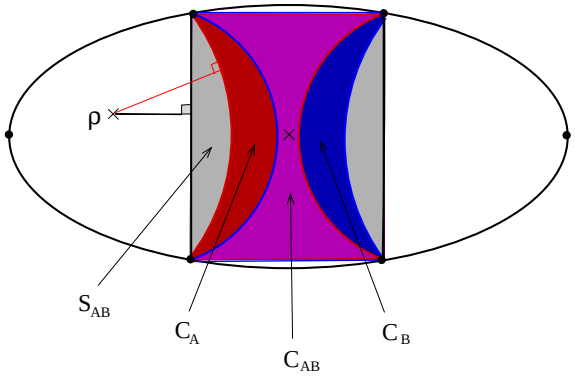

Figure 11.1: Schematic view of the set of quantum states $\mathcal{E}_{\mathsf{AB}} = \mathcal{E}(\mathcal{H}_{\mathsf{AB}})$ of a bipartite system AB. The subset $\mathcal{C}_{\mathsf{AB}}$ of classical states (in magenta) is the intersection of the subsets $\mathcal{C}_{\mathsf{A}}$ and $\mathcal{C}_{\mathsf{B}}$ of A- and B-classical states (in red and blue). The convex hull of $\mathcal{C}_{\mathsf{A}}$ (or $\mathcal{C}_{\mathsf{B}}$) is the subset $\mathcal{S}_{\mathsf{AB}}$ of separable states (gray square). All these subsets intersect the border of $\mathcal{E}_{\mathsf{AB}}$ (pure states of AB) at the pure product states, represented by the four vertices of the square. The maximally mixed state $\rho_{\mathsf{AB}} = 1/(n_{\mathsf{A}} n_{\mathsf{B}})$ lies at the center (cross). The two points at the left and right extremities of the ellipse represent the maximally entangled pure states, which are the most distant states from $\mathcal{S}_{\mathsf{AB}}$ (and also from $\mathcal{C}_{\mathsf{A}}$, $\mathcal{C}_{\mathsf{B}}$, and $\mathcal{C}_{\mathsf{AB}}$). The closest distances of a state ρ to $\mathcal{S}_{\mathsf{AB}}$ (black line) and of ρ to $\mathcal{C}_{\mathsf{A}}$ (red line) define the square roots of the geometric measure of entanglement $E_{\mathrm{Bu}}(\rho)$ and of the geometric discord $D_{\mathsf{A}}(\rho)$, respectively. Note that this picture is for illustrative purposes and does not reflect all geometrical aspects (in particular, $\mathcal{C}_{\mathsf{A}}$, $\mathcal{C}_{\mathsf{B}}$, and $\mathcal{C}_{\mathsf{AB}}$ typically have a lower dimensionality than $\mathcal{E}_{\mathsf{AB}}$ and $\mathcal{S}_{\mathsf{AB}}$).

11.4 Monogamy relation

Consider a tripartite system ABC in a pure state $|\Psi_{\mathsf{ABC}}\rangle$. If B and C are entangled, is there a limit on the amount of entanglement B can have with A? In other words, can entanglement be freely shared between different subsystems? A negative answer to the last question has been highlighted in [64], where it is shown that when A, B, and C are qubits, the sum $C(\rho_{\mathsf{AB}})^2 + C(\rho_{\mathsf{BC}})^2$ of the square concurrences is smaller or equal to $4\det(\rho_{\mathsf{B}})$. It is instructive to consider the limiting case where B and C are maximally entangled. Then, if one also assumes that $rn_{\mathsf{B}} \leq n_{\mathsf{C}} < (r+1)n_{\mathsf{B}}$ with $1 \leq r \leq n_{\mathsf{A}}$, A and B cannot be entangled and even have vanishing discords $\delta_{\mathsf{A}}(\rho_{\mathsf{AB}}) = \delta_{\mathsf{B}}(\rho_{\mathsf{AB}}) = 0$. In fact, the state of BC being maximally entangled, one has $\rho_{\mathsf{BC}} = \sum_k p_k |k\rangle\langle k|$ for some orthogonal maximally entangled states $|k\rangle$ satisfying $D_{kl} = \mathrm{tr}_{\mathsf{C}}(|k\rangle\langle l|) = n_{\mathsf{B}}^{-1}\delta_{kl}$ (see Proposition 10.5.1). Hence the pure state of ABC is $|\Psi_{\mathsf{ABC}}\rangle = \sum_k \sqrt{p_k}|\alpha_k\rangle|k\rangle$ with $\{|\alpha_k\rangle\}$ an orthonormal family of $\mathcal{H}_{\mathsf{A}}$ (Sec. 4.3). Consequently, $\rho_{\mathsf{AB}} = (\sum_k p_k |\alpha_k\rangle\langle\alpha_k|) \otimes (1/n_{\mathsf{B}})$ is a product state and thus a classical state.

The proposition below exhibits an astonishing bound, called the monogamy relation, between the entanglement of formation of ρ_{BC} and the POVM-discord of ρ_{AB} measuring A.

Proposition 11.4.1. (Koashi and Winter [148]) *Let ABC be a tripartite system in the state ρ_{ABC}. Let $\rho_{\mathsf{AB}} = \mathrm{tr}_{\mathsf{C}}(\rho_{\mathsf{ABC}})$ and $\rho_{\mathsf{BC}} = \mathrm{tr}_{\mathsf{A}}(\rho_{\mathsf{ABC}})$ denote the reduced states of the bipartite systems AB and BC, respectively. Then*

$$E_{\mathrm{EoF}}(\rho_{\mathsf{BC}}) \leq S(\rho_{\mathsf{B}}) - J_{\mathsf{B}|\mathsf{A}}(\rho_{\mathsf{AB}}) = \delta_{\mathsf{A}}(\rho_{\mathsf{AB}}) + S(\rho_{\mathsf{AB}}) - S(\rho_{\mathsf{A}}) \,. \tag{11.34}$$

Moreover, the inequality is an equality if ρ_{ABC} is a pure state.

The inequality (11.34) tells us that the more classically correlated are A and B, the less B can be entangled to a third system C. If $n_{\mathsf{B}} \leq n_{\mathsf{C}}$ and B and C are maximally entangled, i.e., $E_{\mathrm{EoF}}(\rho_{\mathsf{BC}}) = \ln(n_{\mathsf{B}})$, then this inequality entails $J_{\mathsf{B}|\mathsf{A}}(\rho_{\mathsf{AB}}) = 0$ (since $S(\rho_{\mathsf{B}}) \leq \ln(n_{\mathsf{B}})$). Thus A and B are not correlated classically, in agreement with the above statement that ρ_{AB} is a product state.

The entropy difference $S_{\mathsf{B}|\mathsf{A}}(\rho_{\mathsf{AB}}) = S(\rho_{\mathsf{AB}}) - S(\rho_{\mathsf{A}})$ in the right-hand side of (11.34) is called the *conditional von Neumann entropy*. It is known that $S_{\mathsf{B}|\mathsf{A}}(\rho_{\mathsf{AB}}) \geq 0$ if ρ_{AB} is separable [129, 56]. Thanks to the subadditivity of S one has $-S(\rho_{\mathsf{B}}) \leq S_{\mathsf{B}|\mathsf{A}}(\rho_{\mathsf{AB}}) \leq S(\rho_{\mathsf{B}})$ (the first inequality is obtained by considering a purification of ρ_{AB} on $\mathcal{H}_{\mathsf{ABC}}$ and using the subadditivity for ρ_{BC} together with the identities $S(\rho_{\mathsf{BC}}) = S(\rho_{\mathsf{A}})$ and $S(\rho_{\mathsf{C}}) = S(\rho_{\mathsf{AB}})$). The quantity $-S_{\mathsf{B}|\mathsf{A}}(\rho_{\mathsf{AB}})$ is the coherent information introduced by Schumacher and Nielsen in the context of the quantum channel capacity [214].

Two consequences of the claim that (11.34) is an equality for tripartite systems ABC in pure states deserve further comments. First, one easily deduces from this claim and the identity (7.4) that [88]

$$E_{\rm EoF}(\rho_{\mathsf{AB}}) + E_{\rm EoF}(\rho_{\mathsf{BC}}) = \delta_{\mathsf{A}}(\rho_{\mathsf{AB}}) + \delta_{\mathsf{C}}(\rho_{\mathsf{BC}}) \; . \tag{11.35}$$

Hence the sum of all entanglement of formations describing the bipartite entanglement shared by B is equal to the sum of the corresponding quantum discords with measurements on the other subsystems. Second, if B is a qubit and ρ_{AB} is of rank two, then ρ_{AB} admits a purification $|\Psi_{\mathsf{ABC}}\rangle$ on $\mathcal{H}_{\mathsf{AB}} \otimes \mathbb{C}^2$ (see (4.12)) and the entanglement of formation of the two-qubit state ρ_{BC} can be computed with the help of the Wootters formula (10.31). One may in this way determine $\delta_{\mathsf{A}}(\rho_{\mathsf{AB}})$ via (11.34).

Proof. We first assume that ABC is in a pure state $|\Psi_{\mathsf{ABC}}\rangle$. Let $\{M^{\rm opt}_{\mathsf{A},i}\}$ be an optimal measurement on A maximizing the gain of information on B, that is, such that $J_{\mathsf{B}|\mathsf{A}}(\rho_{\mathsf{AB}}) = S(\rho_{\mathsf{B}}) - \sum_i \eta_i^{\rm opt} S(\rho^{\rm opt}_{\mathsf{B}|i})$, where $\eta_i^{\rm opt}$ and $\rho^{\rm opt}_{\mathsf{B}|i}$ are the outcome probabilities and conditional states of B for this measurement. Without loss of generality one may assume that $M^{\rm opt}_{\mathsf{A},i} = |\widetilde{\mu}_i^{\rm opt}\rangle\langle\widetilde{\mu}_i^{\rm opt}|$ are of rank one (see the discussion after (11.5)). Since $\rho_{\mathsf{AB}} = {\rm tr}_{\mathsf{C}}(|\Psi_{\mathsf{ABC}}\rangle\langle\Psi_{\mathsf{ABC}}|)$, one has $\eta_i^{\rm opt} = {\rm tr}(\rho_{\mathsf{AB}} M^{\rm opt}_{\mathsf{A},i} \otimes 1) = \|\langle\widetilde{\mu}_i^{\rm opt}|\Psi_{\mathsf{ABC}}\rangle\|^2$. Moreover, the post-measurement conditional state of BC is the pure state

$$|\Psi_{\mathsf{BC}|i}\rangle = (\eta_i^{\rm opt})^{-\frac{1}{2}} \langle\widetilde{\mu}_i^{\rm opt}|\Psi_{\mathsf{ABC}}\rangle \tag{11.36}$$

and the conditional state of B is $\rho^{\rm opt}_{\mathsf{B}|i} = {\rm tr}_{\mathsf{C}}(|\Psi_{\mathsf{BC}|i}\rangle\langle\Psi_{\mathsf{BC}|i}|)$. The ensemble $\{|\Psi_{\mathsf{BC}|i}\rangle, \eta_i^{\rm opt}\}$ gives a pure state decomposition of ρ_{BC}. Actually, let us consider the post-measurement state of ABC in the absence of readout, $\rho'_{\mathsf{ABC}} = \mathcal{M}^{\rm opt}_{\mathsf{A}} \otimes 1(|\Psi_{\mathsf{ABC}}\rangle\langle\Psi_{\mathsf{ABC}}|)$. The measurement being performed on A, it does not change the state of BC, i.e.

$$\rho_{\mathsf{BC}} = \rho'_{\mathsf{BC}} = \sum_i \eta_i^{\rm opt} |\Psi_{\mathsf{BC}|i}\rangle\langle\Psi_{\mathsf{BC}|i}| \; . \tag{11.37}$$

From the definition (10.25) of the entanglement of formation one has

$$E_{\rm EoF}(\rho_{\mathsf{BC}}) \leq \sum_i \eta_i^{\rm opt} S(\rho^{\rm opt}_{\mathsf{B}|i}) = S(\rho_{\mathsf{B}}) - J_{\mathsf{B}|\mathsf{A}}(\rho_{\mathsf{AB}}) \; . \tag{11.38}$$

Conversely, let $\{|\Psi_{\mathsf{BC},i}\rangle, \eta_i\}$ be a pure state decomposition of ρ_{BC} which achieves the minimum in the definition of the entanglement of formation. Let us show that there exists a generalized measurement $\{M_i^{\mathsf{A}}\}$ on A such that η_i is the probability of outcome i and $|\Psi_{\mathsf{BC},i}\rangle$ the corresponding conditional state of BC, i.e.

$${\rm tr}_{\mathsf{A}}(M_i^{\mathsf{A}} \otimes 1 |\Psi_{\mathsf{ABC}}\rangle\langle\Psi_{\mathsf{ABC}}|) = \eta_i |\Psi_{\mathsf{BC},i}\rangle\langle\Psi_{\mathsf{BC},i}| \; . \tag{11.39}$$

In fact, let us observe that $|\Psi'_{\mathsf{ABCE}}\rangle = \sum_i \sqrt{\eta_i} |\Psi_{\mathsf{BC},i}\rangle|\phi_i\rangle$ is a purification of ρ_{BC} on $\mathcal{H}_{\mathsf{ABCE}}$ for some ancilla E, where $\{|\phi_i\rangle\}$ is an orthonormal family of $\mathcal{H}_{\mathsf{AE}}$. Given an arbitrary state $|\epsilon_0\rangle \in \mathcal{H}_{\mathsf{E}}$, $|\Psi_{\mathsf{ABC}}\rangle|\epsilon_0\rangle$ is also a purification of ρ_{BC} on the same space. As a result, there is a unitary U_{AE} on $\mathcal{H}_{\mathsf{AE}}$ such that $|\Psi'_{\mathsf{ABCE}}\rangle = 1 \otimes U_{\mathsf{AE}} |\Psi_{\mathsf{ABC}}\rangle|\epsilon_0\rangle$ (see Sec. 4.3). Define

$$M_i^{\mathsf{A}} = \langle\epsilon_0|U^*_{\mathsf{AE}}|\phi_i\rangle\langle\phi_i|U_{\mathsf{AE}}|\epsilon_0\rangle \tag{11.40}$$

(note the analogy with (5.20)). Then (11.39) is satisfied. Let $\rho_{\mathsf{B}|i} = {\rm tr}_{\mathsf{C}}(|\Psi_{\mathsf{BC},i}\rangle\langle\Psi_{\mathsf{BC},i}|)$ be the post-measurement states of B, so that $E_{\rm EoF}(|\Psi_{\mathsf{BC},i}\rangle) = S(\rho_{\mathsf{B}|i})$. Since by assumption $E_{\rm EoF}(\rho_{\mathsf{BC}}) = \sum_i \eta_i E_{\rm EoF}(|\Psi_{\mathsf{BC},i}\rangle)$, one infers from the definition (11.5) of the classical correlations that

$$J_{\mathsf{B}|\mathsf{A}}(\rho_{\mathsf{AB}}) \geq S(\rho_{\mathsf{B}}) - \sum_i \eta_i S(\rho_{\mathsf{B}|i}) = S(\rho_{\mathsf{B}}) - E_{\rm EoF}(\rho_{\mathsf{BC}}) \; . \tag{11.41}$$

Together with (11.38) this proves that

$$E_{\rm EoF}(\rho_{\mathsf{BC}}) = S(\rho_{\mathsf{B}}) - J_{\mathsf{B}|\mathsf{A}}(\rho_{\mathsf{AB}}) \; . \tag{11.42}$$

Let us now turn to the case of a tripartite system ABC in a mixed state ρ_{ABC}. Consider a purification $|\Psi_{\mathsf{ABCE}}\rangle$ of ρ_{ABC} in the Hilbert space $\mathcal{H}_{\mathsf{ABC}} \otimes \mathcal{H}_{\mathsf{E}}$. Thanks to (11.32) one then has $J_{\mathsf{B}|\mathsf{A}}(\rho_{\mathsf{AB}}) \leq J_{\mathsf{B}|\mathsf{AE}}(\rho_{\mathsf{ABE}})$. The inequality (11.34) then follows by applying (11.42) with $\mathsf{A} \to \mathsf{AE}$. □

Chapter 12

Distance and entropic measures of quantum correlations

Explicar un hecho es unirlo a otro[1] (J.L. Borges).

In this chapter we study the measures of entanglement and quantum correlations based on the Bures distance and the relative entropies. First, we introduce in Sec. 12.1 the geometric measure of entanglement, defined as the minimal square distance between the state ρ and a separable state, as well as similar measures obtained by replacing the square distance by relative entropies. We define analogously in Sec. 12.2 the geometric discord as the minimal square distance between ρ and an A-classical state. We show there that this discord is related to a quantum state discrimination task and determine the closest A-classical states to ρ in terms of the corresponding optimal measurements.

12.1 Geometric and relative-entropy measures of entanglement

12.1.1 Definition and main properties

From a geometrical point of view, it is natural to quantify the amount of entanglement in a state ρ of a bipartite system AB by the distance $d(\rho, \mathcal{S}_{\mathsf{AB}})$ of ρ to the subset $\mathcal{S}_{\mathsf{AB}} \subset \mathcal{E}(\mathcal{H}_{\mathsf{AB}})$ of separable states (see Fig 11.1). As it will become clear below, in order to obtain an entanglement monotone measure the distance d must be contractive. Choosing the Bures distance, it is easy to verify that

$$E_{\mathrm{Bu}}(\rho) = d_{\mathrm{B}}(\rho, \mathcal{S}_{\mathsf{AB}})^2 = \min_{\sigma_{\mathrm{sep}} \in \mathcal{S}_{\mathsf{AB}}} \left\{ d_B(\rho, \sigma_{\mathrm{sep}})^2 \right\} \tag{12.1}$$

satisfies all the axioms of an entanglement measure in Definition 10.3.1. Actually, the axiom (i) holds because d_{B} is a distance on $\mathcal{E}(\mathcal{H}_{\mathsf{AB}})$. The convexity property (ii) is a consequence of the convexity of $\mathcal{S}_{\mathsf{AB}}$ and the joint convexity of the square Bures distance[2] (Corollary 8.2.3). Finally, the monotonicity (iii) is shown in the following way. Let $\sigma_\rho \in \mathcal{S}_{\mathsf{AB}}$ be a closest separable state to ρ, i.e., $E_{\mathrm{Bu}}(\rho) = d_{\mathrm{B}}(\rho, \sigma_\rho)^2$. Let us recall from Sec. 10.3 that any LOCC is a separable quantum operation and can be written as $\mathcal{M}(\rho) = \sum_i A_i \otimes B_i \rho A_i^* \otimes B_i^*$. Furthermore, one has $\mathcal{M}(\mathcal{S}_{\mathsf{AB}}) \subset \mathcal{S}_{\mathsf{AB}}$. One can then use the contractivity of d_{B} to obtain

$$E_{\mathrm{Bu}}(\rho) \geq d_{\mathrm{B}}(\mathcal{M}(\rho), \mathcal{M}(\sigma_\rho))^2 \geq E_{\mathrm{Bu}}(\mathcal{M}(\rho)) \; . \tag{12.2}$$

This shows that E_{Bu} is monotonous with respect to separable operations and, in particular, to LOCCs. The entanglement measure E_{Bu} has been first introduced by Vedral and Plenio [246]. Another measure was considered in [245, 246] by replacing the square distance in (12.1) by the relative entropy $S(\rho||\sigma_{\mathrm{sep}})$. More generally, we can define

$$E_\alpha(\rho) = \min_{\sigma_{\mathrm{sep}} \in \mathcal{S}_{\mathsf{AB}}} \left\{ S_\alpha(\rho||\sigma_{\mathrm{sep}}) \right\} , \tag{12.3}$$

where S_α is the quantum relative Rényi entropy (Sec. 7.3). For $1/2 \leq \alpha \leq 1$, this defines an entanglement measure by the same arguments as above, because S_α is jointly convex and contractive (see Theorem 7.3.1; the

[1] "Explaining a fact is linking it to another".

[2] This justifies the square in our definition (12.1).

property (i) in this theorem ensures that $E_\alpha(\rho) \geq 0$ with equality if and only if $\rho \in \mathcal{S}_{\mathsf{AB}}$). One establishes the following result by invoking the fact that S_α is non-decreasing in α (Proposition 7.3.4) and by using (8.8) and the relation (7.36) between $S_{1/2}(\rho||\sigma)$ and the fidelity $F(\rho,\sigma)$.

Corollary 12.1.1. $\{E_\alpha\}_{1/2\leq\alpha\leq 1}$ *constitutes a non-decreasing family of entanglement measures and*

$$E_{\frac{1}{2}}(\rho) = -2\ln\Big(1 - \frac{E_{\rm Bu}(\rho)}{2}\Big) \;\leq\; E_\alpha(\rho) \quad , \quad \frac{1}{2} \leq \alpha \leq 1 \; . \tag{12.4}$$

The measure E_1 associated to the relative entropy (7.9) is less geometrical than $E_{\rm Bu}$ (it is not associated to a distance) but has the following interesting property.

Proposition 12.1.2. (Vedral and Plenio [246]) *The entanglement measure E_1 coincides with the entanglement of formation $E_{\rm EoF}$ for pure states, and for mixed states $\rho \in \mathcal{E}(\mathcal{H}_{\mathsf{AB}})$ it is bounded from above by $E_{\rm EoF}$,*

$$E_1(\rho) \leq E_{\rm EoF}(\rho) \; . \tag{12.5}$$

Proof. We refer the reader to [246] for a detailed proof of the first statement. It is based on the observation that for a pure state with Schmidt decomposition $|\Psi\rangle = \sum_i \sqrt{\mu_i}|\alpha_i\rangle|\beta_i\rangle$, the minimum in (12.3) is achieved when $\sigma_{\rm sep}$ is the classical state

$$\sigma_* = \sum_{i=1}^n \mu_i |\alpha_i\rangle\langle\alpha_i| \otimes |\beta_i\rangle\langle\beta_i| \; . \tag{12.6}$$

Since $S(\rho_\Psi||\sigma_*) = -\langle\Psi|\ln\sigma_*|\Psi\rangle = -\sum_i \mu_i \ln\mu_i$, the equality $E_1(|\Psi\rangle) = E_{\rm EoF}(|\Psi\rangle)$ follows once one has proven that $S(\rho_\Psi||\sigma_{\rm sep}) \geq S(\rho_\Psi||\sigma_*)$ for all $\sigma_{\rm sep} \in \mathcal{S}_{\mathsf{AB}}$. This is done in Ref. [246] by showing that for any $\sigma_{\rm sep} \in \mathcal{S}_{\mathsf{AB}}$,

$$\frac{{\rm d}f_\Psi(t,\sigma_{\rm sep})}{{\rm d}t}\Big|_{t=0} = 1 - \int_0^\infty {\rm d}t \; {\rm tr}\big((\sigma_*+t)^{-1}\rho_\Psi(\sigma_*+t)^{-1}\sigma_{\rm sep}\big) \;\geq\; 0 \tag{12.7}$$

with $f_\Psi(t,\sigma) = S(\rho_\Psi||(1-t)\sigma_* + t\sigma)$. Indeed, assume that $S(\rho_\Psi||\sigma_{\rm sep}) < S(\rho_\Psi||\sigma_*)$ for some $\sigma_{\rm sep} \in \mathcal{S}_{\mathsf{AB}}$. By taking advantage of the right convexity of the relative entropy, one then finds for any $t \in (0,1]$

$$\frac{f_\Psi(t,\sigma_{\rm sep}) - f_\Psi(0,\sigma_{\rm sep})}{t} \leq -S(\rho_\Psi||\sigma_*) + S(\rho_\Psi||\sigma_{\rm sep}) < 0 \; , \tag{12.8}$$

in contradiction with (12.7). Note that it suffices to prove the non-negativity in (12.7) for the pure product states $\sigma_{\rm sep} = |\phi\otimes\chi\rangle\langle\phi\otimes\chi|$, because of the linearity in $\sigma_{\rm sep}$ of the trace in the right-hand side.

