Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea Magistrale in Fisica

Entanglement detection measuring classical correlations

Relatore:

Prof. Cristian Degli Esposti Boschi

Correlatore:

Prof. Matteo G. A. Paris

Presentata da:

Marina Menga

Sessione III

Anno Accademico 2015/2016

Contents

Abstract

Un sistema fisico presenta *entanglement* quando mostra proprietà globali che non sono riducibili a proprietà locali e/o a correlazioni che possono essere stabilite mediante un canale di comunicazione. L'entanglement é un fenomeno di natura puramente quantistica ed é considerato la principale risorsa dei protocolli quantistici di manipolazione e trasmissione di informazione, in particolare del miglioramento delle prestazioni che questi protocolli possono offrire rispetto alle loro controparti classiche. Lo studio e la caratterizzazione dell'entanglement rivestono dunque un ruolo cruciale nello sviluppo della teoria dell'informazione quantistica e, piú in generale, nella comprensione delle strutture fondamentali della meccanica quantistica non relativistica.

Uno dei principali problemi legati alla teoria dell'entanglement é quello di determinare se un dato sistema quantistico sia entangled, ovvero quello di trovare dei criteri ottimali per la rilevazione, la caratterizzazione e la quantificazione dell'entanglement. Ad oggi, non esiste una soluzione completa per questo problema, ovvero non esiste un criterio generale per discriminare gli stati entangled da quelli che non sono entangled (detti *separabili*) per sistemi descritti in spazi di Hilbert di dimensione arbitraria.

In questo lavoro di tesi, abbiamo esaminato e generalizzato un criterio per la rilevazione dell'entanglement basato sulla misura delle correlazioni classiche esistenti tra osservabili locali complementari di un sistema bipartito [1]. In particolare, ci siamo occupati dei criteri per il riconoscimento dell'entanglement che si basano sulla misura dell'informazione mutua, e abbiamo analizzato le prestazioni di questo criterio per una generica coppia di osservabili, non necessariamente complementari. Le prestazioni del nuovo criterio, in termini di efficienza, robustezza ed applicabilità, sono state analizzate in dettaglio per sistemi bipartiti di qubit, poiché in questo caso l'entanglement é completamente caratterizzato e sono a disposizione strumenti di confronto. In particolare, il criterio é stato utilizzato per la rivelazione di entanglement in presenza di rumore esterno e decoerenza.

I risultati ottenuti mostrano che il caso della misura di osservabili complementari da parte dei due osservatori locali é quello ottimale. Questa configurazione non é peró sempre realizzabile sperimentalmente: l'analisi condotta in questa tesi ha mostrato che il criterio generalizzato é robusto, dal momento che la percentuale di stati rivelati non cala in maniera drammatica quando ci si sposta dalla condizione di complementarietà. Inoltre, il metodo basato sull'informazione mutua, pur non avendo performance ottimali rispetto ad altri metodi di rilevazione dell'entanglement, come ad esempio i cosiddetti *entanglement witness*, é interessante dal punto di vista applicativo, poiché la classe di stati che rivela appartengono a classi differenti da quelle rivelate da altri metodi.

Dal punto di vista dell'applicabilità, il criterio permette di caratterizzare l'entanglement con un numero di misure minore rispetto a quelle richieste per avere completa tomografia dello stato o per ricostruire i piú comuni entanglement witness.

Globalmente, il metodo proposto si é rivelato affidabile e complessivamente efficiente. Sono in corso contatti con gruppi sperimentali per la sua applicazione in sistemi ottico-quantistici discreti, ovvero con qubit codificati nella polarizzazione di singoli fotoni.

Chapter 1

Introduction

Many physicists like to assert that "the world is quantum mechanical". This statement, though sounding very radical, reflects indeed the crucially relevant role that quantum theory plays in the description of reality.

Indeed, natural phenomena which occur at atomic and subatomic scale cannot be explained outside the dominion of quantum physics and even when concerned with the macroscopic objects encountered in everyday life it is necessary to analyse the behaviour of their microscopic constituents, in order to obtain a consistent and complete scientific description.

It is therefore undeniable that the quantum view of reality represents one of the milestones both at the fundamental level of knowledge and a solid basis for incredible experimental applications.

Additionally, quantum mechanics has the particular feature of being at the same time the most successful and the most puzzling of the scientific theories to this day (as Richard Feynman said, *"I think I can safely say that nobody understands quantum mechanics"* [2]).

It is moreover regarded as the most precisely tested theory in the history of science, and one is logically brought to think that what is predicted by quantum formalism is supposed to mirror what happens in laboratory.

Singularly, however, what is generally regarded as the "essence" of quantum formalism, that is **entanglement**, was originally recognised theorically in 1935 but had to wait for over seventy years to be intoduced into the laboratory praxis as a new experimental resource, with a solid connotation of reality, as concrete as matter or energy [3]. Finally, only at the end of the century, entanglement was also equipped with a consistent physical interpretation, on the grounds of the rapidly developing field of quantum information theory, providing answers to fundamental questions of quantum physics and at the same time creating many new open problems in theoretical physics, that are still waiting to be solved.

This singular property of composite quantum systems is a potential resource for many quantum processes, including quantum cryptography, quantum teleportation and

dense coding, and therefore remained a central topic in quantum information theory to the present day. However, since the beginning of the investigation of the entanglement phenomenon, it was clear that this new physical resource is as useful and interesting as complex and difficult to detect and measure.

Therefore a great effort has been put over the years in looking for conceptual and mathematical tools in order to decipher its complex structure, trying to give an answer to the crucial problem of proving that a given quantum state is, or is not, entangled, in terms of its characterization, detection, generation and quantification.