The second statement in the proposition is a consequence of the first one and of the convexity of E_1. Actually, if $\{|\Psi_i\rangle, \eta_i\}$ is a pure state decomposition of ρ minimizing the average entanglement, then

$$E_{\rm EoF}(\rho) = \sum_i \eta_i E_{\rm EoF}(|\Psi_i\rangle) = \sum_i \eta_i E_1(|\Psi_i\rangle) \geq E_1\Big(\sum_i \eta_i |\Psi_i\rangle\langle\Psi_i|\Big) = E_1(\rho) \; . \tag{12.9}$$

□

Note that the inequality (12.5) can be strict. Examples of two-qubit states ρ for which $E_1(\rho) < E_{\rm EoF}(\rho)$ are given in [245]. Thanks to (12.4) and (12.5), one can place an upper bound on $E_{\rm Bu}(\rho)$ by a function of the entanglement of formation $E_{\rm EoF}$. Such a bound does not seem to be known in the literature, but it is not optimal for pure states as a consequence of the next proposition.

Remark 12.1.3. *As shown in [246], E_1 fulfills the stronger monotonicity condition (iii") of Sec. 10.4.2.*

12.1.2 Geometric measure of entanglement and convex roof constructions

Let $F(\rho,\mathcal{S}_{\mathsf{AB}})$ denote the maximal fidelity between ρ and a separable state,

$$F(\rho,\mathcal{S}_{\mathsf{AB}}) = \max_{\sigma_{\rm sep}\in\mathcal{S}_{\mathsf{AB}}} \big\{F(\rho,\sigma_{\rm sep})\big\} \; . \tag{12.10}$$

Proposition 12.1.4. (Streltsov, Kampermann, and Bruß [231]) *The geometric measure of entanglement is given for pure states by*

$$E_{\rm Bu}(|\Psi\rangle) = 2 - 2\sqrt{F(|\Psi\rangle, \mathcal{S}_{\mathsf{AB}})} = 2(1 - \sqrt{\mu_{\max}}) \,, \tag{12.11}$$

where $\mu_{\max} = \max\{\mu_i\}$ *is the largest Schmidt coefficient of* $|\Psi\rangle$*. For mixed states,* $F(\rho, \mathcal{S}_{\mathsf{AB}})$ *is obtained via a maximization over the pure state decompositions of* ρ*,*

$$F(\rho, \mathcal{S}_{\mathsf{AB}}) = \max_{\{|\Psi_i\rangle, \eta_i\}} \left\{ \sum_i \eta_i F(|\Psi_i\rangle, \mathcal{S}_{\mathsf{AB}}) \right\} . \tag{12.12}$$

The nice relation (12.12) is intimately related to Uhlmann's theorem (Sec. 8.2) and to the convexity of $\mathcal{S}_{\mathsf{AB}}$. Note that the relative-entropy measure E_1 does not fulfill a similar property (compare with Proposition 12.1.2). Even though $E_{\rm Bu}$ is not a convex roof, it is a simple function of another entanglement measure $E_{\rm G}$ defined via a convex-roof construction like in (10.25) and from its expression

$$E_{\rm G}(|\Psi\rangle) = 1 - \max_{|\Phi\rangle \in \mathcal{S}_{\mathsf{AB}}} \left\{ |\langle \Phi | \Psi \rangle|^2 \right\} \tag{12.13}$$

for pure states [216, 255]. Actually, we will see that a pure state always admits a pure product state as closest separable state, hence the maximum in (12.13) coincides with $F(|\Psi\rangle, \mathcal{S}_{\mathsf{AB}})$ and $E_{\rm G}(\rho) = 1 - F(\rho, \mathcal{S}_{\mathsf{AB}})$ by the proposition above. According to (12.11), $E_{\rm G}(|\Psi\rangle) = 1 - \mu_{\max}$ is of the form (10.11) with $f_{\rm G}(\rho_{\mathsf{A}}) = 1 - \|\rho_{\mathsf{A}}\|$ satisfying all hypothesis of Proposition 10.3.2. Therefore, by a similar reasoning as in the proof of Proposition 10.4.3, $E_{\rm G}$ is an entanglement measure which fulfills the strong monotonicity property (iii"). In contrast, $E_{\rm Bu}(|\Psi\rangle) = f_{\rm Bu}([\rho_\Psi]_{\mathsf{A}}) = 2(1 - \sqrt{\|[\rho_\Psi]_{\mathsf{A}}\|})$ but $f_{\rm Bu}$ is not concave, whence Proposition 10.3.2 indicates that $E_{\rm Bu}$ does not fulfill (iii'). We should not be bothered too much about that, the two measures $E_{\rm Bu}$ and $E_{\rm G}$ being equivalent (that is, they define the same order of entanglement) and simply related to each other.

Proof. For a pure state $\rho_\Psi = |\Psi\rangle\langle\Psi|$, the fidelity reads $F(\rho_\Psi, \sigma_{\rm sep}) = \langle \Psi | \sigma_{\rm sep} | \Psi \rangle$. Writing the decomposition of separable states into pure product states, $\sigma_{\rm sep} = \sum_i \xi_i |\varphi_i \otimes \chi_i\rangle\langle \varphi_i \otimes \chi_i|$, we get

$$F(\rho_\Psi, \mathcal{S}_{\mathsf{AB}}) = \max_{\{|\varphi_i\rangle, |\chi_i\rangle, \xi_i\}} \left\{ \sum_i \xi_i |\langle \varphi_i \otimes \chi_i | \Psi \rangle|^2 \right\} = \max_{\|\varphi\| = \|\chi\| = 1} \left\{ |\langle \varphi \otimes \chi | \Psi \rangle|^2 \right\} , \tag{12.14}$$

where we have used $\sum_i \xi_i = 1$. For any normalized vectors $|\varphi\rangle \in \mathcal{H}_A$ and $|\chi\rangle \in \mathcal{H}_B$, one derives from the Schmidt decomposition (4.9) and the Cauchy-Schwarz inequality that

$$\begin{aligned} |\langle \varphi \otimes \chi | \Psi \rangle| &\leq \sum_{j=1}^n \sqrt{\mu_j} \big| \langle \varphi | \alpha_j \rangle \langle \chi | \beta_j \rangle \big| \leq \sqrt{\mu_{\max}} \sum_{j=1}^n \big| \langle \varphi | \alpha_j \rangle \langle \chi | \beta_j \rangle \big| \\ &\leq \sqrt{\mu_{\max}} \left(\sum_{j=1}^n |\langle \varphi | \alpha_j \rangle|^2 \right)^{1/2} \left(\sum_{j=1}^n |\langle \chi | \beta_j \rangle|^2 \right)^{1/2} \leq \sqrt{\mu_{\max}} \,. \end{aligned} \tag{12.15}$$

All bounds are saturated for $|\varphi\rangle = |\alpha_{j_{\max}}\rangle$ and $|\chi\rangle = |\beta_{j_{\max}}\rangle$, where $j_{\max}$ is the index for which μ_j is maximum. Thus $F(\rho_\Psi, \mathcal{S}_{\mathsf{AB}}) = \mu_{j_{\max}} = \mu_{\max}$ and the formula (12.11) is proven. It is of interest to note that the pure product state $|\alpha_{j_{\max}}\rangle |\beta_{j_{\max}}\rangle$ is a closest separable state to $|\Psi\rangle$ (a characterization of all these closest separable states will be given in Proposition 12.2.2 below).

We now proceed to show (12.12). Consider a fixed separable state $\sigma_{\rm sep} = \sum_{i=1}^p \xi_i |\Phi_i\rangle\langle\Phi_i|$ with $|\Phi_i\rangle \in \mathcal{S}_{\mathsf{AB}}$ and $\xi_i \geq 0$. Without loss of generality one may assume $p = (n_{\mathsf{A}} n_{\mathsf{B}})^2 + 1$ (see the discussion after Definition 4.4.1). Let $\{|f_i\rangle\}_{i=1}^p$ be an orthonormal basis of an ancilla space $\mathcal{K}$ and $|\Phi\rangle = \sum_i \sqrt{\xi_i} |\Phi_i\rangle |f_i\rangle$ be a purification of $\sigma_{\rm sep}$ on $\mathcal{H} \otimes \mathcal{K}$. Thanks to Theorem 8.2.2, $F(\rho, \sigma_{\rm sep})$ is the maximum over all purifications $|\Psi\rangle$ of ρ on $\mathcal{H} \otimes \mathcal{K}$ of the transition probability $|\langle \Psi | \Phi \rangle|^2$. Writing $|\Psi\rangle$ in the form (4.15) and using the one-to-one correspondence between pure state decompositions and purifications (see Sec. 4.3), one can equivalently maximize over all pure state decompositions $\{|\Psi_i\rangle, \eta_i\}$ of ρ. Moreover, the maximization of $F(\rho, \sigma_{\rm sep})$ over the separable states $\sigma_{\rm sep}$ leads to a maximization over the pure state ensembles $\{|\Phi_i\rangle, \xi_i\}$ in $\mathcal{S}_{\mathsf{AB}}$. This yields

$$F(\rho, \mathcal{S}_{\mathsf{AB}}) = \max_{\{|\Phi_i\rangle, \xi_i\}} \max_{\{|\Psi_i\rangle, \eta_i\}} \left\{ \left| \sum_{i=1}^p \sqrt{\eta_i \xi_i} \langle \Psi_i | \Phi_i \rangle \right|^2 \right\} . \tag{12.16}$$

But, using once more the Cauchy-Schwarz inequality and $\sum_i \xi_i = 1$, one has

$$\max_{\{|\Phi_i\rangle, \xi_i\}} \left\{ \left| \sum_{i=1}^p \sqrt{\eta_i \xi_i} \langle \Psi_i | \Phi_i \rangle \right|^2 \right\} = \sum_{i=1}^p \eta_i \max_{|\Phi\rangle \in \mathcal{S}_{\mathsf{AB}}} \left\{ |\langle \Psi_i | \Phi \rangle|^2 \right\} . \tag{12.17}$$

It has been argued above that the maximal fidelity between $|\Psi_i\rangle$ and a separable state is attained for pure product states, thus $F(|\Psi_i\rangle, \mathcal{S}_{\mathsf{AB}}) = \max_{|\Phi\rangle\in\mathcal{S}_{\mathsf{AB}}} |\langle\Psi_i|\Phi\rangle|^2$. Substituting this expression into (12.17) and (12.16), we arrive at the required relation (12.12). □

According to (12.11), $E_{\rm Bu}(|\Psi\rangle) = 0$ if and only if $|\Psi\rangle$ is a product state, in agreement with the fact that separable pure states are product states. Another consequence of (12.11) and of the bound $\mu_{\max} \geq 1/n$ (which follows from $\sum_i \mu_i = 1$) is $F(|\Psi\rangle, \mathcal{S}_{\mathsf{AB}}) \geq 1/n$, with $n = \min\{n_{\mathsf{A}}, n_{\mathsf{B}}\}$. Furthermore, $F(|\Psi\rangle, \mathcal{S}_{\mathsf{AB}}) = 1/n$ if and only if $|\Psi\rangle$ is maximally entangled (Sec. 4.4). One deduces from (12.12) that

$$E_{\rm Bu}(\rho) \leq 2 - \frac{2}{\sqrt{n}} \,. \tag{12.18}$$

By the same arguments as in the proof of Proposition 10.5.1, this bound is saturated if and only if ρ has maximal entanglement of formation $E_{\rm EoF}(\rho) = \ln n$. This means that $E_{\rm Bu}$ and $E_{\rm EoF}$ capture the same maximally entangled states.

12.1.3 Geometric measure of entanglement for two qubits

In the case of two qubits, a closed formula for $E_{\rm Bu}(\rho)$ can be obtained with the help of Proposition 12.1.4 and of Wootters's result on the concurrence (Sec. 10.4.3). It reads [231]

$$E_{\rm Bu}(\rho) = 2 - \sqrt{2}\big(1 + \sqrt{1 - C(\rho)^2}\big)^{\frac{1}{2}} \tag{12.19}$$

with $C(\rho)$ given by (10.33). Actually, for pure states one finds by comparing $C(|\Psi\rangle) = 2\sqrt{\mu_0\mu_1}$ and (12.11) that $F(|\Psi\rangle, \mathcal{S}_{\mathsf{AB}}) = g(C(|\Psi\rangle))$ with $g(C) = (1 + \sqrt{1 - C^2})/2$. As g is decreasing and concave, (10.27) and (12.12) yield $F(\rho, \mathcal{S}_{\mathsf{AB}}) \leq g(C(\rho))$. But it is shown in [265] that there is an optimal pure state decomposition $\{|\Psi_i\rangle, \eta_i\}$ of ρ such that $C(\rho) = C(|\Psi_i\rangle)$ for any i. Thus

$$g\big(C(\rho)\big) \geq F(\rho, \mathcal{S}_{\mathsf{AB}}) \geq \sum_i \eta_i F(|\Psi_i\rangle, \mathcal{S}_{\mathsf{AB}}) = \sum_i \eta_i g\big(C(|\Psi_i\rangle)\big) = g\big(C(\rho)\big) \,, \tag{12.20}$$

which justifies (12.19).

12.2 Geometric quantum discord

12.2.1 Discord-like measures of quantum correlations

In the same spirit as for the geometric measure of entanglement, one defines the geometric quantum discord as

$$D_{\mathsf{A}}(\rho) = d_{\rm B}(\rho, \mathcal{C}_{\mathsf{A}})^2 = 2(1 - \sqrt{F(\rho, \mathcal{C}_{\mathsf{A}})}) \quad , \quad F(\rho, \mathcal{C}_{\mathsf{A}}) = \max_{\sigma_{\mathsf{A}-\rm cl}\in\mathcal{C}_{\mathsf{A}}} \big\{F(\rho, \sigma_{\mathsf{A}-\rm cl})\big\} \,, \tag{12.21}$$

where $\mathcal{C}_{\mathsf{A}}$ is the (non-convex) set of A-classical states (see Definition 11.2.3). One can introduce similarly the relative-entropy discords

$$D^{(\alpha)}_{\mathsf{A}}(\rho) = \min_{\sigma_{\mathsf{A}-\rm cl}\in\mathcal{C}_{\mathsf{A}}} \big\{S_\alpha(\rho||\sigma_{\mathsf{A}-\rm cl})\big\} \,. \tag{12.22}$$

As in Corollary 12.1.1 one has $D^{(1/2)}_{\mathsf{A}}(\rho) = -2\ln(1 - D_{\mathsf{A}}(\rho)/2) \leq D^{(\alpha)}_{\mathsf{A}}(\rho)$ for any $\alpha \in [1/2, 1]$.

An analog of the geometric discord D_{A} based on the Hilbert-Schmidt distance d_2 has been first introduced by Dakić, Vedral, and Brukner [67]. We hope to have convinced the reader in chapter 8 that the Bures distance is a more natural choice in quantum information. We will see that the discord (12.21) shares many of the properties of the quantum discord δ_{A} of chapter 11, while its analog with the d_2-distance has unpleasant features. In particular, like δ_{A} the Bures geometric discord is invariant under conjugations by local unitaries and contractive with respect to quantum operations $\mathcal{M}_{\mathsf{B}}$ on B. For indeed, the set of A-classical states is invariant under such transformations (see (11.13)), whence

$$D_{\mathsf{A}}(U_{\mathsf{A}} \otimes U_{\mathsf{B}}\, \rho\, U_{\mathsf{A}}^* \otimes U_{\mathsf{B}}^*) = D_{\mathsf{A}}(\rho) \quad , \quad D_{\mathsf{A}}(1 \otimes \mathcal{M}_{\mathsf{B}}(\rho)) \leq D_{\mathsf{A}}(\rho) \tag{12.23}$$

by unitary invariance and contractivity of $d_{\rm B}$. These properties also hold for $D^{(\alpha)}_{\mathsf{A}}$, $1/2 \leq \alpha \leq 1$, because the relative Rényi entropy is also contractive (Theorem 7.3.1). This should be contrasted with the non-monotonicity with respect to operations on B of the Hilbert-Schmidt geometric discord, which is due to the lack of monotonicity of d_2 (Sec. 8.1). An explicit counter-example is given in [198]. We now precise the axioms on discord-like correlation measures.

Definition 12.2.1. *A measure of quantum correlations of a bipartite system* AB *with respect to subsystem* A *is a function* $D_{\mathsf{A}} : \mathcal{E}(\mathcal{H}_{\mathsf{AB}}) \to [0, \infty)$ *satisfying*

(i) $D_{\mathsf{A}}(\rho) = 0$ *if and only if* ρ *is* A*-classical;*

(ii) D_{A} *is invariant under local unitary transformations and contractive under quantum operations on* B*, that is, (12.23) holds true;*

(iii) D_{A} *coincides with an entanglement measure for pure states.*

This definition is at the time of writing of this article believed to capture all relevant physical requirements for quantifying the amount of quantum correlations in AB given that one can access to subsystem A only [203]. The axioms (i-iii) are in particular satisfied by the quantum discord δ_{A} (Propositions 11.2.1 and 11.3.1). This is also true for the geometric discord D_{A}. Actually, we have just shown above that D_{A} satisfies (ii), and (i) is trivial. Since the closest separable state to a pure state is a pure product state, which is A-classical, D_{A} coincides with the geometric measure of entanglement E_{Bu} for pure states (see (12.26) below). Hence D_{A} is a measure of quantum correlations. Similarly, the relative-entropy based discord $D^{(1)}_{\mathsf{A}}$ is a measure of quantum correlations. The property (iii) follows in this case from the fact that if ρ_Ψ is a pure state then a separable state σ_{sep} minimizing $S(\rho_\Psi||\sigma_{\mathrm{sep}})$ is the classical state given by (12.6) (see the proof of Proposition 12.1.2), so that $D^{(1)}_{\mathsf{A}}(\rho_\Psi)$ coincides with the entanglement measure $E_1(\rho_\Psi)$ defined in (12.3). It is an open problem to show that $D^{(\alpha)}_{\mathsf{A}}$ satisfies (iii) when $\alpha \neq 1/2, 1$.

The B-discords D_{B} and $D^{(\alpha)}_{\mathsf{B}}$ are defined by exchanging A and B in (12.21) and (12.22). As for the quantum discord of chapter 11, in general $D_{\mathsf{A}} \neq D_{\mathsf{B}}$. Symmetric measures of quantum correlations are obtained by considering the square distance to the set of classical states $\mathcal{C}_{\mathsf{AB}} = \mathcal{C}_{\mathsf{A}} \cap \mathcal{C}_{\mathsf{B}}$,

$$D_{\mathsf{AB}}(\rho) = 2\Big(1 - \max_{\sigma_{\mathrm{clas}} \in \mathcal{C}_{\mathsf{AB}}} \big\{\sqrt{F(\rho, \sigma_{\mathrm{clas}})}\big\}\Big) \quad , \quad D^{(\alpha)}_{\mathsf{AB}}(\rho) = \min_{\sigma_{\mathrm{clas}} \in \mathcal{C}_{\mathsf{AB}}} \big\{S(\rho||\sigma_{\mathrm{clas}})\big\} \; . \tag{12.24}$$

Let us mention that a similar symmetric information-based discord can be defined by modifying the maximization procedure in (11.12) so as to involve projectors $\pi^{\mathsf{A}}_i \otimes \pi^{\mathsf{B}}_i$ (or generalized measurement operators $M^{\mathsf{A}}_i \otimes M^{\mathsf{B}}_i$), instead of $M^{\mathsf{A}}_i \otimes 1$. It is called the *measurement-induced disturbance* [160]. The relative-entropy symmetric discord $D^{(1)}_{\mathsf{AB}}$ has been studied in [169], together with other quantities characterizing quantum and classical correlations. We will not elaborate further here on the numerous discord-like measures defined in the literature and their operational interpretations (see e.g. [170]).

We emphasize that since $\mathcal{C}_{\mathsf{AB}} \subset \mathcal{C}_{\mathsf{A}} \subset \mathcal{S}_{\mathsf{AB}}$ (see Fig. 11.1), the geometric measures are ordered as

$$E_{\mathrm{Bu}}(\rho) \leq D_{\mathsf{A}}(\rho) \leq D_{\mathsf{AB}}(\rho) \; . \tag{12.25}$$

This ordering is a nice feature of the geometrical approach. It also holds for the relative-entropy measures. In contrast, depending on ρ the entanglement of formation $E_{\mathrm{EoF}}(\rho)$ can be larger or smaller than the quantum discord $\delta_{\mathsf{A}}(\rho)$.

Before going on to general results, let us say few words about explicit calculations of the discords. In the special case of two-qubit states ρ with maximally mixed marginals $\rho_{\mathsf{A}} = \rho_{\mathsf{B}} = 1/2$, the relative-entropy measure $D^{(1)}_{\mathsf{AB}}(\rho)$ coincides with the usual discord $\delta^{\mathrm{v.N.}}_{\mathsf{A}}(\rho)$ [169, 167]. For the same states, a closed formula for $D_{\mathsf{A}}(\rho)$ has been found in [1, 227] and the closest A-classical states to ρ have been determined explicitly[3]. The Hilbert-Schmidt geometric discord is much easier to calculate. A simple formula for arbitrary 2-qubit states is derived in [67] and has been later on extended to higher dimensions. The geometric discord defined with the trace distance d_1 has been determined recently for certain families of two-qubit states (the so-called X-states, containing in particular the states with maximally mixed marginals, and the B-classical states) [63, 177]. Note that since d_1 is contractive, this geometric discord fulfills the axiom (ii) of Definition 12.2.1.

12.2.2 Geometric discord for pure states

We now proceed to determine the geometric discord D_{A} for pure states. It has been seen in the proof of Proposition 12.1.4 that the family of closest separable states of a pure state $|\Psi\rangle$ contains a pure product state, which is a classical state. By inspection of (12.11) and (12.25), one gets

$$D_{\mathsf{A}}(|\Psi\rangle) = D_{\mathsf{B}}(|\Psi\rangle) = D_{\mathsf{AB}}(|\Psi\rangle) = E_{\mathrm{Bu}}(|\Psi\rangle) = 2(1 - \sqrt{\mu_{\max}}) \; . \tag{12.26}$$

[3] This is done in [227] with the help of Corollary 12.2.6 below.

One deduces from the bound $\mu_{\max} \geq 1/n$ (which follows from $\sum_{i=1}^n \mu_i = 1$) that

$$D_{\mathsf{A}}(|\Psi\rangle) \leq 2\Big(1 - \frac{1}{\sqrt{n}}\Big) \quad , \quad n = \min\{n_{\mathsf{A}}, n_{\mathsf{B}}\} \; . \tag{12.27}$$

This bound is saturated when $\mu_i = 1/n$ for any i, that is, for the maximally entangled states. We will see below that this statement is still true for mixed states provided that $n_{\mathsf{A}} \leq n_{\mathsf{B}}$.

The identities (12.26) are analogous to the equality between the entanglement of formation $E_{\rm EoF}$ and the discord δ_{A} for pure states (Sec. 11.1). As said before, they reflect the existence of a pure product state which is closer or at the same distance from $|\Psi\rangle$ than any other separable state. It is of interest to find all the closest A-classical states to $|\Psi\rangle$. This is done in the next proposition.

Proposition 12.2.2. (Spehner and Orszag [226]) *Let $\rho_\Psi = |\Psi\rangle\langle\Psi|$ be a pure state of AB with largest Schmidt coefficient $\mu_{\max}$. If $\mu_{\max}$ is non-degenerate, then the closest A-classical (respectively classical, separable) state to ρ_Ψ for the Bures distance is unique. It is given by the pure product state $|\alpha_{\max}\rangle|\beta_{\max}\rangle$, where $|\alpha_{\max}\rangle$ and $|\beta_{\max}\rangle$ are eigenvectors with eigenvalue $\mu_{\max}$ of $[\rho_\Psi]_{\mathsf{A}}$ and $[\rho_\Psi]_{\mathsf{B}}$, respectively. If $\mu_{\max}$ is r-fold degenerate, say $\mu_{\max} = \mu_1 = \ldots = \mu_r > \mu_{r+1}, \ldots, \mu_n$, then infinitely many A-classical (respectively classical, separable) states σ minimize $d_{\rm B}(\rho_\Psi, \sigma)$. These closest states are convex combinations of the pure product states $|\varphi_l\rangle|\chi_l\rangle$ with*

$$|\varphi_l\rangle = \sum_{i=1}^r u_{il}|\alpha_i\rangle \quad , \quad |\chi_l\rangle = \sum_{i=1}^r \overline{u_{il}}|\beta_i\rangle \quad , \quad l = 1, \ldots, r \; , \tag{12.28}$$

where $\{|\alpha_i\rangle\}_{i=1}^r$ and $\{|\beta_i\rangle\}_{i=1}^r$ are orthonormal families of Schmidt vectors associated to $\mu_{\max}$ in the Schmidt decomposition (4.9), and $(u_{il})_{i,l=1}^r$ is an arbitrary $r \times r$ unitary matrix.

It should be noticed that when $\mu_{\max}$ is degenerate, the vectors (12.28) provide together with $|\alpha_i\rangle$, $|\beta_i\rangle$, $i = r+1, \ldots, n$, a Schmidt decomposition of $|\Psi\rangle$ (in that case this decomposition is not unique, see Sec. 4.2). Conversely, disregarding the degeneracies of the other eigenvalues $\mu_i < \mu_{\max}$, all Schmidt decompositions of $|\Psi\rangle$ are of this form for some unitary matrix $(u_{il})_{i,l=1}^r$. Thus, the existence of an infinite family of closest A-classical states to $|\Psi\rangle$ is related to the non-uniqueness of the Schmidt vectors associated to $\mu_{\max}$, and this family contains the products $|\varphi_l\rangle|\chi_l\rangle$ of these vectors and convex combinations thereof. This shows in particular that the maximally entangled pure states are the pure states with the largest family of closest states[4].

Proof. An arbitrary A-classical state σ can be decomposed as $\sigma = \sum_{ij} q_{ij}|\varphi_i\rangle\langle\varphi_i| \otimes |\chi_{j|i}\rangle\langle\chi_{j|i}|$. In much the same way as in the proof of Proposition 12.1.4, $F(|\Psi\rangle, \mathcal{C}_{\mathsf{A}}) = \mu_{\max}$ and the closest A-classical states to ρ fulfill

$$|\langle\varphi_i \otimes \chi_{j|i}|\Psi\rangle|^2 = \max_{\|\varphi\|=\|\chi\|=1} \big\{|\langle\varphi\otimes\chi|\Psi\rangle|^2\big\} = \mu_{\max} \quad \text{when} \;\; q_{ij} > 0. \tag{12.29}$$

We have thus to determine all $|\varphi\rangle \in \mathcal{H}_{\mathsf{A}}$ and $|\chi\rangle \in \mathcal{H}_{\mathsf{B}}$ such that $|\langle\varphi\otimes\chi|\Psi\rangle|^2 = \mu_{\max}$. This occurs if all inequalities in (12.15) are equalities. Let us first assume that $\mu_1 = \mu_{\max} > \mu_2, \ldots, \mu_n$. After a close look to (12.15) one immediately finds that $|\langle\varphi\otimes\chi|\Psi\rangle|^2 = \mu_{\max}$ if and only if $|\varphi\rangle = |\alpha_1\rangle$ and $|\chi\rangle = |\beta_1\rangle$ up to irrelevant phase factors. Hence (12.29) is satisfied for a single pair (i,j). Therefore, all the q_{ij} vanish except one and the closest A-classical state to $|\Psi\rangle$ is the pure product state $|\alpha_1\rangle|\beta_1\rangle$.

We now proceed to the degenerate case $\mu_1 = \ldots = \mu_r = \mu_{\max} > \mu_{r+1}, \ldots, \mu_n$. Let us establish the necessary and sufficient conditions for the inequalities in (12.15) to be equalities. For the first inequality, the condition is $\arg(\langle\varphi|\alpha_j\rangle\langle\chi|\beta_j\rangle) = \theta$ with θ independent of j. For the second one, the condition is that $|\varphi\rangle$ belongs to $V_{\max} = \mathrm{span}\{|\alpha_j\rangle\}_{j=1}^r$ or $|\chi\rangle$ belongs to $W_{\max} = \mathrm{span}\{|\beta_j\rangle\}_{j=1}^r$. The Cauchy-Schwarz inequality in (12.15) is saturated if and only if $|\langle\varphi|\alpha_j\rangle| = \lambda|\langle\chi|\beta_j\rangle|$ for all j, with $\lambda \geq 0$. Finally, the last inequality holds with equality if and only if $|\varphi\rangle \in \mathrm{span}\{|\alpha_j\rangle\}_{j=1}^n$ and $|\chi\rangle \in \mathrm{span}\{|\beta_j\rangle\}_{j=1}^n$. Putting all conditions together, we obtain $|\varphi\rangle \in V_{\max}$, $|\chi\rangle \in W_{\max}$, and $\langle\chi|\beta_j\rangle = e^{\mathrm{i}\theta}\langle\alpha_j|\varphi\rangle$ for $j = 1, \ldots, r$. Therefore, from any orthonormal family $\{|\varphi_l\rangle\}_{l=1}^r$ of $V_{\max}$ one can construct r orthogonal vectors $|\varphi_l \otimes \chi_l\rangle$ satisfying $|\langle\varphi_l\otimes\chi_l|\Psi\rangle|^2 = \mu_{\max}$ for all $l = 1, \ldots, r$, with $\langle\chi_l|\beta_j\rangle = \langle\alpha_j|\varphi_l\rangle$. The probabilities $\{q_{ij}\}$ are then given by $q_{ij} = q_i$ if $i = j \leq r$ and zero otherwise, $\{q_l\}_{l=1}^r$ being an arbitrary probability distribution. The corresponding A-classical states σ maximizing the fidelity $F(\rho_\Psi, \sigma)$ are the classical states

$$\sigma = \sum_{l=1}^r q_l |\alpha_l \otimes \beta_l\rangle\langle\alpha_l \otimes \beta_l| \; . \tag{12.30}$$

□

[4]This family forms a $(n^2 + n - 2)$ real-parameter sub-manifold of $\mathcal{E}(\mathcal{H}_{\mathsf{AB}})$.

12.2.3 Geometric discord for mixed states and quantum state discrimination

As for all other measures of entanglement and quantum correlations, determining $D_{\mathsf{A}}(\rho)$ is harder for mixed states than for pure states. Interestingly, this problem is related to an ambiguous quantum state discrimination task.

Proposition 12.2.3. (Spehner and Orszag [226]) *For any state ρ of the bipartite system AB, the maximal fidelity between ρ and an A-classical state reads*

$$F(\rho,\mathcal{C}_{\mathsf{A}}) = \max_{\{|\varphi_i\rangle\}}\Big\{P_{\mathrm{S,a}}^{\text{opt v.N.}}(\{\rho_i,\eta_i\})\Big\} = \max_{\{|\varphi_i\rangle\}}\max_{\{\Pi_i\}}\Big\{\sum_{i=1}^{n_{\mathsf{A}}}\eta_i\,\mathrm{tr}(\Pi_i\rho_i)\Big\}\,, \tag{12.31}$$

where the maxima are over all orthonormal bases $\{|\varphi_i\rangle\}_{i=1}^{n_{\mathsf{A}}}$ of $\mathcal{H}_{\mathsf{A}}$ and all von Neumann measurements given by orthonormal families $\{\Pi_i\}_{i=1}^{n_{\mathsf{A}}}$ of projectors of $\mathcal{H}_{\mathsf{AB}}$ with rank n_{B}. Here, $P_{\mathrm{S,a}}^{\text{opt v.N.}}(\{\rho_i,\eta_i\})$ is the maximal success probability in discriminating ambiguously by such measurements the states ρ_i with probabilities η_i defined by

$$\eta_i = \langle\varphi_i|\rho_{\mathsf{A}}|\varphi_i\rangle \quad , \quad \rho_i = \eta_i^{-1}\sqrt{\rho}|\varphi_i\rangle\langle\varphi_i|\otimes 1\sqrt{\rho} \quad , \quad i=1,\ldots,n_{\mathsf{A}} \tag{12.32}$$

(if $\eta_i = 0$ then ρ_i is not defined but does not contribute to the sum in (12.31)). Furthermore, the closest A-classical states to ρ are given by

$$\sigma_\rho = \frac{1}{F(\rho,\mathcal{C}_{\mathsf{A}})}\sum_{i=1}^{n_{\mathsf{A}}}|\varphi_i^{\text{opt}}\rangle\langle\varphi_i^{\text{opt}}|\otimes\langle\varphi_i^{\text{opt}}|\sqrt{\rho}\,\Pi_i^{\text{opt}}\sqrt{\rho}|\varphi_i^{\text{opt}}\rangle\,, \tag{12.33}$$

where $\{|\varphi_i^{\text{opt}}\rangle\}$ and $\{\Pi_i^{\text{opt}}\}$ are any orthonormal basis of $\mathcal{H}_{\mathsf{A}}$ and von Neumann measurement maximizing the right-hand side of (12.31).

The ρ_i are quantum states if $\eta_i > 0$ because $\rho_i \geq 0$ and η_i is chosen such that $\mathrm{tr}(\rho_i) = 1$. Moreover, $\{\eta_i\}_{i=1}^{n_{\mathsf{A}}}$ is a probability distribution (since $\eta_i \geq 0$ and $\sum_i \eta_i = \mathrm{tr}(\rho) = 1$) and the ensemble $\{\rho_i,\eta_i\}_{i=1}^{n_{\mathsf{A}}}$ is a convex decomposition of ρ, i.e., $\rho = \sum_i \eta_i\rho_i$.

Corollary 12.2.4. *If ρ is invertible then one can substitute $P_{\mathrm{S,a}}^{\text{opt v.N.}}(\{\rho_i,\eta_i\})$ in (12.31) by the maximal success probability $P_{\mathrm{S,a}}^{\text{opt}}(\{\rho_i,\eta_i\})$ over all POVMs, given by (6.2).*

Proof. This is a simple consequence of Corollary 6.4.4. Actually, if $\rho > 0$ then the states ρ_i defined in (12.32) are linearly independent, thus the optimal measurement to discriminate them is a von Neumann measurement with projectors of rank $r_i = \mathrm{rank}(\rho_i)$. The linear independence can be justified as follows. Let us first notice that ρ_i has rank $r_i = n_{\mathsf{B}}$ (for indeed, it has the same rank as $\eta_i\rho^{-1/2}\rho_i = |\varphi_i\rangle\langle\varphi_i|\otimes 1\sqrt{\rho}$). A necessary and sufficient condition for $|\xi_{ij}\rangle$ to be an eigenvector of ρ_i with eigenvalue $\lambda_{ij} > 0$ is $|\xi_{ij}\rangle = (\lambda_{ij}\eta_i)^{-1}\sqrt{\rho}|\varphi_i\rangle\otimes|\zeta_{ij}\rangle$, where $|\zeta_{ij}\rangle \in \mathcal{H}_{\mathsf{B}}$ is an eigenvector of $R_i = \langle\varphi_i|\rho|\varphi_i\rangle$ with eigenvalue $\lambda_{ij}\eta_i > 0$. For any i, the Hermitian invertible matrix R_i admits an orthonormal eigenbasis $\{|\zeta_{ij}\rangle\}_{j=1}^{n_{\mathsf{B}}}$. Thanks to the invertibility of $\sqrt{\rho}$, $\{|\xi_{ij}\rangle\}_{i=1,\ldots,n_{\mathsf{A}}}^{j=1,\ldots,n_{\mathsf{B}}}$ is a basis of $\mathcal{H}_{\mathsf{AB}}$ and thus the states ρ_i are linearly independent and span $\mathcal{H}_{\mathsf{AB}}$. □

Before going into the proof of the proposition, let us discuss the state discrimination problems when ρ is pure or A-classical. Of course, the values of $D_{\mathsf{A}}(\rho)$ are already known in these cases, being given by (12.26) and by $D_{\mathsf{A}}(\rho) = 0$, respectively, but it is instructive to recover that from Proposition 12.2.3. If $\rho = \rho_\Psi$ is pure then all states ρ_i with $\eta_i > 0$ are identical and equal to ρ_Ψ, so that $P_{\mathrm{S,a}}^{\text{opt v.N.}} = \max_{\{\Pi_i\}}\{\sum_i \eta_i\langle\Psi|\Pi_i|\Psi\rangle\} = \eta_{\max}$. One gets $F(\rho_\Psi,\mathcal{C}_{\mathsf{A}}) = \mu_{\max}$ by optimization over the basis $\{|\varphi_i\rangle\}$. If ρ is an A-classical state, i.e., if it can be decomposed as in (11.13), then the optimal basis $\{|\varphi_i^{\text{opt}}\rangle\}$ coincides with the basis appearing in this decomposition. With this choice one obtains $\eta_i = q_i$ and $\rho_i = |\varphi_i\rangle\langle\varphi_i|\otimes\sigma_{\mathsf{B}|i}$ for all i such that $q_i > 0$. The states ρ_i are orthogonal and can thus be perfectly discriminated by von Neumann measurements. This yields $F(\rho,\mathcal{C}_{\mathsf{A}}) = 1$ and $D_{\mathsf{A}}(\rho) = 0$ as it should be. Reciprocally, if $F(\rho,\mathcal{C}_{\mathsf{A}}) = 1$ then $P_{\mathrm{S,a}}^{\text{opt v.N.}}(\{\rho_i,\eta_i\}) = 1$ for some basis $\{|\varphi_i\rangle\}$ and the corresponding ρ_i must be orthogonal (chapter 6). Hence one can find an orthonormal family $\{\Pi_i\}$ of projectors with rank n_{B} such that $\rho_i = \Pi_i\rho_i\Pi_i$ for any i with $\eta_i > 0$. It is an easy exercise to show that this implies that $\Pi_i = |\varphi_i\rangle\langle\varphi_i|\otimes 1$ if $\rho|_{\Pi_i\mathcal{H}}$ is invertible. Thus $\rho = \sum_i \eta_i\rho_i$ is A-classical, in agreement with the fact (following directly from the definition) that $D_{\mathsf{A}}(\rho) = 0$ if and only if ρ is A-classical.

The above discussion provides a clear interpretation of the result of Proposition 12.2.3: the states ρ with non-zero discord are characterized by ensembles $\{\rho_i,\eta_i\}$ of non-orthogonal states, which thereby are not perfectly

distinguishable, for any orthonormal basis $\{|\varphi_i\rangle\}$ of $\mathcal{H}_{\mathsf{A}}$. The less distinguishable are the ρ_i's, the most distant is ρ from the set of zero-discord states.

We will establish Proposition 12.2.3 by relying on the slightly more general statement summarized in the following lemma.

Lemma 12.2.5. *For a fixed family $\{\sigma_{\mathsf{A}|i}\}_{i=1}^n$ of states $\sigma_{\mathsf{A}|i} \in \mathcal{E}(\mathcal{H}_{\mathsf{A}})$ having orthogonal supports and spanning $\mathcal{H}_{\mathsf{A}}$, with $1 \leq n \leq n_{\mathsf{A}}$, let us define*

$$\mathcal{C}_{\mathsf{A}}(\{\sigma_{\mathsf{A}|i}\}) = \Big\{\sigma = \sum_{i=1}^n q_i \sigma_{\mathsf{A}|i} \otimes \sigma_{\mathsf{B}|i} \; ; \; \{q_i, \sigma_{\mathsf{B}|i}\}_{i=1}^n \;\; \textit{is a state ensemble on } \mathcal{H}_{\mathsf{B}} \Big\} . \tag{12.34}$$

Then

$$F\big(\rho, \mathcal{C}_{\mathsf{A}}(\{\sigma_{\mathsf{A}|i}\})\big) = \max_{\sigma \in \mathcal{C}_{\mathsf{A}}(\{\sigma_{\mathsf{A}|i}\})} \big\{F(\rho,\sigma)\big\} = \max_U \Big\{ \sum_{i=1}^n \|W_i(U)\|_2^2 \Big\} , \tag{12.35}$$

where the last maximum is over all unitaries U on $\mathcal{H}_{\mathsf{AB}}$ and

$$W_i(U) = \mathrm{tr}_{\mathsf{A}}\big(\sqrt{\sigma_{\mathsf{A}|i}} \otimes 1 \, \sqrt{\rho}\, U\big) . \tag{12.36}$$

Moreover, there exists a unitary $U_{\rm opt}$ achieving the maximum in (12.35) which is such that $W_i(U_{\rm opt}) \geq 0$. The states $\sigma_{\rm opt}$ satisfying $F(\rho, \sigma_{\rm opt}) = F(\rho, \mathcal{C}_{\mathsf{A}}(\{\sigma_{\mathsf{A}|i}\}))$ are given in terms of this unitary by

$$\sigma_{\rm opt} = \frac{1}{F(\rho, \mathcal{C}_{\mathsf{A}}(\{\sigma_{\mathsf{A}|i}\}))} \sum_{i=1}^n \sigma_{\mathsf{A}|i} \otimes W_i(U_{\rm opt})^2 . \tag{12.37}$$

Proof. Using the spectral decompositions of the states $\sigma_{\mathsf{B}|i}$, any $\sigma \in \mathcal{C}_{\mathsf{A}}(\{\sigma_{\mathsf{A}|i}\})$ can be written as

$$\sigma = \sum_{i=1}^n \sum_{j=1}^{n_{\mathsf{B}}} q_{ij} \sigma_{\mathsf{A}|i} \otimes |\chi_{j|i}\rangle\langle\chi_{j|i}| \quad \text{with} \quad q_{ij} \geq 0 \, , \, \sum_{ij} q_{ij} = 1 \, , \tag{12.38}$$

where $\{|\chi_{j|i}\rangle\}_{j=1}^{n_{\mathsf{B}}}$ is an orthonormal basis of $\mathcal{H}_{\mathsf{B}}$ for any i (compare with (11.20)). By assumption, if $i \neq i'$ then $\mathrm{ran}\, \sigma_{\mathsf{A}|i} \perp \mathrm{ran}\, \sigma_{\mathsf{A}|i'}$, so that $\sqrt{\sigma} = \sum_{i,j} \sqrt{q_{ij}} \sqrt{\sigma_{\mathsf{A}|i}} \otimes |\chi_{j|i}\rangle\langle\chi_{j|i}|$. We start by evaluating the trace norm in the definition (8.9) of the fidelity by means of the formula $\|O\|_1 = \max_U |\,\mathrm{tr}(UO)|$ to obtain

$$\begin{aligned} F\big(\rho, \mathcal{C}_{\mathsf{A}}(\{\sigma_{\mathsf{A}|i}\})\big) &= \max_{\sigma \in \mathcal{C}_{\mathsf{A}}(\{\sigma_{\mathsf{A}|i}\})} \max_U \Big\{ \big|\mathrm{tr}(U^* \sqrt{\rho}\sqrt{\sigma})\big|^2 \Big\} \\ &= \max_U \Big\{ \max_{\{q_{ij}\},\{|\chi_{j|i}\rangle\}} \Big| \sum_{i,j} \sqrt{q_{ij}} \langle \chi_{j|i} | W_i(U)^* | \chi_{j|i} \rangle \Big|^2 \Big\} . \end{aligned} \tag{12.39}$$

The square modulus can be bounded by invoking twice the Cauchy-Schwarz inequality and $\sum_{ij} q_{ij} = 1$,

$$\begin{aligned} \Big| \sum_{i,j} \sqrt{q_{ij}} \langle \chi_{j|i} | W_i(U)^* | \chi_{j|i} \rangle \Big|^2 &\leq \sum_{i,j} \big| \langle \chi_{j|i} | W_i(U)^* | \chi_{j|i} \rangle \big|^2 \\ &\leq \sum_{i,j} \big\| W_i(U) | \chi_{j|i} \rangle \big\|^2 = \sum_i \|W_i(U)\|_2^2 . \end{aligned} \tag{12.40}$$

The foregoing inequalities are equalities if the following conditions are satisfied:

(1) $W_i(U) = W_i(U)^* \geq 0$;

(2) $q_{ij} = \langle \chi_{j|i} | W_i(U) | \chi_{j|i} \rangle^2 / (\sum_{i,j} \langle \chi_{j|i} | W_i(U) | \chi_{j|i} \rangle^2)$;

(3) $\{|\chi_{j|i}\rangle\}_{j=1}^{n_{\mathsf{B}}}$ is an eigenbasis of $W_i(U)$ for any i.