Generally, one looks for an entanglement verification procedure who respects several reasonable properties [4]. First of all, it is crucial to consider procedural schemes that are easy to implement in the experimental context, robust against noise and able to detect also weakly entangled states. In addition, the final conclusion that a certain state was or was not entangled should not depend on some assumptions about the nature or the form of the state. Hence, one has to find the measurements that allow to conclude as much as possible about the entanglement content of the given state.

In the present work we are going to review a particular criterion for the characterization of entanglement in the most simple example of a bipartite quantum system, i.e. a two-qubit system, based on the classical correlations existing between the measurement outcomes of the observables related to the local parts of the system.

This criterion was introduced in 2015 by Maccone et al. in [1], as a detecting proposal in order to provide an interpretation of entanglement, rather than to introduce a new entanglement detection scheme, in terms of the classical correlations for *complementary* observables. We mention that in 2014 a similar approach using different measures of correlation was introduced in [5].

Interestingly, they showed that, comparing their results to those obtained with an entanglement scheme based on the use of witness operators [6], the classical correlations method performed better than entanglement witnesses that employ the same measurements and moreover allow the detection of many entangled states that entanglement witness miss.

Moreover, this criterion can be characterized with fewer and simpler measurements than those needed for full tomography. In fact, for the case of a two-qubit system that we are going to consider in what follows, the criterion based on classical correlations would only require four local measurements, since it only entails independent measurements of two local observables on the two systems.

These four measurements are less compared to quantum state tomography that would require measurements of d^4-1 observables [3] (i.e., in the case of $d = 2$, 15 measurements), plus one for the normalization, for a total of 16 measurements.

There is a conceptual difference between entanglement detection and entanglement criteria. The first one refers to a measurement procedure, set to determine whether a given state is entangled or not, using measurements that do not depend on the state. The entanglement criteria, on the other hand, refer to a general characterization of entanglement, and this is the case of the proposal by Maccone et al. [1], that turned out to be quite effective in this sense. However, even if it does not represent an optimal criterion, it can provide a good tool for characterizing entangled states and, as we already underlined, to give an interpretation of entanglement.

In the present thesis work, we take on the direction given in [1], i.e. focusing on the classical correlations between complementary observables, and ask what would happen if we relaxed the complementarity assumption, namely we are going to see how and how much the "goodness" of this criterion will change if we consider non-complementary observables.

1.1 Notation

According to the postulates of quantum mechanics, the states of a physical system correspond to vectors $|\psi\rangle$ of an Hilbert space $\mathcal{H}$, which satisfy the normalization condition $\langle\psi|\psi\rangle = 1$. Consequently, composite systems are described by the tensor product of the individual Hilbert spaces corresponding to each subsystem, and the overall state of the system is a vector in the global space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots$. In this picture is introduced the *superposition principle*, stating that if $|\psi\rangle_1$ and $|\psi\rangle_2$ are possible states of a system, then any normalized linear combination of the two states is also an admissible state of the system. [7]

Moreover, observable quantities are described by Hermitian operators, i.e. operators such that $X = X^\dagger$, which admit a spectral decomposition of the form $X = \sum_x xP_x$. Hence the operator X is expressed in terms of its real eigenvalues x, namely the possible values of the observables, and of the projectors $P_x = |x\rangle\langle x|$ (that satisfy the orthogonality property $P_xP_{x'} = \delta_{xx'}P_x$) on its eigenvectors (according to the eigenvalues equation $X|x\rangle = x|x\rangle$) which form a basis for the Hilbert space. The eigenvalues $|x\rangle$ moreover satisfy the orthonormality property ($\langle x|x'\rangle = \delta_{xx'}$) and the completeness property ($\sum_x |x\rangle\langle x| = \mathbb{I}$), namely they form a complete set of orthonormal states. The linear space of (linear) operators from $\mathcal{H}$ to $\mathcal{H}$ is still an Hilbert space, with scalar product provided by the trace operation, i.e. $\langle A|B\rangle = \mathrm{Tr}[A^\dagger B]$. Then we can introduce the *Born rule* and define the probability of obtaining the outcome x from the measurement of the observable X as:

$$\begin{aligned} p_x &= |\langle\psi|x\rangle|^2 = \langle\psi|x\rangle\langle x|\psi\rangle = \langle\psi|P_x|\psi\rangle \\ &= \sum_n \langle\psi|\phi_n\rangle\langle\phi_n|P_x|\psi\rangle = \sum_n \langle\phi_n|P_x|\psi\rangle\langle\psi|\phi_n\rangle = \mathrm{Tr}[|\psi\rangle\langle\psi| P_x] \end{aligned}$$

and the overall expectation value as:

$$\langle X\rangle = \langle\psi|X|\psi\rangle = \mathrm{Tr}[|\psi\rangle\langle\psi| X].$$

The state of the system after the measurement is the projection of the state before the measurement on the eigenspace of the observed eigenvalue, namely:

$$|\psi_x\rangle = \frac{1}{\sqrt{p_x}} P_x |\psi\rangle .$$

Finally, the dinamical evolution of a physical system is ruled by unitary operators. Given $|\psi_0\rangle$ the initial state of the system at time t_0, the state of the system at time t is given by $|\psi_t\rangle = U(t, t_0)|\psi_0\rangle$, with $UU^\dagger = U^\dagger U = \mathbb{I}$.

However the general postulates just summarized are valid for closed and isolated systems. In what follows, as well as in the great majority of physical systems, we are

dealing with systems that are not isolated and interact with other systems or, in general, are subsystems of the universe and cannot be considered separately. As a consequence, a reformulation of the postulates is required, suitable for the description of any measurement on a quantum system, including those involving external ancillas or noisy environments.