Therefore, (12.35) holds true provided that there is a unitary U on $\mathcal{H}_{\mathsf{AB}}$ satisfying (1). For a given U, let us define $U_{\rm opt} = U\sum_i \pi_i^{\mathsf{A}} \otimes V_i^*$, where π_i^{A} is the projector onto $\operatorname{ran}\sigma_{\mathsf{A}|i}$ and V_i a unitary on $\mathcal{H}_{\mathsf{B}}$ such that $W_i(U) = |W_i(U)^*|V_i$ (polar decomposition). Then $U_{\rm opt}$ is unitary since by hypothesis $\pi_i^{\mathsf{A}}\pi_{i'}^{\mathsf{A}} = \delta_{ii'}\pi_i^{\mathsf{A}}$ and $\sum_i \pi_i^{\mathsf{A}} = 1$, and one readily shows that $W_i(U_{\rm opt}) = W_i(U)V_i^* \geq 0$. As $\sum_i \|W_i(U)\|_2^2 = \sum_i \|W_i(U_{\rm opt})\|_2^2$, the identity (12.35) follows from (12.39) and (12.40). From condition (3) one has $W_i(U_{\rm opt})|\chi_{j|i}^{\rm opt}\rangle = w_{ji}|\chi_{j|i}^{\rm opt}\rangle$ with $\sum_{i,j} w_{ji}^2 = F(\rho, \mathcal{C}_{\mathsf{A}}(\{\sigma_{\mathsf{A}|i}\}))$, see (12.40). Condition (2) entails

$$\sigma_{\mathsf{B}|i}^{\rm opt} = \sum_j q_{ij}^{\rm opt}|\chi_{j|i}^{\rm opt}\rangle\langle\chi_{j|i}^{\rm opt}| = \frac{W_i(U_{\rm opt})^2}{F(\rho, \mathcal{C}_{\mathsf{A}}(\{\sigma_{\mathsf{A}|i}\}))} \,, \tag{12.41}$$

which together with (12.38) leads to (12.37). □

Proof of Proposition 12.2.3. Let $\{|\varphi_i\rangle\}_{i=1}^{n_{\mathsf{A}}}$ be an orthonormal basis of $\mathcal{H}_{\mathsf{A}}$. Applying Lemma 12.2.5 with $\sigma_{\mathsf{A}|i} = |\varphi_i\rangle\langle\varphi_i|$ one gets

$$\begin{aligned} F\big(\rho, \mathcal{C}_{\mathsf{A}}(\{|\varphi_i\rangle\})\big) &= \max_U \left\{ \sum_{i=1}^{n_{\mathsf{A}}} \operatorname{tr}\big[U|\varphi_i\rangle\langle\varphi_i| \otimes 1\, U^* \sqrt{\rho}\, |\varphi_i\rangle\langle\varphi_i| \otimes 1 \sqrt{\rho}\big] \right\} \,, \\ &= \max_{\{\Pi_i\}} \left\{ \sum_{i=1}^{n_{\mathsf{A}}} \operatorname{tr}\big[\Pi_i \sqrt{\rho}|\varphi_i\rangle\langle\varphi_i| \otimes 1 \sqrt{\rho}\big] \right\} = P_{\rm S,a}^{\text{opt v.N.}}(\{\rho_i, \eta_i\}) \,. \end{aligned} \tag{12.42}$$

The last maximum is over all orthonormal families $\{\Pi_i\}_{i=1}^{n_{\mathsf{A}}}$ of projectors of rank n_{B} and the success probability $P_{\rm S,a}^{\text{opt v.N.}}(\{\rho_i, \eta_i\})$ is given by (6.32). Since the fidelity $F(\rho, \mathcal{C}_{\mathsf{A}})$ is the maximum of $F(\rho, \mathcal{C}_{\mathsf{A}}(\{|\varphi_i\rangle\}))$ over all bases $\{|\varphi_i\rangle\}$, this leads to (12.31) and (12.33). □

12.2.4 The qubit case

It has been emphasized in chapter 6 that the optimal success probability and measurement for discriminating ambiguously more than two states are not known explicitly in general. Nonetheless, if the subsystem A is a qubit, the ensemble $\{\rho_i, \eta_i\}$ in Proposition 12.2.3 contains only $n_{\mathsf{A}} = 2$ states and the optimal probability and measurement are easily determined. Following the steps yielding to (6.6) we find

$$P_{\rm S,a}^{\text{opt v.N.}}(\{\rho_i, \eta_i\}) = \frac{1}{2}\big(1 - \operatorname{tr}\Lambda\big) + \sum_{l=1}^{n_{\mathsf{B}}} \lambda_l \,, \tag{12.43}$$

where $\lambda_1 \geq \cdots \geq \lambda_{n_{\mathsf{B}}}$ are the n_{B} largest eigenvalues of $\Lambda = \eta_0\rho_0 - \eta_1\rho_1$. The optimal von Neumann measurement is formed by the spectral projector $\Pi_0^{\rm opt}$ of Λ for these n_{B} eigenvalues and its complement $\Pi_1^{\rm opt} = 1 - \Pi_0^{\rm opt}$. For the states ρ_i associated to the orthonormal basis $\{|\varphi_i\rangle\}_{i=0}^1$ of $\mathbb{C}^2$ via formula (12.32), one has $\Lambda = \sqrt{\rho}\,(|\varphi_0\rangle\langle\varphi_0| - |\varphi_1\rangle\langle\varphi_1|) \otimes 1\sqrt{\rho}$. The operator inside the parenthesis in the last identity is equal to $\sigma_{\mathbf{u}} \equiv \sum_{m=1}^3 u_m\sigma_m$ for some unit vector $\mathbf{u} \in \mathbb{R}^3$ depending on $\{|\varphi_i\rangle\}$ (here σ_1, σ_2, and σ_3 are the Pauli matrices). Conversely, one can associate to any unit vector $\mathbf{u} \in \mathbb{R}^3$ the eigenbasis $\{|\varphi_i\rangle\}_{i=0}^1$ of $\sigma_{\mathbf{u}}$. According to Proposition 12.2.3, $F(\rho, \mathcal{C}_{\mathsf{A}})$ is obtained by maximizing the right-hand side of (12.43) over all Hermitian matrices

$$\Lambda(\mathbf{u}) = \sqrt{\rho}\,\sigma_{\mathbf{u}} \otimes 1\sqrt{\rho} \tag{12.44}$$

with $\mathbf{u} \in \mathbb{R}^3$, $|\mathbf{u}| = 1$. The following corollary of Proposition 12.2.3 is a refinement of a result in [227].

Corollary 12.2.6. *Let A be a qubit, i.e., $n_{\mathsf{A}} = 2$. The fidelity between ρ and the set of A-classical states is given by*

$$F(\rho, \mathcal{C}_{\mathsf{A}}) = \frac{1}{2} \max_{\|\mathbf{u}\|=1} \big\{1 + \|\Lambda(\mathbf{u})\|_1\big\} \,, \tag{12.45}$$

where $\Lambda(\mathbf{u})$ is the $2n_B \times 2n_B$ matrix (12.44). The closest A-classical states to ρ are given by (12.33) where $\Pi_0^{\rm opt}$ is the spectral projector associated to the n_{B} largest eigenvalues of $\Lambda(\mathbf{u}^{\rm opt})$ and $\mathbf{u}^{\rm opt} \in \mathbb{R}^3$ is a unit vector achieving the maximum in (12.45).

Proof. Let $\lambda_l(\mathbf{u})$ be the eigenvalues of $\Lambda(\mathbf{u})$ in non-increasing order. We claim that

$$-\frac{1}{2}\operatorname{tr}(\Lambda(\mathbf{u})) + \sum_{l=1}^{n_{\mathsf{B}}} \lambda_l(\mathbf{u}) = \frac{1}{2}\sum_{l=1}^{n_{\mathsf{B}}} \lambda_l(\mathbf{u}) - \frac{1}{2}\sum_{l=n_{\mathsf{B}}+1}^{2n_{\mathsf{B}}} \lambda_l(\mathbf{u}) = \frac{1}{2}\operatorname{tr}|\Lambda(\mathbf{u})| \,. \tag{12.46}$$

To prove this claim it suffices to show that $\Lambda(\mathbf{u})$ has at most n_{B} positive eigenvalues $\lambda_l(\mathbf{u}) > 0$ and at most n_{B} negative eigenvalues $\lambda_l(\mathbf{u}) < 0$, counting multiplicities. As $\ker\rho \subset \ker\Lambda(\mathbf{u})$ one may without loss of generality restrict $\Lambda(\mathbf{u})$ to the subspace $\Pi\mathcal{H}_{\mathsf{AB}}$, with Π the projector onto $\operatorname{ran}(\rho)$. A standard linear algebra argument shows that if S is a finite invertible matrix and Σ a self-adjoint matrix, then the number of positive (respectively negative) eigenvalues of Σ is equal to the number of positive (respectively negative) eigenvalues of $S^*\Sigma S$. Let $P_\Sigma^\pm$ be the spectral projectors of $\Sigma = \Pi\,\sigma_{\mathbf{u}} \otimes 1\,\Pi$ on $\mathbb{R}_\pm \setminus \{0\}$. Since $\sqrt{\rho} : \Pi\mathcal{H}_{\mathsf{AB}} \to \Pi\mathcal{H}_{\mathsf{AB}}$ is invertible, in order to prove (12.46) it is thus enough to verify that $\operatorname{rank}(P_\Sigma^\pm) \leq n_{\mathsf{B}}$. This is evident if $\operatorname{rank}(\Pi) \leq n_{\mathsf{B}}$. If $\operatorname{rank}(\Pi) > n_{\mathsf{B}}$, then $\pm\langle\Psi|\sigma_{\mathbf{u}} \otimes 1|\Psi\rangle = \pm\langle\Psi|\Sigma|\Psi\rangle > 0$ for any $|\Psi\rangle \in P_\Sigma^\pm\mathcal{H}_{\mathsf{AB}} \subset \Pi\mathcal{H}_{\mathsf{AB}}$. This implies that $\operatorname{rank}(P_\Sigma^\pm) \leq \operatorname{rank}(P_{\sigma_{\mathbf{u}}\otimes 1}^\pm) = n_{\mathsf{B}}$, as otherwise one could find a non-vanishing vector $|\Psi\rangle \in P_\Sigma^\pm\mathcal{H}_{\mathsf{AB}}$ belonging to the n_{B}-dimensional eigenspace of $\sigma_{\mathbf{u}} \otimes 1$ with eigenvalue ∓ 1, in contraction with the foregoing inequality. This establishes (12.46). Then (12.45) follows from (12.43) and Proposition 12.2.3. □

12.2.5 States with the highest geometric discord

The geometric discord D_{A}, as the quantum discord δ_{A}, quantifies the degree of quantumness of a state. Let us recall from Sec. 11.3.2 that when the space dimensions of A and B are such that $n_{\mathsf{A}} \leq n_{\mathsf{B}}$, the "most quantum" states ρ having the highest discord $\delta_{\mathsf{A}}(\rho)$ are the maximally entangled states, i.e., the states with the highest entanglement of formation $E_{\mathrm{EoF}}(\rho) = \ln n_{\mathsf{A}}$. It is comforting that a similar result holds for the geometric discord.

Corollary 12.2.7. *If $n_{\mathsf{A}} \leq n_{\mathsf{B}}$, the highest value of $D_{\mathsf{A}}(\rho)$ on $\mathcal{E}(\mathcal{H}_{\mathsf{AB}})$ is equal to $2 - 2/\sqrt{n_{\mathsf{A}}}$. The most distant states ρ from the set of A-classical states, which are such that $D_{\mathsf{A}}(\rho) = 2 - 2/\sqrt{n_{\mathsf{A}}}$, are the maximally entangled states given by Proposition 10.5.1.*

Comparing with the results of Sec. 12.1.2, we see that when $n_{\mathsf{A}} \leq n_{\mathsf{B}}$ the most distant states from $\mathcal{C}_{\mathsf{A}}$ are also the most distant from the set of separable states $\mathcal{S}_{\mathsf{AB}}$. If $n_{\mathsf{A}} \leq n_{\mathsf{B}} < 2n_{\mathsf{A}}$, these most distant states are maximally entangled pure states, as illustrated in Fig. 11.1.

Proof. This is again a corollary of Proposition 12.2.3. The success probability $P_{\mathrm{S,a}}^{\text{opt v.N.}}$ is clearly larger or equal to the highest prior probability[5] $\eta_{\max} = \max_i\{\eta_i\}$. In view of Proposition 12.2.3 and $\eta_{\max} \geq 1/n_{\mathsf{A}}$, we get

$$F(\rho, \mathcal{C}_{\mathsf{A}}) \geq \frac{1}{n_{\mathsf{A}}} \tag{12.47}$$

for any state ρ. When $n = n_{\mathsf{A}} \leq n_{\mathsf{B}}$ this bound is optimal, the value $1/n$ being achieved for the maximally entangled pure states (Sec. 12.2.2). This proves the first statement. Let ρ be a state such that $F(\rho, \mathcal{C}_{\mathsf{A}}) = 1/n$. According to (12.31) and since it has been argued above that $P_{\mathrm{S,a}}^{\text{opt v.N.}} \geq \eta_{\max} \geq 1/n$, this implies that $P_{\mathrm{S,a}}^{\text{opt v.N.}}(\{\rho_i, \eta_i\}) = 1/n$ whatever the orthonormal basis $\{|\varphi_i\rangle\}$. It is intuitively clear[6] that this can happen only if the receiver gets a collection of identical states ρ_i with equal prior probabilities $\eta_i = 1/n$. From (12.32) and $\rho = \sum \eta_i\rho_i$ one obtains $\rho_{\mathsf{A}} = 1/n$ and $\rho_i = \rho$ for any i and $\{|\varphi_i\rangle\}$. Plugging the spectral decomposition $\rho = \sum_k p_k|k\rangle\langle k|$ into (12.32), the second equality yields $D_{kl} = \operatorname{tr}_{\mathsf{B}}(|k\rangle\langle l|) = n^{-1}\delta_{kl}$ for all k and l such that $p_k p_l \neq 0$. One concludes that ρ has maximal entanglement of formation by following the same steps as in the proof of Proposition 10.5.1. □

One may wonder if Corollary 12.2.7 could also hold for $n_{\mathsf{A}} > n_{\mathsf{B}}$ (modulo the exchange $n_{\mathsf{A}} \leftrightarrow n_{\mathsf{B}}$), as what happens for the geometric measure of entanglement (see Sec. 12.1.2). However, unlike $E_{\mathrm{Bu}}(\rho)$ the geometric discord is not symmetric under the exchange of the two subsystems. The problem of determining its highest value and the corresponding "most quantum" states is still open for $n_{\mathsf{A}} > n_{\mathsf{B}}$. For such space dimensions the bound (12.47) is still correct but it is not optimal, that is, there are no states ρ with fidelity $F(\rho, \mathcal{C}_{\mathsf{A}}) = 1/n_{\mathsf{A}}$.

[5] A receiver would obtain $P_{\mathrm{S,a}} = \eta_{\max}$ by simply guessing that his state is $\rho_{i_{\max}}$, with $\eta_{i_{\max}} = \eta_{\max}$, whatever the measurement outcomes. A better strategy is of course to perform the von Neumann measurement $\{\Pi_i\}$ such that $\Pi_{i_{\max}}$ projects onto a n_{B}-dimensional subspace containing $\operatorname{ran}(\rho_{i_{\max}})$. This range has a dimension $\operatorname{rank}(\rho_{i_{\max}}) \leq n_{\mathsf{B}}$ by a similar reasoning as in the proof of Corollary 12.2.4.

[6] An explicit proof of this fact can be found in [226].

	Entanglement of formation	Quantum discord	Geometric entanglement	Geometric discord
AB in a pure state	$E_{\rm EoF}(\|\Psi\rangle) = \delta_{\mathsf{A}}(\|\Psi\rangle) = H(\{\mu_i\})$		$E_{\rm Bu}(\|\Psi\rangle) = D_{\mathsf{A}}(\|\Psi\rangle) = 2(1-\sqrt{\mu_{\max}})$	
AB in a mixed state	$E_{\rm EoF}(\rho) = \min\{\sum_i \eta_i E_{\rm EoF}(\|\Psi_i\rangle)\}$ (convex roof)	$\delta_{\mathsf{A}}(\rho) = I_{\mathsf{A}:\mathsf{B}}(\rho) - \underbrace{\max\{I_{\mathsf{A}:\mathsf{B}}(\mathcal{M}_{\mathsf{A}}\otimes 1(\rho))\}}_{\text{classical correlations}}$	$E_{\rm Bu}(\rho) = 2(1-\underbrace{\max\{\sqrt{F(\rho,\sigma_{\rm sep})}\}}_{=\text{ convex roof}})$	$D_{\mathsf{A}}(\rho) = 2(1-\underbrace{\max\{\sqrt{F(\rho,\sigma_{\mathsf{A}-\rm cl})}\}}_{=\text{ max. success proba. in state discrimination}})$
Vanishes iff	ρ is separable	ρ is A-classical	ρ is separable	ρ is A-classical
Maximal iff with maximal value	ρ is max. entangled $\ln n$	$E_{\rm EoF}$: true $\forall\ n_{\mathsf{A},\mathsf{B}}$; δ_{A}: true if $n_{\mathsf{A}} \le n_{\mathsf{B}}$	ρ is max. entangled $2(1-1/\sqrt{n})$	$E_{\rm Bu}$: true $\forall\ n_{\mathsf{A},\mathsf{B}}$; D_{A}: true if $n_{\mathsf{A}} \le n_{\mathsf{B}}$
Local unit. invariance	✓	✓	✓	✓
Monotonicity w.r.t.	LOCCs	operations on B	LOCCs	operations on B
Convexity	✓	no	✓	no
Ordering	no		$E_{\rm Bu}(\rho) \le D_{\mathsf{A}}(\rho)$	
ABC in a pure state	$E_{\rm EoF}(\rho_{\mathsf{BC}}) = \delta_{\mathsf{A}}(\rho_{\mathsf{AB}}) + S(\rho_{\mathsf{AB}}) - S(\rho_{\mathsf{A}})$		?	

Table 12.1: Summary of the definitions and properties of the entanglement of formation (chapter 10), quantum discord (chapter 11), geometric measure of entanglement (Sec. 12.1), and geometric discord (Sec. 12.2). Here n_{A} and n_{B} are the space dimensions of the subsystems A and B, $n = \min\{n_{\mathsf{A}}, n_{\mathsf{B}}\}$, and μ_i are the Schmidt coefficients in (4.9).

Indeed, one can show as in the proof above that if $F(\rho, \mathcal{C}_{\mathsf{A}}) = 1/n_{\mathsf{A}}$ then the eigenvectors $|k\rangle$ of ρ with eigenvalues $p_k > 0$ have maximally mixed marginals $D_{kk} = (|k\rangle\langle k|)_{\mathsf{A}} = 1/n_{\mathsf{A}}$. But this is impossible since $\mathrm{rank}(D_{kk}) \le n_{\mathsf{B}}$ by (4.9).

Remark 12.2.8. *One can place a lower bound on $F(\rho, \mathcal{C}_{\mathsf{A}})$ for $n_{\mathsf{A}} > n_{\mathsf{B}}$ by invoking the inequality [226]*

$$F(\rho, \mathcal{C}_{\mathsf{A}}) \ge \frac{\|\rho\|}{n_{\mathsf{B}}} + \frac{1 - \|\rho\|}{n_{\mathsf{A}}} \frac{n_{\mathsf{B}} - \delta_\rho}{n_{\mathsf{B}}} \tag{12.48}$$

where $\delta_\rho = 0$ if $\mathrm{rank}(\rho) \le n_{\mathsf{B}}$ and 1 otherwise.

Table 12.1 presents a comparison of the properties of the entanglement of formation, the quantum discord, and their geometrical analogs based on the Bures distance.

12.2.6 Geometric discord and least square measurements

The ensemble $\{\rho_i, \eta_i\}$ in the discrimination task associated to the geometric discord in Proposition 12.2.3 turns out to be related to the transpose operation of the von Neumann measurement in the basis $\{|\varphi_i\rangle\}$. In fact, let us denote by $\mathcal{M}_{\mathsf{A}}$ the measurement on A with rank-one orthonormal projectors $\pi_i^{\mathsf{A}} = |\varphi_i\rangle\langle\varphi_i|$. Let

$$\eta_i = \langle\varphi_i|\rho_{\mathsf{A}}|\varphi_i\rangle \quad , \quad \rho_{\mathsf{AB}|i} = \eta_i^{-1} |\varphi_i\rangle\langle\varphi_i| \otimes \langle\varphi_i|\rho|\varphi_i\rangle \tag{12.49}$$

be the corresponding probabilities and post-measurement conditional states when the initial state is ρ. The transpose operation of $\mathcal{M}_{\mathsf{A}}$ for ρ is (see (5.24))

$$\mathcal{R}_{\mathcal{M}_{\mathsf{A}},\rho}(\sigma) = \sum_{i=1}^{n_{\mathsf{A}}} \sqrt{\rho}|\varphi_i\rangle\langle\varphi_i| \otimes \langle\varphi_i|\rho|\varphi_i\rangle^{-\frac{1}{2}} \langle\varphi_i|\sigma|\varphi_i\rangle \langle\varphi_i|\rho|\varphi_i\rangle^{-\frac{1}{2}} \sqrt{\rho}\,. \tag{12.50}$$

We observe that

$$\rho_i = \mathcal{R}_{\mathcal{M}_{\mathsf{A}},\rho}(\rho_{\mathsf{AB}|i}) \quad , \quad i = 1, \ldots, n_{\mathsf{A}}\,. \tag{12.51}$$

Comparing (5.23) and (12.51), one expects from the discussion in Sec. 5.5.3 that the least square measurement $\{M_i^{\rm lsm}\}$ for the state ensemble $\{\rho_i, \eta_i\}$ is associated to the transpose operation of $\mathcal{R}_{\mathcal{M}_{\mathsf{A}},\rho}$ for $\mathcal{M}_{\mathsf{A}}(\rho) = \sum_i \eta_i \rho_{\mathsf{AB}|i}$. But this two-fold transpose operation coincides with $\mathcal{M}_{\mathsf{A}}$, hence $\{M_i^{\rm lsm}\}$ is nothing but the von Neumann measurement on A in the basis $\{|\varphi_i\rangle\}$. This can be readily checked: since $\{\rho_i, \eta_i\}$ is a convex decomposition of ρ, (5.36) leads to

$$M_i^{\rm lsm} = \eta_i \rho^{-1/2} \rho_i \rho^{-1/2} = \pi_i^{\mathsf{A}} \otimes 1\,. \tag{12.52}$$

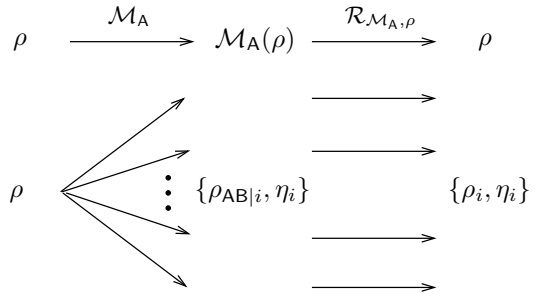

Figure 12.1: State changes under the von Neumann measurement $\mathcal{M}_{\mathsf{A}}$ with rank-one projectors $\pi_i^{\mathsf{A}} = |\varphi_i\rangle\langle\varphi_i|$ followed by its transpose operation $\mathcal{R}_{\mathcal{M}_{\mathsf{A}},\rho}$. The upper line corresponds to a measurement without readout and the other lines to the different measurement outcomes.

One can bound $P_{\mathrm{S,a}}^{\text{opt v.N.}}(\{\rho_i, \eta_i\})$ from below by the success probability obtained by discriminating the ρ_i with $\{M_i^{\mathrm{lsm}}\}$, and from above by the square root of this probability, see (6.21). By Proposition 12.2.3, this yields

$$\max_{\{|\varphi_i\rangle\}} \left\{ \sum_{i=1}^{n_{\mathsf{A}}} \mathrm{tr}_{\mathsf{B}} \big[\langle\varphi_i|\sqrt{\rho}|\varphi_i\rangle^2 \big] \right\} \leq F(\rho, \mathcal{C}_{\mathsf{A}}) \leq \max_{\{|\varphi_i\rangle\}} \left\{ \sum_{i=1}^{n_{\mathsf{A}}} \mathrm{tr}_{\mathsf{B}} \big[\langle\varphi_i|\sqrt{\rho}|\varphi_i\rangle^2 \big] \right\}^{\frac{1}{2}} . \tag{12.53}$$

The left- and right-hand sides become nearly equal when $F(\rho, \mathcal{C}_{\mathsf{A}})$ is almost one, that is, if ρ is close to $\mathcal{C}_{\mathsf{A}}$. Other inequalities on $F(\rho, \mathcal{C}_{\mathsf{A}})$ can be obtained in terms of the fidelities $F(\rho_i, \rho_j)$ with the help of Proposition 6.5.1.

The aforementioned observations are summarized by Fig. 12.1.

Part III

Perspectives

PERSPECTIVES FROM PART I: DYNAMICS

La question ne se pose pas. Elle en est absolument incapable : il y a trop de vent (B. Vian).

1) Quantum measurement processes

A deeper understanding of the implications of the recent results of Allahverdyan, Balian, and Nieuwenhuizen [6] discussed in chapter 1 is desirable. Revisiting the principles of quantum mechanics is fascinating and I plan to keep thinking in the next years about these fundamental issues, in parallel with solving more concrete problems. The model presented in chapter 1 describes ideal von Neumann projective measurements. Non-ideal measurements are also discussed in [6]. In quantum information theory, one is interested in extracting as much information as possible from a system with some given experimental constraints. The work presented in chapter 1 could be pursued in the following directions:

(a) **Weak measurements.**
Weak measurements are measurements in which the system weakly interacts with the measuring apparatus, so that the eigenvalues of the measured observable are not fully correlated with the pointer values at the end of the measurement. This means that the system-apparatus interaction time $t_{\rm int}$ is smaller than the system-apparatus entanglement time $\tau_{\rm ent}$ (see Sec. 1.3.4), as opposed to the strong measurement case considered in chapter 1. Weak measurements combined with an appropriate post-selection of the measurement outcomes may be used to amplify small signals thanks to the weak value amplification protocol of Y. Aharonov, D.Z. Albert, and L. Vaidman [Phys. Rev. Lett. **60**, 1351 (1988)]. It would be of interest to change the model of Sec. 1.3 in such a way that the apparatus performs a weak measurement, and to determine the corresponding time scales. A related issue, which is also connected to chapter 3, is the use of weak measurements in parameter estimation (such as quantum interferometry). The practical advantage of such measurements has been pointed out, and a protocol to saturate the quantum Cramér-Rao bound (Sec. 9.2.2) with weak measurements has been proposed recently [G. Bié Alves *et al.*, *Weak value amplification as an optimal metrological protocol*, arXiv:1410.7415].

(b) **Time-evolution of correlations during a measurement process.**
It would be interesting to determine the time evolution of entanglement and quantum discord between the system and pointer during the measurement process of Sec. 1.3. As far as classical correlations are concerned, Cornelio *et al.* argue that they are either time-independent or reach a constant value after a finite time [Phys. Rev. Lett. **109**, 190402 (2012)].

2) Evolution of entanglement for quantum trajectories

The results presented in chapter 2 on the average entanglement of quantum trajectories is limited to two-qubit systems and Markovian dynamics. Relevant extensions of this work are thus:

(a) **Entanglement decay in bipartite systems with space dimensions higher than four.**
In the case of two subsystems coupled to independent baths with local jump operators, the general argument based on the no-jump trajectory (Sec. 2.4.1) tells us that the average entanglement can only vanish asymptotically. It would still be nice to find how fast this entanglement decays for higher-dimensional subsystems than qubits. For good choices of the entanglement measure (see the discussion in Sec. 2.4.2), a strategy would be to try to bound this measure by an exponential in time and to calculate the corresponding decay rate κ.

(b) **Non-Markovian quantum trajectories.**
It is natural to ask whether our results with S. Vogelsberger extend to the non-Markovian case. Non-Markovian quantum trajectories have been studied in Refs. [47, 80, 267, 267].

3) Macroscopic superpositions in Bose-Josephson junctions

As stated in chapter 3, the tuning of the interaction energies that we have proposed for protecting the superposition of coherent states from decoherence does not apply directly to the experiments performed in Basel. P. Treutlein has developed in the last years a quantum tomography protocol to characterize fully the state in the junction at a given time. He is now interested in observing a superposition of coherent states. Some theoretical efforts are needed to study the experimentally realizable ways to reduce decoherence on the superposition. The main source of decoherence is atom losses. Two-body losses could be suppressed by using initial coherent states with a small number of atoms in the upper internal state (as stressed in Sec. 3.5.1, the two-body losses are negligible in the lower state; moreover, the mean number of lost atoms at time t is roughly $\Gamma_{2-\mathrm{body}} N_0 t$ for two-body losses, whereas it is of the order of $\Gamma_{1-\mathrm{body}} t$ for one body losses, with $\Gamma_{1-\mathrm{body}} \gg \Gamma_{2-\mathrm{body}}$). An idea would be to find an optimal time-dependent potential for the upper mode (splitting of the two condensates) so as to reduce the effects of one-body losses as well (see Sec. 3.6).

PERSPECTIVES FROM PART II: QUANTUM CORRELATIONS AND DISTINGUISHABILITY OF QUANTUM STATES

Une sortie, c'est une entrée qu'on prend dans l'autre sens (B. Vian).

Our results on the Bures geometric discord (Sec. 12.2) indicate that contractive distances on the set of quantum states provide useful tools to describe quantum correlations in composite quantum systems. They show that the geometric approach has clear links with other quantum information problems such as the discrimination of non-orthogonal states. This approach should be developed in the future. Several related papers have appeared in the last years, with in particular the use of other contractive distances like the trace distance [63] and the quantum Hellinger distance [L. Chang and S. Luo, Phys. Rev. A **87**, 062303 (2013)], the determination of classical correlations instead of quantum correlations [T.R. Bromley *et al.*, J. Phys. A: Math. Theor. **47**, 405302 (2014)], the study of the geometric discord for Gaussian states in continuous variable systems [P. Marian and T.A. Marian, *Hellinger distance as a measure of Gaussian discord*, arXiv:1408.4477]. Further investigations could go in the following directions.

1) Concrete physical systems

(a) **Time evolution of quantum correlations and decoherence.**
One should investigate the time evolution of the geometric discord in specific models of dissipative dynamics for optical or atomic systems. The main advantage of the geometric approach with respect to the entropy approach is that it provides additional information, namely, the closest classical state(s) to a given state ρ of the bipartite system. The general form of the closest A-classical states σ_ρ for the Bures distance has been given in Proposition 12.2.3 in the general case and in Proposition 12.2.6 when subsystem A is a qubit. We have determined σ_ρ explicitly for two-qubit states ρ with maximally mixed marginals (see (4b) in the publication list). The information on ρ provided by its closest classical state(s) would certainly be relevant for a deeper understanding of the way quantum correlations decrease with time when the system is subject to decoherence. For instance, in order to maintain longer the quantum correlations in the system, the trajectory $t \mapsto \rho(t)$ should be as orthogonal as possible to the geodesics linking the time-evolved state to its closest classical state. We are not aware of any work in the literature in this direction.

(b) **Ground state properties and quantum phase transitions.**
In the last decade it has been pointed out that one may see the signature of a quantum phase transition by looking at the entanglement properties of the ground state of the system in the thermodynamic limit. In particular, right at the transition one sometimes observes a degeneracy of the highest eigenvalue of the reduced density matrix of the ground state for given bipartitions of the system. Furthermore, it has been shown in some spin models that the gap between the first and second eigenvalues exhibits a scaling behavior near the transition with the critical exponent of the universality class of the model [G. De Chiara, L. Lepori, M. Lewenstein, and A. Sanpera, Phys. Rev. Lett. **109**,

237208 (2012)] and that the derivative of the concurrence between two spins diverges logarithmically [A. Osterloh, L. Amico, G. Falci, and R. Fazio, Nature **416**, 608 (2002)]. Proposition 12.2.2 above shows that the aforementioned degeneracy is related to the existence of a symmetry characterized by infinitely many closest classical states to the ground state. It is also suggested in the literature that the phase transition can be singled out by operational properties on the ground state (such as local convertibility using LOCC operations and an ancilla system acting as a catalyst, see the remark of Sec. 10.2), but this has been criticized by H. Bragança *et al.* [*Nonuniversality of entanglement convertibility*, arXiv:1312.0619].

2) Multipartite systems, systems with continuous variables

(a) **Measures of quantum correlations in systems with a large number of particles.**
In chapter 3 we have used the quantum Fisher information to quantify quantum correlations between the atoms. It has been argued that this Fisher information gives the amount of quantum correlations useful for phase estimation in interferometry. Moreover, k-producible states, i.e., states admitting pure state decompositions involving tensor products of states of at most k particles, have Fisher informations bounded from above by a known increasing function of k [135, 236]. This shows that states with high Fisher informations present some truly multipartite entanglement. However, the very notion of particle entanglement is still a subject of controversy for indistinguishable particles [G.C. Ghirardi and L. Marinatto, Phys. Rev. A **70**, 012109 (2004); N. Killoran, M. Cramer, and M.B. Plenio, Phys. Rev. Lett. **112**, 150501 (2014)]. Since the quantum Fisher information coincides up to a numerical factor with the Bures metric, it has a geometrical interpretation in terms of the Bures distance (in fact, it gives the speed at which the state separates from its time-evolved state, see Sec. 8.5). It has been argued recently that the Fisher information optimized over all Hamiltonians H generating the time evolution is a measure of quantum correlations satisfying the three axioms of Definition 12.2.1 [D. Girolami *et al.*, Phys. Rev. Lett. **112**, 210401 (2014)]. It would be interesting to find its relation with the Bures geometric discord, or at least to establish some bounds of one measure in terms of the other.

(b) **Monogamy relation between the geometric discord and geometric entanglement?**
A natural question is whether in three-partite systems the geometric measure of entanglement and the geometric discord satisfy a monogamy bound similar to that given in Sec. 11.4 for the entanglement of formation and the entropy-based quantum discord.

(c) **Geometric discord for systems with continuous variables.**
In some experiments like those performed at the LKB in Paris, the quantum information is encoded in optical modes of the electromagnetic field. These modes are harmonic oscillators with infinite dimensional Hilbert spaces (continuous variables). When two modes are in a Gaussian state, the quantum correlations between them can be characterized by the Gaussian discord introduced by Adesso and Datta and by Giorda and Paris, which has been shown recently to coincide with the usual quantum discord for a large family of Gaussian states [Pirandola *et al.*, *Optimality of Gaussian discord*, arXiv:1309.2215]. The generalization of the results presented in Sec. 12.2 to bipartite systems with continuous variables is an interesting open problem.

3) Relations between the Bures geometric discord and other measures of quantum correlations

(a) **"Local" geometric measures of quantum correlations.**
Although we have defined geometric measures of quantum correlations as distances of a given quantum state ρ to the set of separable or classical states, one could also think of measuring them locally (in a geometrical meaning). An attempt in this direction has been given in Sec. 10.1 by substituting the usual correlator between two local observables by the Bures metric. From a physical viewpoint it is quite natural to characterize quantum correlations (in particular entanglement) as correlations between local observables, but the maximization procedure over all such observables proposed in Sec. 10.1 is artificial and somehow arbitrary. Further work is required to obtain a more natural definition, even for pure states. Up to our knowledge, our definition of the concurrence (10.5) is new for bipartite systems with subsystems of space dimensions larger than two, albeit there have been

other generalizations in the literature. A different idea would be to see if quantum correlations could be related to the Gauss curvature of the Bures metric.

(b) **The choice of the distance.**

At this stage there is no clear reason for choosing the Bures distance within the family of all contractive Riemannian distances for quantifying quantum correlations. This family has been fully characterized by Petz (Sec. 8.6). I am currently working with W. Roga and F. Illuminati on the geometric discord obtained from the quantum Hellinger distance. The latter is defined as the Hilbert-Schmidt distance of the square roots of the two states; it is contractive thanks to Lieb's concavity theorem (Sec. 8.3). We have found that the corresponding measure of quantum correlations has similar properties than the Bures geometric discord, and have obtained bounds of one measure in terms of the other. It turns out that the quantum Hellinger discord is particularly easy to compute. In fact, it is simply related to the Hilbert-Schmidt geometric discord defined in Ref. [67], evaluated for the square root of the state instead of the state itself. The Hilbert-Schmidt discord is known to be easily computable, especially if subsystem A is a qubit and subsystem B has arbitrary dimensions [67], but it is not a *bona fide* measure of quantum correlations because of the lack of monotonicity of the Hilbert-Schmidt distance (see Sec. 12.2). The geometric discord with quantum Hellinger distance does not suffer from this drawback. It is a promising measure of quantum correlations satisfying all the axioms of Definition 12.2.1, while being at the same time an easily computable quantity.

(c) **Measurement-induced geometric discord and discord of response.**

Apart from the geometric discord, two related measures of quantum correlations have attracted a lot of attention in recent years: (1) the distance of the state ρ to the set of states obtained from ρ by the action of local von Neumann measurements on subsystem A (measurement-induced geometric discord) [162], and (2) the distance of ρ to the set of states obtained from ρ by the action of local unitary operations on A with a given non-degenerate spectrum (discord of response) [203]. Since the output of a local von Neumann measurement on A is always A-classical, the measure defined in (1) cannot be smaller than the geometric discord. The two measures are equal for the Hilbert-Schmidt distance, but differ for the Bures and quantum Hellinger distances. It is highly desirable to compare these different measures of quantum correlations defined in the literature. Together with W. Roga and F. Illuminati, we aim to obtain explicit relations or bounds of one measure in terms of the other. One can also compare the measures obtained from different contractive distances. An unpleasant fact is that a state can be more quantum correlated than a given state for one distance and less quantum correlated for another distance.

4) Other open problems

(a) **Dissipative evolutions which do not decrease quantum correlations in bipartite systems.**

Let us beg the following question: given a pair of quantum states (ρ, σ) on a finite-dimensional Hilbert space, what are the quantum operations $\mathcal{M}$ such that the distance between the transformed states $\mathcal{M}(\rho)$ and $\mathcal{M}(\sigma)$ is the same as the distance between ρ and σ? A sufficient condition is that $\mathcal{M}$ admits an approximate reversal, as in Petz's theorem on the monotonicity of the relative entropy (Theorem 7.2.1). In fact, the equality $d(\rho, \sigma) = d(\mathcal{M}(\rho), \mathcal{M}(\sigma))$ then follows from the contractivity of the distance. One may wonder if this condition is actually necessary for any contractive Riemannian distances d, as it is for the relative von Neumann entropy. As a preliminary step, one could try to generalize the proof of Petz's theorem given in Sec. 7.2 to the quantum relative Rényi entropies defined in Sec. 7.3, which contain both the von Neumann entropy and the Bures distance as special cases. The characterization of the quantum operations which do not decrease the distance between two given states would enable to obtain the operations which do not decrease quantum correlations in a given state. One may hope that such operations will not depend on the contractive distance chosen to define the measure of quantum correlations, or at least that one may be able to identify subclasses of distances for which the operations are the same. Results in this direction have been obtained recently by M. Cianciaruso, T.R. Bromley, W. Roga, R. Lo Franco, and G. Adesso [*Universality of the freezing of geometric quantum correlations*, arXiv:1411.2978].

(b) **Quantum relative Réyni entropies.**

The relative α-entropies S_α defined by M. Müller-Lennert *et al.* [175] and by M. Wilde, A. Winter, and D. Yang [260] (Sec. 7.3) are likely to find many applications in quantum information theory.

Since they are simply related for $\alpha = 1/2$ to the Bures distance, one may wonder if an analog of Proposition 8.3.1 could be true also for S_α with α different from $1/2$, that is, if $S_\alpha(\rho, \sigma)$ could be obtained as a supremum of the classical Rényi divergence for the set of outcome probabilities $\mathbf{p}$ and $\mathbf{q}$ corresponding to measurements on ρ and σ.

(c) **Properties of the state ensemble $\{\rho_i, \eta_i\}$ in Proposition 12.2.3.**

Our main result in chapter 12 is a link between the geometric discord and the minimal error probability in the discrimination of the states from the ensemble $\{\rho_i, \eta_i\}$ specified in Proposition 12.2.3. We have shown that this state ensemble is given in terms of the conditional states of a von Neumann measurement, as summarized in Fig. 12.1. This suggests that the ensemble $\{\rho_i, \eta_i\}$ has special properties from a quantum information perspective, which should be investigated.

RELATED PROJECTS

La science est surtout une prise de conscience de plus en plus complète de ce qui peut et doit être découvert.
(B. Vian)

1) Dynamics of open quantum systems: rigorous derivation of non-Markovian master equations

Starting with the pioneering works of V. Jakšić and C.-A. Pillet [Ann. Inst. H. Poincaré Phys. Théor. **62**, 47 (1995); Comm. Math. Phys. **176**, 619 (1996)], the mathematical analysis of open quantum systems has mainly focused in the last decades on large time behaviors, in particular the relaxation to equilibrium or evolution towards stationary states. However, physicists are more often interested by intermediate times. For instance, as stressed in chapters 1-3, decoherence effects are studied with the help of master equations giving the time evolution of the system state. In the weak coupling limit (van Hove limit), which consists in letting the system-environment coupling constant λ go to zero while rescaling the time like $t = \lambda^{-2}\tau$ with τ fixed, one obtains a Markovian master equation of the Lindblad type [159]. This limit was put on a rigorous footing by E.B. Davies in the seventies [73]. However, it is well known that the weak coupling limit does not exist if the system Hamiltonian H_{S} has dense or continuous spectrum, because it is based on a spectral averaging. For small nonzero coupling constants λ, one obtains a good approximation of the exact dynamics provided that the spectral gaps between neighboring eigenvalues of H_{S} are much larger than the damping constant γ appearing in the Lindblad equation. To describe the dynamics of open systems with small energy gaps, which are common in condensed matter, physicists often rely on the time-dependent Bloch-Redfield master equation [46], which is perturbative in λ but does not involve any spectral averaging. This equation also includes some small-time non-Markovian effects. I have a long-standing project to justify rigorously this time-dependent Bloch-Redfield equation, by showing that the trace norm of the difference between the solution of this equation and the exact reduced density matrix $\rho_{\mathsf{S}}(t)$ is small for small coupling constants λ and *fixed* times t. More precisely, according to unpublished calculations I have made some years ago, the error should be of the order of $\lambda^4 t + \lambda^6 t^2$. I would like to show that this bound is optimal, that is, to find a specific example of system coupled to a bath which does not return to a stationary state and for which the error grows with time like $\lambda^4 t$. An intriguing fact is that I discovered another master equation, slightly different from the Bloch-Redfield equation, which leads to comparable errors. I suspect that there exists a whole family of such equations. In order to find which one is better to use, it would be of interest to compare the corresponding stationary states. Although in principle the master equation approximates badly the exact dynamics at large times, in many situations one may expect firstly that this equation has a stationary state $\rho_{\mathsf{S}}(\infty)$ coinciding with the exact stationary state up to fourth order in λ, and secondly that its solution is already close to $\rho_{\mathsf{S}}(\infty)$ at intermediate times. Another interesting problem is to derive master equations including higher order terms in the coupling constant λ.