Let us consider a quantum system whose preparation is not completely under control, but we only know that it is prepared in the state $|\psi_k\rangle$ with probability p_k. We say than that such system is described by the statistical ensemble $\{p_k, |\psi_k\rangle\}$, with $\sum_k p_k = 1$ and the states $|\psi_k\rangle$ are in general not orthogonal.

We can thus evaluate the expected value of an observable X in the following way:

$$\begin{aligned}\langle X\rangle &= \sum_k p_k \langle X\rangle_k = \sum_k p_k \langle\psi_k|X|\psi_k\rangle = \sum_{nmk} p_k \langle\psi_k|\phi_n\rangle\langle\phi_n|X|\phi_m\rangle\langle\phi_m|\psi_k\rangle \\ &= \sum_{nmk} p_k \langle\phi_m|\psi_k\rangle\langle\psi_k|\phi_n\rangle\langle\phi_n|X|\phi_m\rangle = \sum_{nm}\langle\phi_m|\rho|\phi_n\rangle\langle\phi_n|X|\phi_m\rangle \\ &= \sum_m \langle\phi_m|\rho X|\phi_m\rangle = \mathrm{Tr}[\rho X]\end{aligned}$$

where

$$\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k|$$

is defined as the statistical *density operator* of our system. A density operator associated to an ensemble $\{p_k, |\psi_k\rangle\}$ is defined as a positive (hence selfadjoint) operator with unit trace, i.e. $\mathrm{Tr}[\rho] = 1$.

If we denote with ρ the state before the measurement, we have that any observable quantity is associated to an Hermitian operator X with spectral decomposition $X = \sum_x x |x\rangle\langle x|$, and the eigenvalues are real and conventionally non-degenerate. The eigenvectors $|x\rangle$ form a basis for the Hilbert space and the corresponding projectors $P_x = |x\rangle\langle x|$ span the entire Hilbert space and satisfy to $\sum_x P_x = \mathbb{I}$.

As we said, the projectors are orthogonal and hence $P_x^2 = P_x$. It follows that the eigenvalues of any projector are 0 and 1.

A measurement on X yelds one of the eigenvalues x as a possible outcome, and the probability that a particular outcome is found as the measurement result is given by the aforementioned Born rule, given by:

$$p_x = \mathrm{Tr}[P_x \rho P_x] = \mathrm{Tr}[\rho P_x^2] = \mathrm{Tr}[\rho P_x],$$

where the probabilities p_x are non-negative ($0 \le p_x \le 1$) and normalized ($\sum_x p_x = 1$).

Moreover, the state after the measurement where the outcome x has been obtained is given by the *reduction rule* or *projection postulate*, namely:

$$\rho_x = \frac{1}{p_x} P_x \rho P_x$$

or, for a measurement in which we did not record the result:

$$\tilde{\rho} = \sum_x p_x \rho_x = \sum_x P_x \rho P_x.$$

The conceptual meaning under this generalization is that the measurement process is random and we cannot predict its outcome with precision, but we can only predict a *spectrum* of the possible outcomes together with the probability that a given outcome is found in an actual measurement. Indeed, we note that the state ρ does not represent a single system, but rather an *ensemble* of identically prepared systems: if we perform the same measurement on each element of the ensemble we can predict the probabilities with which every possible outcome would occur, but we cannot predict the result of each individual measurement (except the extreme cases of probability 0 or 1).

We notice that the number of the possible outcomes of a measurement is limited by the number of terms in the orthogonal resolution of the identity, which is itself never greater than the dimensionality of the Hilbert space. However it would be often desiderable to have more outcomes than this limited number while preserving the properties of positivity and normalization of probability distribution. This is formally possible by relaxing the assumptions on the mathematical objects that describe the measurement process, still mantaining a consistent prescription to generate probabilities.

Indeed, from the general expression of the Born rule, we notice that in order to generate probabilities it is sufficient that the P_x^2 are positive operators, and therefore the requirement that they are projectors is not necessary. We introduce thus a generalization of the projectors P_x, in terms of positive operators $\Pi_x \geq 0$, and write a new prescription to generate probabilities as: $p_x = \mathrm{Tr}[\rho \Pi_x]$. Naturally, this must represent a true probability distribution, i.e. it must be normalized, and from this requirement follows that $\sum_x \Pi_x = \mathbb{I}$, which is analogous to the condition for the P_x.

We call a decomposition of the identity in terms of positive operators $\sum_x \Pi_x = \mathbb{I}$ a *"probability operator-valued measure"* (or POVM) and the $\Pi_x \geq 0$ the elements of the POVM.

Definition: We define $\mathcal{M}_x$, namely the post-measurement operator, or *detection operator*, as the operator that satisfies the following:

$$p_x = \mathrm{Tr}[M_x \rho M_x^\dagger] = \mathrm{Tr}[\rho \Pi_x]$$

hence obtaining:

$$\Pi_x = M_x^\dagger M_x,$$

which, by construction, is a positive operator.

The state after the measurement will be:

$$\rho_x = \frac{1}{p_x} M_x \rho M_x^\dagger.$$

Overall, we can say that every set of detection operators M_x that satisfy:

$$\sum_x M_x^\dagger M_x = \mathbb{I} \tag{1.1}$$

represents a generalization of the projectors P_x, while the POVM elements generalize P_x^2.

Therefore, our reformulation of the postulates consists in substituting projectors with POVMs and associate them with observable quantities. Thus, a measurement will yield as a result one of the elements of the POVM and both the Born rule and the projection postulate are reformulated in terms of the post-measurement operators M_x. We can then say that every set of detection operators M_x satisfying (1.1) is a legitimate measure operation that leads to a suitable probability distribution, which is said "*generalized measurement*".