2) Rigorous derivation of the Bose-Hubbard model for bosons trapped in a double well potential

A lot of progress has been made in the last decade in the mathematical physics community on the rigorous treatment of interacting bosons. The starting issue has been the justification of the mean-field approximation, which consists in assuming that the atoms are independent but each atom feels an effective mean-field potential created by the other atoms. The linear many-body problem is turned in the mean-field limit into a non-linear one-body problem. For trapped neutral atoms, the interactions are short range and are usually replaced by a contact interaction, leading to the celebrated Gross-Pitaevskii equation. This equation has been justified rigorously from the many-body problem in a seminal series of works by E.H. Lieb, R. Seiringer, and J. Yngvason [Phys. Rev. A. **61**, 043602 (2000); Phys. Rev. Lett. **94**, 080401 (2005)], E.H. Lieb, R. Seiringer, J.P. Solovej, and J. Yngvason [Oberwolfach Seminar Series **34** (Birkhäuser, Boston, 2005)], and E.H. Lieb and R. Seiringer [Comm. Math. Phys. **264**, 505-537 (2006)]. In these works, it was shown that the ground state energy of the N-body Hamiltonian converges to the minimal energy of the Gross-Pitaevskii energy functional in the limit of a large number of atoms N. More recently, a rigorous treatment of Bogoliubov's theory has been obtained (see [J. Derezinski and M. Napiórkowski, *Excitation spectrum of interacting bosons in the mean-field infinite-volume limit*, arXiv:1305.3641], [P. Grech and R. Seiringer, Comm. Math. Phys. **322**, 559 (2013)], [M. Lewin, P.T. Nam, S. Serfaty, and J.P. Solovej, *Bogoliubov spectrum of interacting Bose gases*, arXiv:1211.2778]). The Bogoliubov theory gives the sub-leading correction (beyond the mean-field) to the ground state energy and predicts the excitation spectrum, which is of crucial importance for the description of superfluid properties of Bose gases.

I started recently to work with N. Rougerie on the many-body problem for interacting bosons trapped in a symmetric double-well potential. The main goal is to justify rigorously in the limit $N \to \infty$ the two-mode Bose-Hubbard Hamiltonian (3.1) used in chapter 3 to describe the dynamics of Bose-Josephson junctions. In a first stage, one can obtain the correct ground state when the tunneling between the two wells is much smaller than the atomic interactions and in the opposite regime where tunneling dominates interactions. In the first limit it is energetically favorable to have half of the atoms in each well, so that the ground state is the Fock state $|N/2, N/2\rangle$, whereas in the second limit the ground state is the spin coherent state (3.7), i.e., all atoms are in the same one-body state delocalized in the two wells. For a periodic trapping potential with infinitely many wells, these localized and delocalized states characterize the Mott insulator and superfluid phases, respectively. The energy difference per atom between the localized and delocalized states is the sum of a term of the order of $1/N$ and of the tunneling amplitude between the two wells (which is exponentially small in the inter-well distance). Therefore, to determine the ground state energy one has to include the sub-leading terms in N given by Bogoliubov's theory. Note that a possibly better Hamiltonian for describing the dynamics of the atoms trapped in the double-well potential could be obtained as follows. One first determines the minimizer of the Gross-Pitaevskii energy functional with a fixed number of atoms in each well, then replaces these atom numbers by the number operators n_1 and n_2 in each mode, and finally adds the tunneling term $K(a_1^\dagger a_2 + a_2^\dagger a_1)$. This Hamiltonian differs from the Bose-Hubbard Hamiltonian (3.1) by additional terms involving powers of n_1 and n_2 higher than two.

Appendix A

Operator monotone and operator convex functions

We recall in this appendix some basic facts about operator monotone and operator convex functions. We refer the reader to the lecture notes [50] and the book [38] for more complete presentations of these notions.

We denote by $\mathcal{B}(\mathcal{H})_+$ the set of non-negative operators on $\mathcal{H}$, with $\dim(\mathcal{H}) = n < \infty$. A function $f : \mathbb{R}_+ \to \mathbb{R}$ is *operator convex* if for any $n \times n$ matrices $A, B \in \mathcal{B}(\mathcal{H})_+$ and any $0 \leq \eta \leq 1$, it holds $f((1-\eta)A + \eta B) \leq (1-\eta)f(A) + \eta f(B)$. It is *strictly operator convex* if the inequality holds with equality if and only if $\eta \in 0, 1$ or $A = B$. It is *operator concave* if $-f$ is operator convex. It is *operator monotone-increasing* if for any $A, B \in \mathcal{B}(\mathcal{H})_+$, $A \leq B \Rightarrow f(A) \leq f(B)$, and *operator monotone-decreasing* if the reverse equality holds.

It is not hard to show (see e.g. [50]) that $f(x) = x^{-1}$ is operator monotone-decreasing and strictly operator convex. Clearly, this is then also true for $f(x) = (x+t)^{-1}$ for any $t \geq 0$. According to the integral representation

$$A^\alpha = \frac{\sin(\alpha\pi)}{\pi} \int_0^\infty \mathrm{d}t\, t^\alpha \Big(\frac{1}{t} - \frac{1}{t+A} \Big) \,, \tag{1}$$

it follows that $f_\alpha(x) = x^\alpha$ is operator monotone-increasing and strictly operator concave for $0 < \alpha < 1$. Similarly, one shows that f_α is operator monotone-decreasing and operator convex for $\alpha \in [-1, 0]$ and operator convex for $\alpha \in [1, 2]$. However, for instance the square function f_2 is not operator monotone and the cube function f_3 is not operator convex. One can establish that $g(x) = \ln x$ and $f(x) = x \ln x$ are operator concave and operator convex, respectively, thanks to the identities

$$\ln A = \lim_{\alpha \to 0} \alpha^{-1}(A^\alpha - 1) \quad , \quad A \ln A = \lim_{\alpha \to 1} \frac{A^\alpha - A}{\alpha - 1} \,. \tag{2}$$

Another example of monotone-increasing function is $f(x) = (x-1)/\ln x = \int_0^1 \mathrm{d}\alpha\, x^\alpha$.

Operator monotonicity is much stronger than usual monotonicity of real functions. This is clear from Löwner's theorem, which states that if $f : (-1, 1) \to \mathbb{R}$ is operator monotone and non-constant, then f admits the integral representation

$$f(x) = f(0) + f'(0) \int_{-1}^1 \mathrm{d}\mu(t) \frac{x}{1 - xt} \,, \tag{3}$$

where μ is a probability measure on $[-1, 1]$ (see [38], Corollary V.4.5). Furthermore, if $f : \mathbb{R}_+ \to \mathbb{R}_+$ is continuous, then f is operator monotone if and only if it is operator concave ([38], Theorem V.2.5). The fact that concavity implies monotonicity is easily obtained by noting that if $0 \leq A \leq B$, $C = B - A \geq 0$, and $0 \leq \eta < 1$, then $f(\eta B) \geq \eta f(A) + (1-\eta) f(\eta(1-\eta)^{-1} C)$ (by concavity). As $f(x) \geq 0$ the second term in the right-hand side is non-negative and thus $f(\eta B) \geq \eta f(A)$. Letting $\eta \to 1$ we get $f(B) \geq f(A)$. The converse implication can be shown by similar arguments as those used to establish (4) below and by invoking the fact that if (4) is satisfied for any contraction C then f is operator convex (see [38] for more detail).

Another remarkable result valid for continuous functions $f : [0, a) \to \mathbb{R}$ is that f is operator convex and $f(0) \leq 0$ if and only if $g(x) = x^{-1} f(x)$ is operator monotone on $(0, a)$ ([38], Theorem V.2.9). Similarly, for functions $f : (-1, 1) \to \mathbb{R}$ of class C^2, if f is operator convex and $f(0) = 0$ then $g(x)$ is operator monotone ([38], Corollary V.3.11). An integral representation for non-linear operator convex functions f can be obtained with the help of the last property, by applying (3) to $g(x)$.

If $f : \mathbb{R}_+ \to \mathbb{R}$ is operator convex and $f(0) \leq 0$, then

$$f(C^*AC) \leq C^* f(A) C \tag{4}$$

for any contraction $C \in \mathcal{B}(\mathcal{H})$, $\|C\| \leq 1$, and any $A \in \mathcal{B}(\mathcal{H})_+$. This inequality can be shown as follows [112]. Let us consider the matrices

$$\widehat{A} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \quad , \quad \widehat{U}_\pm = \begin{pmatrix} C & \pm D \\ E & \mp C^* \end{pmatrix} \tag{5}$$

with $D = \sqrt{1 - CC^*}$ and $E = \sqrt{1 - C^*C}$ (the latter operators are well defined since $\|C\| \leq 1$). An explicit calculation shows that $\widehat{U}_\pm$ is unitary and

$$\begin{pmatrix} C^*AC & 0 \\ 0 & DAD \end{pmatrix} = \frac{1}{2} \sum_{\epsilon=\pm} \widehat{U}_\epsilon^* \widehat{A} \widehat{U}_\epsilon \, . \tag{6}$$

If f is operator convex and $f(0) \leq 0$, then

$$\begin{aligned} \begin{pmatrix} f(C^*AC) & 0 \\ 0 & f(DAD) \end{pmatrix} &= f \begin{pmatrix} C^*AC & 0 \\ 0 & DAD \end{pmatrix} \\ &\leq \frac{1}{2} \sum_{\epsilon=\pm} f(\widehat{U}_\epsilon^* \widehat{A} \widehat{U}_\epsilon) \\ &\leq \frac{1}{2} \sum_{\epsilon=\pm} \widehat{U}_\epsilon^* \begin{pmatrix} f(A) & 0 \\ 0 & 0 \end{pmatrix} \widehat{U}_\epsilon = \begin{pmatrix} C^* f(A) C & 0 \\ 0 & D f(A) D \end{pmatrix} . \end{aligned} \tag{7}$$

This implies in particular the bound (4). Conversely, it is shown in [112] that if this bound is satisfied for any orthogonal projection C and any $A \in \mathcal{B}(\mathcal{H})_+$, then f is operator convex and $f(0) \leq 0$.

Let $\mathcal{M}$ be a quantum operation on $\mathcal{B}(\mathcal{H})$ and $f : \mathbb{R}_+ \to \mathbb{R}$ be operator convex. Then the following Jensen-type inequality holds [75]:

$$f(\mathcal{M}^*(A)) \leq \mathcal{M}^*(f(A)) \quad , \quad A \in \mathcal{B}(\mathcal{H})_+ \, . \tag{8}$$

A simple justification of this inequality is as follows. Since $\mathcal{M}^*(c\,1) = c\,1$ for any constant $c \in \mathbb{R}$, one may assume without loss of generality that $f(0) = 0$. Let $A \in \mathcal{B}(\mathcal{H})_+$. According to Stinespring's theorem (Sec. 5) one can find a unitary operator U on an enlarged space $\mathcal{H} \otimes \mathcal{H}_{\mathsf{E}}$ and a vector $|\epsilon_0\rangle \in \mathcal{H}_{\mathsf{E}}$ such that $\mathcal{M}^*(A) = \langle \epsilon_0 | U^* A \otimes 1 U | \epsilon_0 \rangle$. Let us set $P_0 = |\epsilon_0\rangle\langle\epsilon_0|$. Applying (4) with $C = 1 \otimes P_0$, one gets

$$\begin{aligned} f(\mathcal{M}^*(A)) \otimes P_0 &= f(1 \otimes P_0 \, U^* A \otimes 1 U \, 1 \otimes P_0) \\ &\leq 1 \otimes P_0 f(U^* A \otimes 1 U) 1 \otimes P_0 = \mathcal{M}^*(f(A)) \otimes P_0 \, . \end{aligned} \tag{9}$$

Appendix B

Trace inequalities

In this appendix some inequalities involving the $\|\cdot\|_p$-norms are stated or derived.

1) Let us first recall the triangle and "inverse triangle" inequalities: for any matrices A and B one has

$$\|A+B\|_p \begin{cases} \leq \|A\|_p + \|B\|_p & \text{if } p \geq 1 \\ \geq \|A\|_p + \|B\|_p & \text{if } 0 < p < 1. \end{cases} \tag{1}$$

This shows that the map $A \mapsto \|A\|_p$ defined by (4.2) is a norm for $p \geq 1$, but this is not the case for $p < 1$. One deduces the bound

$$\mathrm{tr}[\sqrt{|A|^2 + |B|^2}] \leq \mathrm{tr}\,|A| + \mathrm{tr}\,|B| \tag{2}$$

by applying (1) for $p = 1$ to the matrices

$$\widehat{A} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \quad , \quad \widehat{B} = \begin{pmatrix} 0 & 0 \\ B & 0 \end{pmatrix} .$$

2) Another standard result is the Lieb-Thirring inequality [155]. We quote here without proof a generalization of this inequality derived by Araki [14]. Let $k > 0$ and A and B be non-negative operators. If $\alpha \geq 1$ then

$$\left\| B^{\frac{1}{2}} A B^{\frac{1}{2}} \right\|_{\alpha k}^{\alpha} \leq \left\| B^{\frac{\alpha}{2}} A^{\alpha} B^{\frac{\alpha}{2}} \right\|_k . \tag{3}$$

Taking $\alpha \to \alpha^{-1}$ and $k \to k/\alpha$, one can deduce that the reverse inequality holds true if $0 \leq \alpha \leq 1$.

3) Next, let us show that for any square matrices A, B, C, and D of the same size, the following bound generalizing the Cauchy-Schwarz inequality $\|AB\|_1 \leq \|A\|_2 \|B\|_2$ holds true [173]

$$\|AB + CD\|_1^2 \leq \left(\|A\|_2^2 + \|D\|_2^2\right)\left(\|B\|_2^2 + \|C\|_2^2\right) . \tag{4}$$

Actually, let us form the 2×2 block matrices

$$\widehat{E} = \begin{pmatrix} A^* & 0 \\ C^* & 0 \end{pmatrix} \quad , \quad \widehat{F} = \begin{pmatrix} B & 0 \\ D & 0 \end{pmatrix} .$$

Then

$$\|AB + CD\|_1^2 = \left\|\widehat{E}^* \widehat{F}\right\|_1^2 \leq \|\widehat{E}\|_2^2 \|\widehat{F}\|_2^2 = \left(\|A\|_2^2 + \|C\|_2^2\right)\left(\|B\|_2^2 + \|D\|_2^2\right) .$$

But $CD = UD^*C^*U$ with U unitary by the polar decomposition. Applying the above inequality with C and D replaced by UD^* and C^*U and using the unitary invariance of $\|\cdot\|_2$, one gets the desired result (4).

4) Let $B = (B_{ij})_{i,j=1}^m$ be a non-negative $m \times m$ operator-valued matrix, whose entries B_{ij} are given by $p_i \times p_j$ matrices. Denote by $A = \sqrt{B} = (A_{ij})_{i,j=1}^m$ the square root of B. Then for any $j = 1, \ldots, m$, one has [29]

$$\sum_{i, i \neq j} \left\|A_{ij}\right\|_2^2 \leq \frac{1}{2} \sum_{i, i \neq j} \left\|B_{ij}\right\|_1 . \tag{5}$$

Let us first establish (5) for $m = 2$. Thanks to the singular value decomposition and the unitary invariance of the $\|\cdot\|_p$-norms, we may assume without loss of generality that A_{12} is a diagonal $p_1 \times p_2$ matrix, i.e.$A_{12} = \sum_{k=1}^{p} \sqrt{\nu_k}|k\rangle\langle k|$ with $p = \min\{p_1, p_2\}$. By a standard argument, the non-negativity of A implies

$$|\langle\varphi_1|A_{12}|\varphi_2\rangle|^2 \leq \langle\varphi_1|A_{11}|\varphi_1\rangle\langle\varphi_2|A_{22}|\varphi_2\rangle$$

for any vectors $|\varphi_1\rangle \in \mathbb{C}^{p_1}$ and $|\varphi_2\rangle \in \mathbb{C}^{p_2}$. Using this bound and the relation $B_{12} = A_{11}A_{12} + A_{12}A_{22}$, we find

$$\|A_{12}\|_2^2 = \sum_{k=1}^{p} \nu_k \leq \sum_{k=1}^{p} \sqrt{\nu_k\langle k|A_{11}|k\rangle\langle k|A_{22}|k\rangle} \leq \frac{1}{2}\sum_{k=1}^{p} \sqrt{\nu_k}\big(\langle k|A_{11}|k\rangle + \langle k|A_{22}|k\rangle\big) = \frac{1}{2}\|B_{12}\|_1 \ .$$

Consider now the general case $m \geq 2$. The idea is to write B as a 2×2 block matrix such that the upper left and lower right blocks are the $(m-1) \times (m-1)$ matrix $(B_{ij})_{i,j=1}^{m-1}$ and the single entry B_{mm}, respectively, whereas the upper right (lower left) block forms a column (line) vector with entries B_{im} (B_{mi}). A similar block decomposition can be made for A. Applying the foregoing result for $m = 2$, one gets

$$\begin{aligned}\sum_{i,i\neq m} \big\|A_{im}\big\|_2^2 = \left\| \begin{pmatrix} A_{1m} \\ \vdots \\ A_{(m-1)m} \end{pmatrix} \right\|_2^2 &\leq \frac{1}{2}\left\| \begin{pmatrix} B_{1m} \\ \vdots \\ B_{(m-1)m} \end{pmatrix} \right\|_1 \\ &= \frac{1}{2}\left\| \sqrt{\sum_{i,i\neq m} |B_{im}|^2} \right\|_1 \leq \frac{1}{2}\sum_{i,i\neq m} \|B_{im}\|_1 \ ,\end{aligned}$$

where we have used (2) in the last bound. This proves (5) for $j = m$. By an appropriate unitary conjugation, one deduces that the bound holds for any j.

5) The following trace inequality plays a central role in the derivation of the quantum Chernoff bound [19]: for any positive square matrices $A > 0$ and $B > 0$ and any $0 \leq s \leq 1$,

$$\frac{1}{2}\Big(\mathrm{tr}(A) + \mathrm{tr}(B) - \mathrm{tr}\,|A - B|\Big) \leq \mathrm{tr}(A^{1-s}B^s) \ . \tag{6}$$

This inequality was first shown in [19], but the proof in this reference is not very transparent. We present here a much simpler proof due to Ozawa, which has been first reported in [140]. Denoting by $O_\pm = (|O| \pm O)/2 \geq 0$ the positive and negative parts of O, one may express $\mathrm{tr}\,|A-B|$ as $2\,\mathrm{tr}(A-B)_+ - \mathrm{tr}(A) + \mathrm{tr}(B)$. Thus (6) is equivalent to

$$\mathrm{tr}\big((A^s - B^s)A^{1-s}\big) \leq \mathrm{tr}(A-B)_+ \ .$$

Since $f(x) = x^s$ is operator monotone (see Appendix A) and $A \leq A + (A-B)_- = B + (A-B)_+$, one has $A^s \leq (B + (A-B)_+)^s$. Hence

$$\begin{aligned}\mathrm{tr}\big((A^s - B^s)A^{1-s}\big) &\leq \mathrm{tr}\big(\big[(B + (A-B)_+)^s - B^s\big]A^{1-s}\big) \\ &\leq \mathrm{tr}\big(\big[(B + (A-B)_+)^s - B^s\big](B + (A-B)_+)^{1-s}\big) \ ,\end{aligned}$$

where the second inequality relies on the similar bound $B^s \leq (B+(A-B)_+)^s$. By rearranging the product in the last trace and using the latter bound with $s \leftrightarrow (1-s)$, one gets

$$\mathrm{tr}\big((A^s - B^s)A^{1-s}\big) \leq \mathrm{tr}(B) + \mathrm{tr}(A-B)_+ - \mathrm{tr}\big(B^s(B + (A-B)_+)^{1-s}\big) \leq \mathrm{tr}(A-B)_+ \ .$$

This concludes the justification of (6).