We briefly mention here, for the sake of clearness on generalized measurements we just introduced, an important theorem, due to Naimark, which basically states that any generalized measurement can be considered as a standard measurement in a larger Hilbert space and, the other way round, if we focus on a portion of a composite system where a standard measurement takes place, then the statistics of the outcomes and the post-measurement states of the subsystem may be obtained with the tools of generalized measurements.

Naimark's Theorem: Let us consider a system coupled with an additional, or auxiliary, system, usually called *"ancilla"*. The Hilbert space of the overall system is $\mathcal{H}_A \otimes \mathcal{H}_B$ and we assume that the system and the ancilla are initially independent (i.e. the global initial state is $R = \rho_A \otimes \rho_B$) and let us assume moreover that the ancilla is prepared in the pure state $\rho_B = |\omega_B\rangle \langle\omega_B|$. Then we let them evolve and we want to obtain information on the state of the system by measuring an observable X on the ancilla, after its interaction with the main system. Let the unitary operation U describe this interaction. According to the Born rule, the probability of the outcomes is given by:

$$\begin{aligned} p_x &= \mathrm{Tr}_{AB}\left[U\rho_A \otimes \rho_B U^\dagger \mathbb{I} \otimes |x\rangle \langle x|\right] \\ &= \mathrm{Tr}_A\left[\rho_A \mathrm{Tr}_B[\mathbb{I} \otimes \rho_B U^\dagger \mathbb{I} \otimes |x\rangle \langle x| U]\right] \end{aligned}$$

and therefore the set of operators $\Pi_x = \mathrm{Tr}_B[\mathbb{I} \otimes \rho_B U^\dagger \mathbb{I} \otimes |x\rangle \langle x| U] = \langle\omega_B|U^\dagger \mathbb{I} \otimes P_x U|\omega_B\rangle$ is the object that would permit to write the Born rule at the level of the subsystem A, i.e. it is our candidate POVM. This setup is called the *"Naimark extension"* of the POVM.

Then, one can prove that the measurement procedure performed on the ancilla is adequately described by a set of detection operators which realizes a POVM on the Hilbert space of the system. Additionally, one can also prove that, conversely, given a set of detection operators M_x which realizes a POVM, this is the only description for the system of an indirect measurement performed on a larger Hilbert space.

Eventually, we can say that the Naimark theorem asserts that there exists a one-to-one correspondence between POVMs and indirect measurements of the type described above, namely measurements involving the coupling of a given system with an ancillary system. This means that an indirect measurement can be seen as the physical implementation of a POVM and, on the other hand, a POVM can be realized by an indirect measurement.

"Information is physical" - Rolf Landauer

Having briefly introduced the fundamental elements of quantum theory, we can now go on to analyse the question of the quantum description of the nature of information, since it has been clear since the early days of quantum theory that classical ideas about information would need a revision under the new formulation of physics.

The fundamental concept that lies at the basis of classical computation and information is the *bit*, which is usually defined as the indivisible unit of classical information, and can be in one of the two classical stable states, conventionally denoted by 0 and 1 [8].

Quantum computation and information are based on an analogous concept to that of the bit, namely the notion of quantum bit (or *qubit*). Just like bits, qubits are also physical objects and they describe a state in the simplest possible quantum system, i.e. the quantum system associated to the smallest non-trivial Hilbert space: a two-demensional Hilbert space $\mathcal{H}$.

Therefore, a vector of $\mathcal{H}$ represents a possible state of the system, and every couple of orthogonal and unitary states of $\mathcal{H}$ constitutes an ortho-normal basis for $\mathcal{H}$. We can choose a particular basis of this kind and denote it as $\{|0\rangle, |1\rangle\}$.

Then we could say that a qubit has two fundamental vector states, namely:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{1.2}$$

These states represent basis vectors in the complex two-dimensional space $\mathbb{C}^2$, equivalent to the Hilbert space $\mathcal{H}^2$.

The $\{|0\rangle, |1\rangle\}$ is usually called the *computational basis*, and the states $|0\rangle$ and $|1\rangle$ are the quantum analog of the states 0 and 1 for a bit.

Then we can say that the essential difference between bits and qubits is the fact that a qubit can be in a state which is different from either $|0\rangle$ or $|1\rangle$.

Indeed, every linear combination of states, i.e. *superposition*, of the form:

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle\,, \tag{1.3}$$

where a and b are complex numbers that satisfy $|a|^2 + |b|^2 = 1$ is the most general expression of a normalized state in $\mathcal{H}$.

Definition: a **qubit** is a state in a two-dimensional Hilbert space $\mathcal{H}$ that can take any value of the form (1.3).

If we perform a measurement on the state, it will project the qubit onto the basis $\{|0\rangle, |1\rangle\}$. Hence we will obtain the outcome $|0\rangle$ with probability $|a|^2$ and the outcome $|1\rangle$ with probability $|b|^2$.

Moreover, we find here another important difference between the classical and the quantum bit: we can measure a classical bit without disturbing it and decipher all the information it encodes. On the other hand, a measurement on the qubit will irrevocably disturb its state: if the initial value of the qubit is unknown, than it is not possible to determine a and b through only one measurement. However, after the measurement, the qubit will be prepared in a *known* state, i.e. either $|0\rangle$ or $|1\rangle$, that would genereally differ from the initial state.

We point out that the coefficients a and b in (1.3) represent more than just the probabilities of the outcomes of a measurement in the computational basis. In particular, they also encode information on the *relative phase* of a and b, which has also physical relevance.

It is natural to interpret the qubit as written in (1.3) as the spin state of an object with spin-1/2. Therefore the basis states $|0\rangle$ and $|1\rangle$ will represent the spin up ($|\uparrow\rangle$) and spin down ($|\downarrow\rangle$) states along a particular direction, conventionally the $\hat{z}$-axis.

Therefore, the two complex coefficients a and b that characterize the qubit (modulo the normalization and the overall phase) describe the *orientation* of the spin in three-dimensional space (i.e. the polar angle θ and the azimuthal angle ϕ).

Let us consider the spin operator in an arbitrary direction θ, i.e., the general observable:

$$\sigma_\theta = \cos\theta\,\sigma_{\mathbf{n}} + \sin\theta\,\sigma_{\mathbf{n}_\perp}$$

where $\{\mathbf{n}, \mathbf{n}_\perp\}$ is the orthogonal basis for the state space and the normalization on

probabilities, $|\cos^2\theta| + |\sin\theta|^2 = 1$, holds.

Let us now consider a general (qubit) state described by the density matrix ρ.
We can write the probability distributions of the outcomes of measurements as:

$$p_0 = \langle 0|_\theta \, \rho \, |0\rangle_\theta \,, \quad p_1 = \langle 1|_\theta \, \rho \, |1\rangle_\theta$$

where $|0\rangle_\theta$ and $|1\rangle_\theta$ are eigenstates of σ_θ.
The probabilities on all possible θ therefore will be:

$$P_0 = \int \mathrm{d}\theta \, p(\theta) \, p_0(\theta) \tag{1.4}$$

and

$$P_1 = \int \mathrm{d}\theta \, p(\theta) \, p_1(\theta) = 1 - P_0, \tag{1.5}$$

where $p(\theta)$ is the probability that the state is found along the direction θ. Moreover, we have, as for probabilities:

$$P_0 + P_1 = 1.$$

We can write this probabilities as follows:

$$P_0 = \int \mathrm{d}\theta \, p(\theta) \, p_0(\theta) = \int \mathrm{d}\theta \, p(\theta) \; \langle 0|_\theta \, \rho \, |0\rangle_\theta = \mathrm{Tr}[\rho \, \Pi_0], \tag{1.6}$$

and analogously for P_1, where use has been made of the *Born rule*.

In (1.6) we introduced the *POVM* defined as:

$$\Pi_0 = \int \mathrm{d}\theta \, p(\theta) \, |0\rangle_\theta \; \langle 0|_\theta \,,$$

whose elements satisfy to the following properties:

$$\begin{cases} \Pi_k \geq 0; \\ \sum_k \Pi_k = \mathbb{I}. \end{cases}$$

1.2 Entanglement

As it is well known, Albert Einstein, along with Nathan Rosen and Boris Podolski (EPR) [9], and, independently, Erwin Schrödinger [10] were the first who recognized in 1935 a quite uncanny feature of quantum mechanics, questioning the completeness of the theory, which lies since then at the center of interest of the physics of the XXI century. This feature implies that global states of composite system exist, which cannot be written as a product of the states of their individual subsystems, and was famously referred to as "the spooky action at a distance" by Einstein himself. This phenomenon, now known as "entanglement", was originally called by Schrödinger "Verschränkung" (from the german verb *verschränken*, to cross, or to join crosswise), which referred to the intrinsic order of statistical relations existing between the subsystems composing a global quantum system.

Later, John Stewart Bell recognized that entanglement leads to considerable deviations of quantum mechanics from classical physics, which could be tested experimentally [11]. In fact, he showed that the attempt of ascribing values to physical quantities prior to measurement was ruled out by entanglement itself, accepting the EPR conclusion that the quantum description of physical reality is incomplete and formalizing the *local hidden variable model* [12], based on the assumptions of *realism, locality* and *free will*. He then introduced the so-called *"Bell inequalities"*, proving that those assumptions impose experimental constraints on statistical correlations in bipartite systems, namely that the probabilities for the measurement outcomes of an entangled quantum state violate the Bell inequalities. Therefore, entanglement is considered that feature of quantum formalism which makes it impossible to simulate quantum correlations with any classical formalism.

Anyway, hidden variable programs caused several controversies and are in general refused by standard quantum mechanics. In particular, the *"Bell Theorem"* was demonstrated without any explicit reference to hidden variables and states the *non-locality* of quantum mechanics, namely, the outcome of a measurement can be influenced by the measurement context even at a distance, i.e., non locally [12]. Hence, a physical system can, or cannot, exhibit a certain property depending on the outcomes of measurements on a second physical system at a distance, *correlated* with the first one.

In the mid-60s of the last century there began the transition from the theory of entanglement to its experimental reality, and since then many interesting and consistent experiments have been performed, firmly confirming the predictions of quantum description.

In fact, a fundamental non-classical axpect of entanglement was already been recognized in the mid-30s by Schrödinger, in relation to the notion of "knowledge" in quantum

context. Only at the end of the century this singular aspect of entanglement was formalized in terms of entropic inequalities [13] and a physical interpretation was given, in which the fundamental quantities involved were seen as responsible for the capabilities of transmission of quantum *information* [14], [15]. In particular, transmission is possible exactly in those situations in which the entropy of the output system exceeds the entropy of the total system.

In the following years, enormous experimental progress and a growing interest in entanglement theory were the basis for the advent of quantum information theory, which finally recognised entanglement as a central notion and an important resource, enabling useful applications like quantum cryptography [16], quantum teleportation [17] or measurement based quantum computation [18], otherwise impossible to be performed by means of classical resources.