Appendix C

List of publications

(1a) D. Spehner, F. Haake, *Quantum measurements without macroscopic superpositions*, Phys. Rev. A 77 (2008), 052114

(1b) D. Spehner, F. Haake, *Decoherence bypass of macroscopic superpositions in quantum measurement*, J. Phys. A: Math. Theor. 41 (2008), 072002

(1c) D. Spehner, F. Haake, *Quantum measurements without Schrödinger cat states*, J. Phys.: Conf. Series 84 (2007), 012018 (conference proceedings)

(2a) S. Vogelsberger, D. Spehner, *Average entanglement for Markovian quantum trajectories*, Phys. Rev. A 82 (2010), 052327

(2b) S. Vogelsberger et D. Spehner, *Entanglement evolution for quantum trajectories*, J. Phys.: Conf. Series 306 (2011), 012029 (conference proceedings)

(3a) D. Spehner, K. Pawlowski, G. Ferrini, A. Minguzzi, *Effect of one-, two-, and three-body atom loss processes on superpositions of phase states in Bose-Josephson junctions*, Eur. Phys. J. B 87 (2014), 157

(3b) K. Pawlowski, D. Spehner, A. Minguzzi, G. Ferrini, *Macroscopic superpositions in Bose-Josephson junctions: Controlling decoherence due to atom losses*, Phys. Rev. A 88 (2013), 013606

(3c) G. Ferrini, D. Spehner, A. Minguzzi, F.W.J. Hekking, *Effect of phase noise on quantum correlations in Bose-Josephson junctions*, Phys. Rev. A 84 (2011), 043628

(3d) G. Ferrini, D. Spehner, A. Minguzzi, F.W.J. Hekking, *Noise in Bose-Josephson junctions: Decoherence and phase relaxation*, Phys. Rev. A 82 (2010), 033621

(4a) D. Spehner, *Quantum correlations and Distinguishability of quantum states*, J. Math. Phys. 55 (2014), 075211 (review article)

(4b) D. Spehner, M. Orszag, *Geometric quantum discord with Bures distance: the qubit case*, J. Phys. A: Math. Theor. 47 (2014), 035302

(4c) D. Spehner, M. Orszag, *Geometric quantum discord with Bures distance*, New J. of Phys. 15 (2013), 103001

Other publications since 2006 not described in the present manuscript:

- S.A. Reyes, L. Morales-Molina, M. Orszag, D. Spehner, *Harnessing gauge fields for maximally entangled state generation*, Eur. Phys. Lett. 108 (2014), 20010
- W. De Roeck, D. Spehner, *Derivation of some translation-invariant Lindblad equations for a quantum Brownian particle*, J. Stat. Phys. 150 (2013), 320
- A. Faggionato, H. Schulz-Baldes, D. Spehner, *Mott law as lower bound for a random walk in a random environment*, Comm. Math. Phys. 263 (2006), 21-64

Bibliography

[1] B. Aaronson, R.L. Franco, and G. Adesso, *Comparative investigation of the freezing phenomena for quantum correlations under nondissipative decoherence*, Phys. Rev. A **88**, 012120 (2013)

[2] P.M. Alberti, *A note on the transition-probability over C^*-algebras*, Lett. Math. Phys. **7**, 25-32 (1983)

[3] M. Ali, A.R.P. Rau, and G. Alber, *Quantum discord for two-qubit X states*, Phys. Rev. A **81**, 042105 (2010)

[4] A.E. Allahverdyan, R. Balian, and T.M. Nieuwenhuizen, *Quantum measurement as a driven phase transition: An exactly solvable model*, Phys. Rev. A **64**, 032108 (2001)

[5] A.E. Allahverdyan, R. Balian, and T.M. Nieuwenhuizen, *Curie-Weiss model of the quantum measurement process*, Europhys. Lett. **61**, 452-458 (2003)

[6] A.E. Allahverdyan, R. Balian, and T.M. Nieuwenhuizen, *Understanding quantum measurement from the solution of dynamical models*, Phys. Rep. **525**, 1-166 (2013)

[7] A.E. Allahverdyan, R. Balian, and T.M. Nieuwenhuizen, *Statistical theory of ideal quantum measurement processes*, preprint arXiv:1303.7257

[8] M.P. Almeida, F. de Melo, M. Hor-Meyll, A. Salles, S.P. Walborn, P.H. Souto Ribeiro, and L. Davidovich, *Environment-Induced Sudden Death of Entanglement*, Science **316**, 579-582 (2007)

[9] T. Ando, *Convexity of certain maps on positive definite matrices and applications to Hadamard products*, Lin. Alg. and Appl. **26**, 203-241 (1979)

[10] J. Anglin, *Cold, dilute, trapped bosons as an open quantum system*, Phys. Rev. Lett. **79**, 6-9 (1997)

[11] H. Araki, *A remark on Bures distance function for normal states*, Publ. RIMS Kyoto Univ. **6**, 477-482 (1970)

[12] H. Araki, *Relative entropy for states of von Neumann algebras*, Publ. RIMS Kyoto Univ. **11**, 809-833 (1976)

[13] H. Araki and T. Masudai, *Positive cones and L_p-spaces for von Neumann algebras*, Publ. RIMS Kyoto Univ. **18**, 339-411 (1982)

[14] H. Araki, *On an Inequality of Lieb and Thirring*, Lett. Math. Phys. **19**, 167-170 (1990)

[15] A. Aspect, J. Dalibard, and G. Roger, *Experimental Test of Bell's Inequalities Using Time-Varying Analyzers*, Phys. Rev. Lett. **49**, 1804 (1982)

[16] S. Attal and Y. Pautrat, *From repeated to continuous quantum interactions*, Ann. H. Poincaré Phys. Théor. **7**, 59-104 (2006)

[17] S. Attal and A. Joye, *Weak Coupling and Continuous Limits for Repeated Quantum Interactions*, J. Stat. Phys. **126**, 1241-1283 (2007)

[18] S. Attal and C. Pellegrini, *Stochastic Master Equations in Thermal Environment*, Open Syst. and Inf. Dyn. **17**, 389-408 (2010)

[19] K.M.R. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, L.I. Masanes, A. Acin, and F. Verstraete, *Discriminating States: The Quantum Chernoff Bound*, Phys. Rev. Lett. **98**, 160501 (2007)

[20] R. Balian, Y. Alhassid, and H. Reinhardt, *Dissipation in many-body systems: a geometric approach based on information theory*, Phys. Rep. **131**, 1 (1986)

[21] R. Balian, *On the principles of quantum mechanics and the reduction of the wave packet*, Am. J. Phys. **57**, 1019-1027 (1989)

[22] R. Balian, *From Microphysics to Macrophysics: Methods and Applications of Statistical Physics* (Springer, 2007)

[23] R. Balian, *The entropy-based quantum metric*, Entropy 2014 **16**(7), 3878-3888 (2014)

[24] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, *Optimum measurements for discrimination among symmetric quantum states and parameter estimation*, Int. J. Theor. Phys. **36**, 1269-1288 (1997)

[25] A. Barchielli and V.P. Belavkin, *Measurements continuous in time and a posteriori states in quantum mechanics*, J. Phys. A: Math. Gen. **24**, 1495 (1991)

[26] S.M. Barnett, *Minimum error discrimination between multiply symmetric states*, Phys. Rev. A **64**, 030303 (2001)

[27] S.M. Barnett, A. Chefles, and I. Jex, *Comparison of two unknown pure quantum states*, Phys. Lett. A **307**, 189-195 (2003)

[28] H. Barnum, C.M. Caves, C.A. Fuchs, R. Jozsa, and B. Schumacher, *Noncommuting mixed states cannot be broadcast*, Phys. Rev. Lett. **76**, 2818 (1996)

[29] H. Barnum and E. Knill, *Reversing quantum dynamics with near-optimal quantum and classical fidelity*, J. Math. Phys. **43**, 2097-2106 (2002)

[30] V.P. Belavkin, *A posterior Schrödinger equation for continuous non demolition measurement*, J. Math. Phys. **31**, 2930-2934 (1990)

[31] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, 2006)

[32] C.H. Bennett, *Quantum Cryptography Using Any Two Nonorthogonal States*, Phys. Rev. Lett. **68**, 3121 (1992)

[33] C.H. Bennett, H.J. Bernstein, S. Popescu, and B. Schumacher, *Concentrating partial entanglement by local operations*, Phys. Rev. A **53**, 2046 (1996)

[34] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin, and W.K. Wootters, *Mixed-state entanglement and quantum error correction*, Phys. Rev. A **54**, 3824 (1996)

[35] C.H. Bennett, D.P. DiVincenzo, C.A. Fuchs, T. Mor, E. Rains, P.W. Shor, J.A. Smolin, and W.K. Wootters, *Quantum nonlocality without entanglement*, Phys. Rev. A **59**, 1070 (1999)

[36] J.A. Bergou, *Discrimination of quantum states*, J. Mod. Opt. **57**(3), 160-180 (2010)

[37] J.A. Bergou, U. Herzog, and M. Hillery, *Discrimination of quantum states* in: *Quantum State Estimation*, (Lecture Notes in Physics vol **649**), M. Paris and J. Rehacek Eds. (Springer, Berlin, 2004), pp. 417-465

[38] R. Bhatia, *Matrix Analysis* (Springer, 1991)

[39] P. Böhi, M.F. Riedel, J. Hoffrogge, J. Reichel, T.W. Hänsch, and P. Treutlein, *Coherent manipulation of Bose-Einstein condensates with state-dependent microwave potentials on an atom chip*, Nat. Phys. **5**, 592-597 (2009)

[40] D. Bohm, *Quantum Theory* (Prentice-Hall, Englewood Cliffs NJ, 1951)

[41] D. Bouwmeester, A. Ekert and A. Zeilinger Eds., *The Physics of Quantum Information: quantum cryptography, quantum teleportation, quantum computation* (Springer, 2000)

[42] O. Bratteli and D.W. Robinson, *Operator Algebras and Quantum Statistical Mechanics*, Vol. 1 and 2 (Springer, Berlin, 1997)

[43] D. Braun, F. Haake, and W.T. Strunz, *Universality of Decoherence*, Phys. Rev. Lett. **86**, 2913 (2001)

[44] D. Braun, *Creation of Entanglement by Interaction with a Common Heat Bath*, Phys. Rev. Lett. **89**, 277901 (2002)

[45] S.L. Braunstein and C.M. Caves, *Statistical Distance and the Geometry of Quantum States*, Phys. Rev. Lett **72**, 3439-3443 (1994)

[46] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, 2002)

[47] H.-P. Breuer, B. Kappler, and F. Petruccione, *Stochastic wave-function method for non-Markovian quantum master equations*, Phys. Rev. A **59**, 1633 (1999)

[48] J. Bub, *How to Solve the Measurement Problem of Quantum Mechanics*, Found. Phys. **18**, 701-722 (1988)

[49] D. Bures, *An extension of Kakutani's theorem on infinite product measures to the tensor product of semifinite w^*-algebras*, Trans. Am. Math. Soc. **135**, 199-212 (1969)

[50] E.A. Carlen, *Trace inequalities and quantum entropy: An introductory course*, in: *Entropy and the quantum*, R. Sims and D. Ueltschi Eds., Contemp. Math. **529** (Amer. Math. Soc., Providence, RI, 2010), pp. 73-140; see http://www.mathphys.org/AZschool/material/AZ09-carlen.pdf

[51] H.J. Carmichael, *An open system approach to Quantum Optics*, Lecture Notes in Physics **m18** (Springer, 1991)

[52] A.R.R. Carvalho, M. Busse, O. Brodier, C. Viviescas, and A. Buchleitner, *Optimal Dynamical Characterization of Entanglement*, Phys. Rev. Lett. **98**, 190501 (2007)

[53] A.R.R. Carvalho and M. França Santos, *Distant entanglement protected through artificially increased local temperature*, New J. Phys. **13**, 013010 (2011)

[54] D. Cavalcanti, L. Aolita, S. Boixo, K. Modi, M. Piani, and A. Winter, *Operational interpretations of quantum discord*, Phys. Rev. A **83**, 032324 (2011)

[55] N.N. Cencov, *Statistical Decision Rules and Optimal Interferences*, Transl. Math. Monogr. **53** (Amer. Math. Soc., Providence, 1982)

[56] N.J. Cerf and C. Adami, *Negative Entropy and Information in Quantum Mechanics*, Phys. Rev. Lett. **79**, 5194-5197 (1997)

[57] A. Chefles, *Unambiguous discrimination between linearly independent quantum states*, Phys. Lett A **239**, 339-347 (1998)

[58] A. Chefles, *Quantum State Discrimination*, Contemporary Physics **41**, 401-424 (2000)

[59] K. Chen and L.-A. Wu, *A matrix realignment method for recognizing entanglement*, Quantum Information and Computation **3**, 193-202 (2003)

[60] H. Chernoff, *A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the sum of Observations*, Ann. Math. Statist. **23**, 493-507 (1952)

[61] C.-L. Chou and L.Y. Hsu, *Minimal-error discrimination between symmetric mixed quantum states*, Phys. Rev. A **68**, 042305 (2003)

[62] R.B.M. Clarke, V.M. Kendon, A. Chefles, S.M. Barnett, E. Riis, and M. Sasaki, *Experimental realization of optimal detection strategies for overcomplete states*, Phys. Rev. A **64**, 012303 (2001)

[63] F. Ciccarello, T. Tufarelli, and V. Giovannetti, *Towards computability of trace distance discord*, New J. Phys. **16**, 013038 (2014)

[64] V. Coffman, J. Kundu, and W.K. Wootters, *Distributed entanglement*, Phys. Rev. A **61**, 052306 (2000)

[65] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Atom-photon Interactions: Basic Processes and Applications* (Wiley, New York, 1992)

[66] G.M. D'Ariano, P. Lo Presti, and P. Perinotti, *Classical randomness in quantum measurements*, J. Phys. A: Math. Gen. **38**, 5979-5991 (2005)

[67] B. Dakić, V. Vedral, and C. Brukner, *Necessary and Sufficient Condition for Nonzero Quantum Discord*, Phys. Rev. Lett **105**, 190502 (2010)

[68] B. Dakić *et al.*, *Quantum discord as resource for remote state preparation*, Nature Physics **8**, 666-670 (2012)

[69] J. Dalibard, Y. Castin, and K. Mølmer, *Wave-function approach to dissipative processes in quantum optics*, Phys. Rev. Lett. **68**, 580 (1992)

[70] J. Dalibard, cours au collège de France, http://www.college-de-france.fr/site/jean-dalibard/course-2013-2014.htm

[71] A. Datta, S.T. Flammia, and C.M. Caves, *Entanglement and the power of one qubit*, Phys. Rev. A **72**, 042316 (2005)

[72] A. Datta, A. Shaji, and C.M. Caves, *Quantum discord and the power of one qubit*, Phys. Rev. Lett **100**, 050502 (2008)

[73] E.B. Davies, *Markovian Master Equations*, Commun. Math. Phys. **39**, 91-110 (1974)

[74] E.B. Davies, *Information and quantum measurement*, IEEE Trans. Inf. Theory **24**, 596-599 (1978)

[75] C. Davis, *A Schwarz inequality for convex operator functions*, Proc. Amer. Math. Soc. **8**, 42-44 (1957)

[76] S. Deléglise, I. Dotsenko, C. Sayrin, J. Bernu, M. Brune, J.M. Raimond, S. Haroche, *Reconstruction of non-classical cavity field states with snapshots of their decoherence*, Nature **455**, 510 (2008)

[77] J. Derezinski and W. De Roeck, *Extended weak coupling limit for Pauli-Fierz operators*, Commun. Math. Phys. **279**, 1-30 (2008)

[78] D. Dieks, *Overlap and distinguishability of quantum states*, Phys. Lett. A **126**, 303-306 (1988)

[79] L. Diósi, in *Irreversible Quantum Dynamics*, Lecture Notes in Physics **622**, F. Benatti and R. Floreanini Eds. (Springer, Berlin, 2003), p.157

[80] L. Diósi, N. Gisin, and W.T. Strunz, *Non-Markovian quantum state diffusion*, Phys. Rev. A **58**, 1699-1712 (1998)

[81] P.J. Dodd and J.J. Halliwell, *Disentanglement and decoherence by open system dynamics*, Phys. Rev. A **69**, 052105 (2004)

[82] R. Durrett, *Probability: Theory and Examples*, 2nd Ed. (Duxbury Press, USA, 1996)

[83] Y.C. Eldar and G.D. Forney Jr, *On quantum detection and the square-root measurement*, IEEE Trans. Inf. Theory **47**, 858-872 (2001)

[84] Y.C. Eldar, A. Megretski, and G.C. Verghese, *Designing optimal quantum detectors via semidefinite programming*, IEEE Trans. Inf. Theory **49**, 1007-1012 (2003)

[85] Y.C. Eldar, *von Neumann measurement is optimal for detecting linearly independent mixed quantum states*, Phys. Rev. A **68**, 052303 (2003)

[86] Y.C. Eldar, A. Megretski, and G.C. Verghese, *Optimal Detection of Symmetric Mixed Quantum States*, IEEE Trans. Inf. Theory **50**, 1198 (2004)

[87] J. Esteve, C. Gross, A. Weller, S. Giovanazzi, and M.K. Oberthaler, *Squeezing and entanglement in a Bose-Einstein condensate*, Nature **455**, 1216 (2008)

[88] F.F. Fanchini, M.F. Cornelio, M.C. de Oliveira, and A.O. Caldeira, *Conservation law for distributed entanglement of formation and quantum discord*, Phys. Rev. A **84**, 012313 (2011)

[89] U. Fano, *Effects of configuration interaction on intensities and phase shifts*, Phys. Rev. **124**, 1866-1878 (1961)

[90] Y. Feng, R. Duan, and M. Ying, *Unambiguous discrimination between mixed quantum states*, Phys. Rev. A **70**, 012308 (2004)

[91] H. Feshbach, *Unified theory of nuclear reactions*, Ann. Phys. **5**, 357-390 (1958)

[92] Z. Ficek and R. Tanás, *Dark periods and revivals of entanglement in a two-qubit system*, Phys. Rev. A **74**, 024304 (2006)

[93] B.V. Fine, F. Mintert, and A. Buchleitner *Equilibrium entanglement vanishes at finite temperature*, Phys. Rev. B **71**, 153105 (2005)

[94] M.C. Fischer, B. Gutierrez-Medina, and M.G. Raizen, *Observation of the Quantum Zeno and Anti-Zeno Effects in an Unstable System*, Phys. Rev. Lett **87**, 040402 (2001)

[95] R.L. Frank and E.H. Lieb, *Monotonicity of a relative Rényi entropy*, J. Math. Phys. **54**, 122201 (2013)

[96] A. Fujiwara, *Quantum channel identification problem*, Phys. Rev. A **63**, 042304 (2001)

[97] M.R. Gallis and G.N. Fleming, *Environmental and spontaneous localization*, Phys. Rev. A **42**, 38 (1990)

[98] F. Galve, G.L. Giorgi, and Z. Zambrini, *Orthogonal measurements are almost sufficient for quantum discord of two qubits*, Eur. Phys. Lett. **96**, 40005 (2011)

[99] G.C. Ghirardi, P. Pearle, and A. Rimini, *Markov processes in Hilbert space and continuous spontaneous localization of systems of identical particles*, Phys. Rev. A **42**, 78-89 (1990)

[100] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum Metrology*, Phys. Rev. Lett. **96**, 010401 (2006)

[101] N. Gisin and I.C. Percival, *The quantum state diffusion model applied to open systems*, J. Phys A: Math. Gen. **25**, 5677 (1992)

[102] D. Giulini *et al.*, *Decoherence and the appearance of a Classical World in Quantum Theory* (Springer, 1996)

[103] V. Gorini and A. Kossakowski, *N-level system in contact with a singular reservoir*, J. Math. Phys. **17**, 1298 (1976)

[104] M. Greiner, O. Mandel, T. Esslinger, T.W. Hänsch, and I. Bloch, *Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms*, Nature **415**, 39-44 (2002)

[105] S. Gröblacher, T. Paterek, R. Kaltenbaek, C. Brukner, M. Żukowski, M. Aspelmeyer, and A. Zeilinger, *An experimental test of non-local realism*, Nature **446**, 871-875 (2007)

[106] C. Gross, T. Zibold, E. Nicklas, J. Estève, and M.K. Oberthaler, *Nonlinear atom interferometer surpasses classical precision limit*, Nature **464**, 1165 (2010)

[107] C. Gross, *Dissertation: Spin squeezing and non-linear atom interferometry with Bose-Einstein condensates*, PhD thesis, Universität Heidelberg (2010)

[108] M. Gu, H.M. Chrzanowski, S.M. Assad, T. Symul, K. Modi, T.C. Ralph, V. Vedral, and P. Koy Lam, *Observing the operational significance of discord consumption*, Nature Physics **8**, 671 (2012)

[109] C. Guerlin, J. Bernu, S. Deléglise, C. Sayrin, S. Gleyzes, S. Kuhr, M. Brune, J.-M. Raimond, and S. Haroche, *Progressive field-state collapse and quantum non-demolition photon counting*, Nature **448**, 889-893 (2007)

[110] O. Gühne and G. Tóth, *Entanglement detection*, Physics Reports **474**, 1 (2009)

[111] S. Hamieh, R. Kobes, and H. Zaraket, *Positive-operator-valued measure optimization of classical correlations*, Phys. Rev. A **70**, 052325 (2004)

[112] F. Hansen and G.K. Pedersen, *Jensen's inequality for operator and Löwner's theorem*, Math. Ann. **258**, 229-241 (1982)

[113] S. Haroche and J.-M. Raimond, *Exploring the quantum: atoms, cavities and photons* (Oxford Univ. Press, 2006)

[114] P. Hausladen and W.K. Wootters, *A "pretty good" measurement for distinguishing quantum states*, J. Mod. Opt. **41**, 2385-2390 (1994)

[115] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W.K. Wootters, *Classical information capacity of a quantum channel*, Phys. Rev. A **54**, 1869-1876 (1996)

[116] P. Hayden, R. Jozsa, D. Petz, and A. Winter, *Structure of States Which Satisfy Strong Subadditivity of Quantum Entropy with Equality*, Commun. Math. Phys. **246**, 359-374 (2004)

[117] W. Heisenberg, *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*, Zeit. für Phys. **43**, 172-198 (1927); English translation in [258]

[118] W. Heisenberg, *Les principes physiques de la théorie des quanta* (Ed. J. Gabay 1990, reprinted from : Gauthier-Villars, 1932, French translation of the conferences given by W. Heisenberg in Chicago during Spring 1929)

[119] C.W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976)

[120] L. Henderson and V. Vedral, *Classical, quantum and total correlations*, J. Phys. A: Math. Gen. **34**, 6899-6905 (2001)

[121] K. Hepp and E.H. Lieb, *Phase Transition in Reservoir Driven Open Systems with Applications to Lasers and Superconductors*, Helv. Phys. Acta **46**, 573-602 (1973)

[122] F. Hiai and D. Petz, *The proper formula for relative entropy and its asymptotics in quantum probability*, Commun. Math. Phys. **143**, 99-114 (1991)

[123] A.S. Holevo, *Statistical decisions in quantum theory*, J. Multivar. Anal. **3**, 337-394 (1973)

[124] A.S. Holevo, *Remarks on optimal quantum measurements*, Problemy Peredachi Informatsii **10**, 51-55 (1974)

[125] A.S. Holevo, *On Asymptotically Optimal Hypothesis Testing in Quantum Statistics*, Theory Probab. Appl. **23**, 411-415 (1979)

[126] K. Hornberger and J. Sipe, *Collisional decoherence reexamined*, Phys. Rev. A **68**, 012105 (2003)

[127] M. Horodecki, P. Horodecki, and R. Horodecki, *Separability of mixed states: necessary and sufficient conditions*, Phys. Lett. A **223**, 1-8 (1996)

[128] P. Horodecki, *Separability criterion and inseparable mixed states with positive partial transposition*, Phys. Lett. A **232**, 333-339 (1997)

[129] R. Horodecki and M. Horodecki, *Information-theoretic aspects of inseparability of mixed states*, Phys. Rev. A **54**, 1838-1843 (1996)

[130] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Quantum entanglement*, Rev. Mod. Phys. **81**, 865-942 (2009)

[131] Y.P. Huang and M.G. Moore, *Creation, detection, and decoherence of macroscopic quantum superposition states in double-well Bose-Einstein condensates*, Phys. Rev. A **73**, 023606 (2006)

[132] M. Hübner, *Explicit computation of the Bures distance for density matrices*, Phys. Lett. A **163**, 239-242 (1992)

[133] P. Hyllus, O. Gühne, and A. Smerzi, *Not all pure entangled states are useful for sub-shot-noise interferometry*, Phys. Rev. A **82**, 012337 (2010)

[134] P. Hyllus, L. Pezzé, and A. Smerzi, *Entanglement and sensitivity in precision measurements with states of a fluctuating number of particles*, Phys. Rev. Lett. **105**, 120501 (2010)

[135] P. Hyllus, W. Laskowski, R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, L. Pezzé, and A. Smerzi, *Fisher information and multiparticle entanglement*, Phys. Rev. A **85**, 022321 (2012)

[136] I.D. Ivanovic, *How to differentiate between non-orthogonal states*, Phys. Lett. A **123**, 257-259 (1987)

[137] M.W. Jack, *Decoherence due to three-body loss and its effect on the state of a Bose-Einstein Condensate*, Phys. Rev. Lett. **89**, 140402 (2002)

[138] M.W. Jack, *Effect of atom loss on collapse and revivals of phase coherence in small atomic samples*, Phys. Rev. A **67**, 043612 (2003)

[139] G. Jaeger and A. Shimony, *Optimal distinction between two non-orthogonal quantum states*, Phys. Lett. A **197**, 83-87 (1995)