Many experiments nowadays aim at the generation of entanglement (as fully presented in [4]). Entanglement is also of big relevance in tasks such as quantum communication between parties separated by a macroscopic distance and the development of quantum computing. On the other hand, it has also given new insights on the understanding of many and various physical phenomena and on the interpretation of important and fundamental questions of the theory.

One of the first definitions of entangled states comes in a negative form in a seminal paper by R. F. Werner, who gave an accurate definition of *separable* states, i.e. those states that are not entangled. [19]

Then, Asher Peres ([20]) introduced the so-called *positive partial transpose (PPT) criterion* for density matrices, or *Peres criterion*, for the discrimination of separable states, which appeared to be a strong test for entanglement. Consequently to the introduction of this criterion, it was noticed that in general positive maps (like the partial transpose) can be used as strong detectors of entanglement. Unfortunately, they are unphysical and therefore cannot be implemented directly in laboratory contexts. However, the "Jamiołkowski isomorphism" (which we are going to discuss briefly later) comes to help, offering a duality between these unphysical objects and physical measurable quantities, i.e. Hermitian operators, and providing a characterization of entanglement that would serve as a basis for a general theory of detection of entanglement.

A completely different type of separability criteria has been later introduced on the basis of the pioneering work of Barbara Terhal [21] and it relates on the fundamental concept of *entanglement witnesses*.

The concept of entanglement witness was applied to different problems in quantum physics, but their main virtue is that they provide a smart and consistent way of detecting entanglement, since it does not need full (tomographic) information about the state [3].

Unfortunately quantum entanglement has some bothersome but interesting features: it generally has a rather complex structure, it is fragile to environment and it can not

be increased on average when systems are not directly in contact but are distributed in spatially separated regions.

Thus, the theory of entanglement tries to answer some fundamental questions such as how to optimally detect entanglement, both theoretically and in laboratory, how to reverse an inevitable process of entanglement degradation and how to characterize, control and quantify entanglement [3].

Therefore one can say that the fundamental question in quantum entanglement theory is to discern which states are entangled and which are not. Only in few cases this question has a simple answer. The simplest case is that of pure bipartite states.

1.2.1 Definition

As we discussed in section 1.1, the general axioms of quantum mechanics are intended to characterize the quantum behaviour of a global quantum system, i.e. the entire universe. Most of the time, however we wish to limit our observations only to a small part of the universe, that is, we observe a small section of a larger quantum system. This means that we concentrate on a system inside its *environment*, and therefore we must also consider the interactions between them, for a complete physical description.

According to Bell's interpretation, *quantum information* is encoded in *non-local correlations* between the different parts of a physical system.
Almost all the information that discerns a state from another in a given system is encoded in the non-local correlations between the results of the measures of its subsystems.

Another important feature is that the content of information can be quantified by *entropy*, where big entropy means little information. In fact, for big entropy we have access only to an exponentially little quantity of information, considering every (sub)system separately.
The measures reveal very little information if we do not consider how the results of a measure obtained on all the subsystems are correlated between them.

When these non-local correlations exist between the parts of the system, we say that these parts are **entangled.** This means that we cannot completely decode the state of the system by deviding it and studying its individual parts.

Generally, quantum systems are inevitably in contact with the *environment*. This interaction estabilishes non-local correlations between them as expressed by the *decoherence* concept of the impossibility if isolating the system from the environment. Therefore quantum information, initially encoded in the system, turns out to be encoded in the *correlations* between the system and the environment.

Let us consider now the most simple case of a *bipartite quantum system* [4]: for instance, we consider a system made up of two qubits, say A and B, and we observe only one of them. Let us define the orthonormal bases for the two qubits, A and B respectively, as follows:

$$\{|0\rangle_A, |1\rangle_A\}$$
$$\{|0\rangle_B, |1\rangle_B\}$$

Then we could write the state of the global system as follows:

$$|\psi_{AB}\rangle = a\,|0\rangle_A \otimes |0\rangle_B + b\,|1\rangle_A \otimes |1\rangle_B$$

and we say that the qubits A and B are *correlated*.

Let us suppose that we measure the qubit A by projecting onto the $\{|0\rangle_A, |1\rangle_A\}$ basis. Then, the measure will return $|0\rangle_A$, with probability $|a|^2$, and will prepare the system in the state $|0\rangle_A \otimes |0\rangle_B$, or will return $|1\rangle_A$, with probability $|b|^2$, and will prepare the system in the state $|1\rangle_A \otimes |1\rangle_B$.

In either case, we could say that a definite state of qubit B is *selected* by the measurement: if we act a subsequent measure on B then we are guaranteed (with probability one) to find $|0\rangle_B$ if we had found $|0\rangle_A$ or $|1\rangle_B$ if we had found $|1\rangle_A$. In this sense, we say that the outcomes of the $\{|0\rangle_A, |1\rangle_A\}$ and $\{|0\rangle_B, |1\rangle_B\}$ measurements are perfectly correlated in the state $|\psi\rangle_{AB}$.

Entanglement of pure states

Let us assume that we are given two quantum systems. The first one is owned by one physicist, called Alice, and the second one by another one, called Bob. The physical states of Alice's system may be described by states in a Hilbert space $\mathcal{H}_A$ of dimension d_A, and in Bob's system in a Hilbert space $\mathcal{H}_B$ of dimension d_B. The composite system of both parties is then described by vectors in the tensor-product of the two spaces $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Thus, any vector in $\mathcal{H}_A \otimes \mathcal{H}_B$ can be written as

$$|\psi\rangle = \sum_{i,j=1}^{d_A,d_B} c_{i,j}\,|a_i\rangle \otimes |b_j\rangle \in \mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$$

with a complex $d_A \times d_B$ matrix $C = (c_{ij})$. To keep the notation simple, we often write tensor products of vectors as $|a\rangle \otimes |b\rangle \equiv |a\rangle\,|b\rangle \equiv |ab\rangle$. Now one can define separability and entanglement for pure states.

Definition: Entanglement for Pure States:
A bipartite pure state $|\psi_{AB}\rangle \in \mathcal{H}$ is said to be a product state or *"separable"* if it is the direct product of pure states in the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, i.e., if we can find states $|\phi_A\rangle \in \mathcal{H}_A$ and $|\phi_B\rangle \in \mathcal{H}_B$ such that:

$$|\psi_{AB}\rangle = |\phi_A\rangle \otimes |\phi_B\rangle \,.$$

Otherwise the state $|\psi_{AB}\rangle$ is called *entangled*.

The definition of product states means, physically, that the state is not correlated. Therefore a product state can be easily prepared locally: for instance, Alice produces the state $|\phi_A\rangle$ and Bob produces in an independent way the state $|\phi_B\rangle$. Subsequently, when Alice measures any observable A and Bob measures B, the probabilities of the different outcomes will factorize. Thus, we could say that the measurement outcomes for Alice are independent from the outcomes on Bob's part of the system.

We will now give a definition of entanglement for mixed states and to do so we need to introduce a very useful tool in the description of entanglement for bipartite systems, namely the so-called *Schmidt decomposition* [22].

The Schmidt decomposition is of central rilevance in the characterization and quantification of entanglement associated with pure states. For a vector in the tensor product of two Hilbert spaces

$$|\psi\rangle = \sum_{i,j=1}^{d_a,d_B} c_{ij} \, |a_i b_j\rangle \in \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}},$$

there exists an orthonormal basis $|\alpha_j\rangle$ of $\mathcal{H}_{\mathcal{A}}$ and an orthonormal basis $|\beta_i\rangle$ of $\mathcal{H}_{\mathcal{B}}$ such that the following holds:

$$|\psi\rangle = \sum_{k=1}^{R} \lambda_k \, |\alpha_k \beta_k\rangle \,, \tag{1.7}$$

where the coefficients $\lambda_k \geq 0$ are real and the number $R \leq \min\{d_A, d_B\}$ is called the *Schmidt rank* of $|\psi\rangle$.

One has, from the general definition of a Schmidt decomposition, that the λ_k are unique, up to a permutation, and, if they are pairwise different, then the $|\alpha_k\rangle$ and $|\beta_k\rangle$ are also unique, up to a phase.

Then, one can give an alternative definition of separability associating with any bipartite pure state $|\psi\rangle$ a positive integer, namely the Schmidt rank R and saying that the state is entangled (or non-separable) if its Schmidt rank is greater than one. Therefore, since a separable bipartite pure state can be written as a direct product of pure states in each of the local Hilbert spaces (as in the definition above), then the reduced density matrices $\rho_A = |\phi\rangle_A \langle\phi|_A$ and $\rho_B = |\phi\rangle_B \langle\phi|_B$ are pure. On the other hand, any entangled

state cannot be expressed as a direct product of pure states, hence ρ_A and ρ_B are mixed states.

Entanglement of mixed states

We can now generalize what we said for pure states to the situation in which one does not know exactly the state of a quantum system. Indeed, due to the decoherence phenomenon, in the experimental context we unavoidably deal with mixed states rather than pure ones. However mixed state can still contain some kind of "noisy" entanglement.
Therefore, what one can say is only that the state is, with some probability p_i, in one of some states $|\phi_i\rangle \in \mathcal{H}$, and is consequently described by a so-called *density matrix*:

$$\rho = \sum_i p_i \left|\phi_i\right\rangle \left\langle\phi_i\right|,$$

$$\text{with} \quad \sum_i p_i = 1 \text{ and } p_i \geq 0$$

Therefore, given a certain basis, this density matrix, or state, will be represented by a complex matrix. This matrix is positive semidefinite and hermitian, since all the operators $|\phi_i\rangle \langle\phi_i|$ are positive and hermitian, and has unitary trace, due to the conditions on the probabilities.
Vice versa, any matrix that follows these properties can represent a density matrix of a certain state. This gives a geometrical picture of the set of all states as a convex state, i.e., given the density matrices ρ_i we can define their *convex combination* as:

$$\rho = \sum_i p_i \rho_i,$$

where $p_i \geq 0$ and $\sum_i p_i$ are often called *convex weights*. The resulting matrix ρ is again a state.

Now one can define separability and entanglement for mixed states, according to Werner [19]. The reasoning is the same as that for the pure case and we can therefore say that a state is separable if it can be produced locally, otherwise it is entangled.

Definition: Entanglement for Mixed States:
Any bipartite state represented by a density matrix ρ_{AB} defined on Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ is said to be a *product state* if there exist a state ρ_A in $\mathcal{H}_A$ and ρ_B in $\mathcal{H}_B$ such that:

$$\rho_{AB} = \rho_A \otimes \rho_B.$$

Therefore we can say that the state is separable if there exist convex weights p_i and product states $\rho_A^i \otimes \rho_B^i$ such that the following holds:

$$\rho = \sum_i p_i \rho_A^i \otimes \rho_B^i,$$

where ρ_A^i and ρ_B^i are defined on the local Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. Otherwise we say that the state is entangled.