[140] V. Jakšić, Y. Ogata, C.-A. Pillet, and R. Seiringer, *Quantum hypothesis testing and non-equilibrium statistical mechanics*, Rev. Math. Phys. **24**(6), 1230002 (2012)

[141] V. Jakšić and C.-A. Pillet, *Entropic Functionals in Quantum Statistical Mechanics*, in *Proceedings of XVIIth International Congress of Mathematical Physics, Aalborg 2012* (World Scientific, Singapore, 2013), pp. 336-343

[142] V. Jakšić, C.-A. Pillet, and M. Westrich, *Entropic fluctuations of quantum dynamical semigroups*, J. Stat. Phys. **154**, 153-187 (2014)

[143] D. Jonathan and M.B. Plenio, *Entanglement-assisted local manipulation of pure quantum states*, Phys. Rev. Lett. **83**, 3566-3569 (1999)

[144] R. Jozsa, *Fidelity for mixed quantum states*, J. Mod. Opt. **41**, 2315-2323 (1994)

[145] R. Jozsa and N. Linden, *On the role of entanglement in quantum-computational speed-up*, Proc. R. Soc. Lond. A **459**, 2011-2032 (2003)

[146] M. Kitagawa, and M. Ueda, *Squeezed spin states*, Phys. Rev. A **47**, 5138–5143 (1993)

[147] E. Knill and R. Laflamme, *Power of One Bit of Quantum Information*, Phys. Rev. Lett. **81**, 5672 (1998)

[148] M. Koashi and A. Winter, *Monogamy of quantum entanglement and other correlations*, Phys. Rev. A **69**, 022309 (2004)

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

[150] H. Krauter, C.A. Muschik, K. Jensen, W. Wasilewski, J.M. Petersen, I. Cirac, and E. Polzik, *Entanglement generated by dissipation and steady state entanglement of two macroscopic objects*, Phys. Rev. Lett. **107**, 080503 (2011)

[151] F. Kubo and T. Ando, *Means of positive linear operators*, Math. Ann. **246**, 205-224 (1980)

[152] B.P. Lanyon, M. Barbieri, M.P. Almeida, and A.G. White, *Experimental Quantum Computing without Entanglement*, Phys. Rev. Lett. **101**, 200501 (2008)

[153] E.H. Lieb, *Convex trace functions and the Wigner-Yanase-Dyson conjecture*, Adv. Math. **11**, 267-288 (1973)

[154] E.H. Lieb and M.B. Ruskai, *Proof of the strong subadditivity of quantum-mechanical entropy*, J. Math. Phys. **14**, 1938-1941 (1973)

[155] E.H. Lieb and W. Thirring, *Inequalities for the moments of the eigenvalues of the Schrödinger Hamiltonian and their relation to Sobolev inequalities*, in *Studies in mathematical physics: essays in honor of Valentine Bargman*, E.H. Lieb, B. Simon, A.S. Wightman Eds. (Princeton University Press, Princeton, 1976), pp. 269-297

[156] G. Lindbald, *An entropy inequality for quantum measurements*, Commun. Math. Phys. **28**, 245-249 (1972)

[157] G. Lindblad, *Expectations and Entropy Inequalities for Finite Quantum Systems*, Commun. Math. Phys. **39**, 111-119 (1974)

[158] G. Lindblad, *Completely Positive Maps and Entropy Inequalities*, Commun. Math. Phys. **40**, 147-151 (1975)

[159] G. Lindblad, *On the Generators of Quantum Dynamical Semigroups*, Commun. Math. Phys. **48**, 119 (1976)

[160] S. Luo, *Using measurement-induced disturbance to characterize correlations as classical or quantum*, Phys. Rev. A **77**, 022301 (2008)

[161] S. Luo, *Quantum discord for two-qubit systems*, Phys. Rev. A **77**, 042303 (2008)

[162] S. Luo and S. Fu, *Geometric measure of quantum discord*, Phys. Rev. A **82**, 034302 (2010)

[163] V. Madhok and A. Datta, *Interpreting quantum discord through quantum state merging*, Phys. Rev. A **83**, 032323 (2011)

[164] V. Madhok and A. Datta, *Role of quantum discord in quantum communication*, arXiv:1107.09994 [quant-ph]

[165] E. Mascarenhas, D. Cavalcanti, V. Vedral, and M. França Santos, *Physically realizable entanglement by local continuous measurements*, Phys. Rev. A **83**, 022311 (2011)

[166] L. Mazzola, S. Maniscalco, J. Piilo, K.-A. Suominen, and B.M. Garraway, *Sudden death and sudden birth of entanglement in common structured reservoirs*, Phys. Rev. A **79**, 042302 (2009)

[167] L. Mazzola, J. Piilo, and S. Maniscalco, *Sudden Transition between Classical and Quantum Decoherence*, Phys. Rev. Lett. **104**, 200401 (2010)

[168] G.J. Milburn, J. Corney, E.M. Wright, and D.F. Walls, *Quantum dynamics of an atomic Bose-Einstein condensate in a double-well potential*, Phys. Rev. A **55**, 4318-4325 (1997)

[169] K. Modi, T. Parerek, W. Son, V. Vedral, and M. Williamson, *Unified View of Quantum and Classical Correlations*, Phys. Rev. Lett. **104**, 080501 (2010)

[170] K. Modi, A. Brodutch, H. Cable, T. Paterek, and V. Vedral, *The classical-quantum boundary for correlations: Discord and related measures*, Rev. Mod. Phys. **84**, 1655-1707 (2012)

[171] M. Mohseni, A.M. Steinberg, and J.A. Bergou, *Optical Realization of Optimal Unambiguous Discrimination for Pure and Mixed Quantum States*, Phys. Rev. Lett. **93**, 200403 (2004)

[172] C. Monroe, D.M. Meekhof, B.E. King, and D.J. Wineland, *A "Schrödinger cat" superposition state of an atom*, Science **272**, 1131-1136 (1996)

[173] A. Montanaro, *A lower bound on the probability of error in quantum state discrimination*, in *Proceedings of IEEE Information Theory Workshop ITW'08* (IEEE, Piscataway, NJ, 2008), p.378

[174] E.A. Morozova and N.N. Chentsov, *Markov invariant geometry on state manifolds* (in Russian), Itogi Nauki i Tekhniki **36**, 69-102 (1990)

[175] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, *On quantum Rényi entropies: a new generalization and some properties*, J. Math. Phys. **54**, 122203 (2013)

[176] W. Nagourney, J. Sandberg, and H. Dehmelt, *Shelved optical electron amplifier: observation of quantum jumps*, Phys. Rev. Lett. **56**, 2797-2799 (1986)

[177] T. Nakano, M. Piani, and G. Adesso, *Negativity of quantumness and its interpretations*, Phys. Rev. A **88**, 012117 (2013)

[178] H. Nha and H.J. Carmichael, *Entanglement within the Quantum Trajectory Description of Open Quantum Systems*, Phys. Rev. Lett. **93**, 120408 (2004)

[179] M.A. Nielsen, *Conditions for a Class of Entanglement Transformations*, Phys. Rev. Lett. **82**, 436-439 (1999)

[180] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Information* (Cambridge University Press, 2000)

[181] M. Nussbaum and A. Szkola, *The Chernoff lower bound for symmetric quantum hypothesis testing*, The Annals of Statistics **37** (Institute of Mathematical Statistics, 2009), pp. 1040-1057

[182] T. Ogawa and H. Nagaoka, *Strong converse and Stein's lemma in quantum hypothesis testing*, IEEE Trans. Inf. Theory **46**(7), 2428-2433 (2000)

[183] M. Ohya and D. Petz, *Quantum entropy and its use* (Springer-Verlag, Berlin/Heidelberg, 1993)

[184] H. Ollivier and W.H. Zurek, *Quantum Discord: A Measure of the Quantumness of Correlations*, Phys. Rev. Lett. **88**, 017901 (2001)

[185] M. Orszag, *Quantum Optics*, 2nd Ed. (Springer, 2008)

[186] M. Orszag and M. Hernandez, *Coherence and Entanglement in two-qubit systems*, Adv. in Optics and Photonics **2**, 229-286 (2010)

[187] M. Ozawa, *Entanglement measures and the Hilbert-Schmidt distance*, Phys. Lett. A **268**, 158-160 (2000)

[188] G.M. Palma, K.-A. Suominen, and A.K. Ekert, *Quantum Computers and Dissipation*, Proc. Roy. Soc. Lond. A **452**, 567-584 (1996)

[189] P.F. Palmer, *The singular coupling and weak coupling limits*, J. Math. Phys. **18**, 527 (1977)

[190] G. Passante, O. Moussa, D.A. Trottier, and R. Laflamme, *Experimental detection of nonclassical correlations in mixed-state quantum computation*, Phys. Rev. A **84**, 044302 (2011)

[191] A. Peres, *How to differentiate between non-orthogonal states*, Phys. Lett. A **128**, 19 (1988)

[192] A. Peres, *Neumark's theorem and Quantum Inseparability*, Foundations of Physics **20**, 1441-1453 (1990)

[193] A. Peres, *Separability Criterion for Density Matrices*, Phys. Rev. Lett. **77**, 1413-1415 (1996)

[194] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer Academic Publisher, 1995)

[195] D. Petz, *Monotone Metrics on Matrix Spaces*, Lin. Alg. and its Appl. **244**, 81-96 (1996)

[196] D. Petz, *Monotonicity of quantum relative entropy revisited*, Rev. Math. Phys. **15**, 79-91 (2003)

[197] L. Pezzé and A. Smerzi, *Entanglement, Nonlinear Dynamics, and the Heisenberg limit*, Phys. Rev. Lett. **102**, 100401 (2009)

[198] M. Piani, *Problem with geometric discord*, Phys. Rev. A **86**, 034101 (2012)

[199] M.B. Plenio and P.L. Knight, *The quantum-jump approach to dissipative dynamics in quantum optics.*, Rev. Mod. Phys. **70**, 101 (1998)

[200] D. Qiu and L. Li, *Minimum-error discrimination of quantum states: Bounds and comparison*, Phys. Rev. A **81**, 042329 (2010)

[201] A.G. Redfield, *On the Theory of Relaxation Processes*, IBM J. Res. Dev. **1**, 19 (1957)

[202] F. Riedel, P. Böhi, Yun Li, T.W. Hänsch, A. Sinatra, and P. Treutlein, *Atom-chip-based generation of entanglement for quantum metrology*, Nature **464**, 1170 (2010)

[203] W. Roga, S.M. Giampaolo, and F. Illuminati, *Discord of response*, J. Phys. A: Math. Gen. **47**, 365301 (2014)

[204] T. Rudolph, R.W. Spekkens, and P.S. Turner, *Unambiguous discrimination of mixed states*, Phys. Rev. A **68**, 010301 (2003)

[205] M.B. Ruskai, *Beyond strong subadditivity: improved bounds on the contraction of the generalized relative entropy*, Rev. Math. Phys. **6**(5a), 1147-1161 (1994)

[206] A. Sanpera, R. Tarrach, and G. Vidal, *Quantum inseparability as local pseudomixture*, Phys. Rev. A **58**, 826-830 (1998)

[207] M. Sasaki, K. Kato, M. Izutsu, and O. Hirota, *Quantum channels showing superadditivity in classical capacity*, Phys. Rev. A **58**, 146-158 (1998)

[208] A. Sawicki, A. Huckleberry, and M. Kuś, *Symplectic Geometry of Entanglement*, Commun. Math. Phys. **305**, 441-468 (2011)

[209] M.H. Schleier-Smith, I.D. Leroux, and V. Vuletić, *States of an Ensemble of Two-Level Atoms with Reduced Quantum Uncertainty*, Phys. Rev. Lett. **104**, 073604 (2010)

[210] M. Schlosshauer, *Decoherence, the measurement problem, and interpretations of quantum mechanics*, Rev. Mod. Phys. **76**, 1267 (2004)

[211] E. Schrödinger, *Die gegenwärtige Situtation in der Quantenmechanik*, Naturwiss. **23**, 807-812, 823-828, 844-849 (1935); English translation in [258]

[212] L.S. Schulman, *Techniques and Applications of Path Integration* (Wiley Classics Library, New York, 1996)

[213] B.W. Schumacher, *Sending entanglement through noisy quantum channels*, Phys. Rev. A **54**, 2614-2628 (1996)

[214] B.W. Schumacher and M.A. Nielsen, *Quantum data processing and error correction*, Phys. Rev. A **54**, 2629 (1996)

[215] C.E. Shannon, *A Mathematical Theory of Communication*, Bell Syst. Tech. J. **27**, 379-423 and 623-656 (1948)

[216] A. Shimony, *Degree of entanglement*, Ann. N.Y. Acad. Sci. **755**, 675-679 (1995)

[217] P.W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM J. Comp. **26**, 1484-1509 (1997)

[218] A. Sinatra and Y. Castin, *Phase dynamics of Bose-Einstein condensates: losses versus revivals*, Eur. Phys. J. D **4**, 247-260 (1998)

[219] A. Sinatra, J.-C. Dornstetter, and Y. Castin, *Spin squeezing in Bose-Einstein condensates: Limits imposed by decoherence and non-zero temperature*, Front. Phys. **7**, 86-97 (2012)

[220] H.-J. Sommers and K. Życzkowski, *Bures volume of the set of mixed quantum states*, J. Phys. A: Math. Gen. **36**, 10083-10100 (2003)

[221] D. Spehner and J. Bellissard, *A kinetic model of quantum jumps*, J. Stat. Phys. **104**, 525-566 (2001)

[222] D. Spehner and M. Orszag, *Quantum jump dynamics in cavity QED*, J. Math. Phys. **43**, 3511-3537 (2002)

[223] D. Spehner and F. Haake, *Decoherence bypass of macroscopic superpositions in quantum measurement*, J. Phys. A: Math. Theor. **41**, 072002 (2008)

[224] D. Spehner and F. Haake, *Quantum measurements without macroscopic superpositions*, Phys. Rev. A **77**, 052114 (2008)

[225] D. Spehner, *Models for quantum measurements*, Lecture notes (Winter school "Aspects of quantum dynamics", Grenoble, Nov. 2008), http://www-fourier.ujf-grenoble.fr/~spehner/notes_de_cours_q_meas6.pdf

[226] D. Spehner and M. Orszag, *Geometric quantum discord with Bures distance*, New J. Phys. **15**, 103001 (2013)

[227] D. Spehner and M. Orszag, *Geometric quantum discord with Bures distance: the qubit case*, J. Phys. A: Math. Theor. **47**, 035302 (2014)

[228] D. Spehner, K. Pawlowski, G. Ferrini, and A. Minguzzi, *Effect of one-, two-, and three-body atom loss processes on superpositions of phase states in Bose-Josephson junctions*, Eur. Phys. J. B **87**, 157 (2014)

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

[230] D. Stoler, *Generalized Coherent States*, Phys. Rev. D **4**, 2309-2312 (1971)

[231] A. Streltsov, H. Kampermann, and D. Bruß, *Linking a distance measure of entanglement to its convex roof*, New J. Phys. **12**, 123004 (2010)

[232] W.T. Strunz, *Open system dynamics with non-Markovian quantum trajectories*, Phys. Rev. Lett. **82**, 1801-1805 (1999)

[233] W.T. Strunz, F. Haake, and D. Braun, *Universality of decoherence for macroscopic quantum superpositions*, Phys. Rev. A **67**, 022101 (2003)

[234] M.O. Terra Cunha, *The geometry of entanglement sudden death*, New J. Phys. **9**, 237 (2007)

[235] P.E. Toscheck and C. Wunderlich, *What does an observed quantum system reveal to its observer?*, Eur. Phys. J. D **14**, 387396 (2001)

[236] G. Toth, *Multipartite entanglement and high-precision metrology*, Phys. Rev. A **85**, 022322 (2012)

[237] J. Tyson, *Two-sided estimates of minimum-error distinguishability of mixed quantum states via generalized Holevo-Curlander bounds*, J. Math. Phys. **50**, 032106 (2009)

[238] J. Tyson, *Two-sided bounds on minimum-error quantum measurement, on the reversibility of quantum dynamics, and on the maximum overlap problem using directional iterates*, J. Math. Phys. **51**, 092204 (2010)

[239] A. Uhlmann, *Endlich-dimensionale Dichtematrizen II*, Wiss. Z. Karl-Marx-Univ. Leipzig, Math.-Nat R. **22**, 139-177 (1973)

[240] A. Uhlmann, *The "transition probability" in the state space of a $*$-algebra*, Rep. Math. Phys. **9**, 273-279 (1976)

[241] A. Uhlmann, *Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory*, Commun. Math. Phys. **54**, 21-32 (1977)

[242] A. Uhlmann, *Parallel transport and "quantum holonomy" along density operators*, Rep. Math. Phys. **24**, 229-240 (1986)

[243] A. Uhlmann, *Geometric phases and related structures*, Rep. Math. Phys. **36**, 461-481 (1995)

[244] U. Umegaki, *Conditional expectations in an operator algebra IV (entropy and information)*, Kodai Math. Sem. Rep. **14**, 59-85 (1962)

[245] V. Vedral, M.B. Plenio, M.A. Rippin, and P.L. Knight, *Quantifying Entanglement*, Phys. Rev. Lett. **78**, 2275-2279 (1997)

[246] V. Vedral and M.B. Plenio, *Entanglement measures and purifications procedures*, Phys. Rev. A **57**, 1619-1633 (1998)

[247] A.F. Verbeure, in *Encyclopedia of mathematical physics* Vol. **4**, J.-P. Françoise, G.L. Naber, and T.S. Tsou Eds. (Elsevier, 2006), p.130

[248] G. Vidal, *Entanglement monotones*, J. Mod. Opt. **47**, 355-376 (2000)

[249] C. Viviescas, I. Guevara, A.R.R. Carvalho, M. Busse, and A. Buchleitner, *Entanglement dynamics in open two-qubit systems via diffusive quantum trajectories*, Phys. Rev. Lett. **105**, 210502 (2010)

[250] B. Vlastakis, G. Kirchmair, Z. Leghtas, S.E. Nigg, L. Frunzio, S.M. Girvin, M. Mirrahimi, M.H. Devoret, and R.J. Schoelkopf, *Deterministically encoding quantum information using 100-Photon Schrödinger cat states*, Science **342**, 607-610 (2013)

[251] S. Vogelsberger and D. Spehner, *Average entanglement for markovian quantum trajectories*, Phys. Rev. A **82**, 052327 (2010)

[252] S. Vogelsberger, *Dynamique des systèmes quantiques ouverts : décohérence et perte d'intrication* (in French), PhD thesis, Université Joseph Fourier, Grenoble (2012)

[253] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1955)

[254] W. Wasilewski, K. Jensen, H. Krauter, J.J. Renema, M.V. Balabas, and E.S. Polzik, *Quantum Noise Limited and Entanglement-Assisted Magnetometry*, Phys. Rev. Lett. **104**, 133601 (2010)

[255] T.C. Wei and P.M. Goldbart, *Geometric measure of entanglement and applications to bipartite and multipartite quantum states*, Phys. Rev. A **68**, 042307 (2003)

[256] U. Weiss, *Quantum Dissipative Systems* (World Scientific, Singapore, 1993)

[257] R.F. Werner, *Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model*, Phys. Rev. A **40**, 4277-4281 (1989)

[258] J.A. Wheeler and W.H. Zurek Eds., *Quantum Theory and Measurement* (Princeton University Press, Princeton NJ, 1983)

[259] E. Wigner, *The problem of measurement*, Am. J. Phys. **31**, 6 (1963)

[260] M.M. Wilde, A. Winter, D. Yang, *Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy*, Commun. Math. Phys. **331**, 593-622 (2014)

[261] H.M. Wiseman and G.J. Milburn, *Quantum theory of field-quadrature measurements*, Phys. Rev. A **47**, 642 (1993)

[262] H.M. Wiseman and L. Diósi, *Complete parametrization, and invariance, of diffusive quantum trajectories for Markovian open systems*, Chem. Phys. **268**, 91-104 (2001)

[263] M.M. Wolf, *Quantum Channels and Operations Guided Tour* (2002), see http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf

[264] W.K. Wootters, *Statistical distance and Hilbert space*, Phys. Rev. D **23**, 357-362 (1981)

[265] W.K. Wootters, *Entanglement of Formation of an Arbitrary State of Two Qubits*, Phys. Rev. Lett. **80**, 2245 (1998)

[266] S.L. Woronowicz, *Positive maps of low dimensional matrix algebras*, Rep. Math. Phys. **10**, 165-183 (1976)

[267] T. Yu, L. Diósi, N. Gisin, and W.T. Strunz, *Non-markovian quantum-state diffusion: Perturbation approach*, Phys. Rev. A **60**, 91-103 (1999)

[268] T. Yu and J.H. Eberly, *Finite-Time Disentanglement Via Spontaneous Emission*, Phys. Rev. Lett. **93**, 140404 (2004)

[269] H.P. Yuen, R.S. Kennedy, and M. Lax, *Optimum testing of multiple hypotheses in quantum detection theory*, IEEE Trans. Inf. Theory **21**, 125-134 (1975)

[270] Yun Li, Y. Castin, and A. Sinatra, *Optimum spin squeezing in Bose-Einstein condensates with particle losses*, Phys. Rev. Lett. **100**, 210401 (2008)

[271] Yun Li, P. Treutlein, J. Reichel, and A. Sinatra, *Spin squeezing in a bimodal condensate: spatial dynamics and particle losses*, Eur. Phys. J. B **68**, 365-381 (2009)

[272] B. Yurke and D. Stoler, *Generating quantum mechanical superpositions of macroscopically distinguishable states via amplitude dispersion*, Phys. Rev. Lett. **57**, 13-16 (1986)

[273] B. Yurke, S.L. McCall, and J.R. Klauder, *SU(2) and SU(1,1) interferometers*, Phys. Rev. A **33**, 4033-4054 (1986)

[274] W.M. Zhang, D.H. Feng, and R. Gilmore, *Coherent states: Theory and some applications*, Rev. Mod. Phys. **62**, 867 (1990)

[275] W.H. Zurek, *Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse?*, Phys. Rev. D **24**, 1516-1525 (1981)

[276] W.H. Zurek, *Decoherence and the transition from quantum to classical*, Physics Today **44** (10), 36-44 (1991); e-print: quant-ph/0306072 (updated version)

[277] W.H. Zurek, *Decoherence, einselection, and the quantum origins of the classical*, Rev. Mod. Phys. **75**, 715-775 (2003)

[278] K. Życzkowski, P. Horodecki, M. Horodecki, and R. Horodecki, *Dynamics of quantum entanglement*, Phys. Rev. A **65**, 012101 (2001)