From a physical point of view, separable states are said to be classically correlated, i.e., for producing a separable state are necessary only local operations and classical communication. For instance, Alice and Bob can share a random number generator through classical communication, producing the outcomes i with probabilities p_i. Then they can agree to produce locally the product state $\rho_A^i \otimes \rho_B^i$ for each of the outcomes. This kind of procedure leads then to the production of the state ρ. We notice that this procedure is not specific for the quantum theory, and this is the reason why we call this kind of correlation "classical". On the other hand, for an entangled state the correlations cannot be created through the local procedure we just described and therefore in this sense we say that entanglement is a typical feature of the quantum theory.

We shall mention, finally, that while the theory of bipartite quantum correlations is well developed, the generalization to a multipartite case is still an open problem (see e.g. [3] and [23]), and a lot of effort is put in this direction since it is relevant expecially in the case of entanglement and its applications to quantum information theory [24]. Indeed, while for bipartite quantum systems there exists (up to some local unitary operation) one single most entangled state, namely the *maximally entangled state*, this is not the case for multipartite systems. In fact, multipartite quantum systems can be entangled in many inequivalent ways, therefore leading to the existence of different entanglement classes, and in general a maximally entangled state is not unique.
A maximally entangled state dues its name to the fact that spatially separated parties can deterministically obtain any other bipartite state from it via local operations assisted by classical communication (*LOCC*, for short) [25].

This means that the maximally entangled state is the **optimal** bipartite entanglement resource, since it maximizes any measure of entanglement and serves as a standard with which the resourcefulness of other entangled bipartite quantum states can be compared [26].

This situation is drastically different in the multipartite case, in which one cannot find a unique maximally entangled state to use as a test of resourcefulness of other quantum states. This peculiarity however is the reason why a lot of effort nowadays goes to the study of multipartite entanglement, in order to find entanglement detections schemes capable of distinguishing between the different classes.

1.2.2 Separability criteria

Given the above definitions of entanglement and separability, it is very natural to ask whether a given state or density matrix is separable or entangled. This is the so-called *separability problem* and it represents one of the fundamental problems in entanglement theory. There are several known criteria that imply separability or entanglement of a state. However, up until now, no general solution for the separability problem is known. We are now going to present some of the most common criteria for bipartite entanglement. One of the most known is the socalled *PPT criterion*, as first presented in [20].

Positive Partial Transpose (PPT) criterion:

Let us first note that we can expand any density matrix of a composite quantum system in a chosen product basis as:

$$\rho = \sum_{i,j}^{N} \sum_{k,l}^{M} \rho_{ij,kl} \left|i\right\rangle \left\langle j\right| \otimes \left|k\right\rangle \left\langle l\right|$$

We can now define the *partial transposition* of ρ as the transposition with respect to one subsystem. It follows that, in a bipartite system, there are two partial transposition, one for each sub-system. For instance, the partial transposition with respect to the subsystem A is given by:

$$\rho^{T_A} = \sum_{i,j}^{N} \sum_{k,l}^{M} \rho_{ji,kl} \left|i\right\rangle \left\langle j\right| \otimes \left|k\right\rangle \left\langle l\right|$$

and similarly the one with respect to B will be:

$$\rho^{T_B} = \sum_{i,j}^{N} \sum_{k,l}^{M} \rho_{ij,lk} \left|i\right\rangle \left\langle j\right| \otimes \left|k\right\rangle \left\langle l\right|$$

It is then straightforward to note that the partial transposition is related to the usual transposition by: $\rho^T = (\rho^{T_A})^{T_B}$, and so $\rho^{T_B} = (\rho^{T_A})^T$.

We can now say that a density matrix ρ has a *positive partial transpose*, or the matrix is *PPT*, if its partial transposition has no negative eigenvalues, i.e. is positive semidefinite:

$$\rho^{T_A} \geq 0 \iff \rho^{T_B} \geq 0.$$

We can now state the

PPT Criterion: Let ρ be a bipartite separable state. Then ρ is PPT.
This fact follows directly from the definition of separability. In fact for a separable state $\rho = \sum_k p_k {\rho_k}^A \otimes {\rho_k}^B$ we have

$$\rho^{T_A} = \sum_k p_k ({\rho_k}^A)^T \otimes {\rho_k}^B = \sum_k p_k \tilde{\rho}_k^A \otimes {\rho_k}^B \geq 0.$$

This theorem provides a very strong criterion for the detection of entanglement: for a given density matrix one can easily calculate the partial transpose and compute its spectrum. If one finds negative eigenvalues one can conclude that the state is entangled. Given this result, the question arises if this criterion is also sufficient for separability, i.e., whether $\rho^{T_A} \geq 0$ implies separability. As it was shown immediately after the discovery of the PPT criterion, this is the case only in low dimensional systems, as stated by the following

Horodecki Theorem [27]: If ρ is a state in a 2×2 or 2×3 system, then $\rho^{T_A} \geq 0$ implies that ρ is separable. In other dimensions this is not the case.

Although the PPT criterion does not constitute a necessary and sufficient criterion, it is the most popular criterion. This fact is due to several reasons, beside its simplicity. First of all, the fact that it provides a complete characterization of entanglement for two-qubit systems makes it very appealing, since two-qubit systems are the most studied bipartite systems (however, the PPT criterion is of limited use for the investigation of multipartite entanglement). Second, it has been shown that the amount of violation of the PPT condition can be used to quantify entanglement (see [28] and [29]).

Since the PPT criterion does not detect all states, the question arises of how can one prove that a state is entangled, if the PPT criterion fails. For this problem, many criteria have been proposed. One of the most simple and strong is the *computable cross norm or realignment criterion* (CCNR criterion) [30]. In order to formulate the CCNR criterion, we make use of the *Schmidt decomposition* in the operators space, defined in 1.2.1. The λ_k in (1.7) can then be computed as in the general case, and we can formulate the

CCNR criterion: if the state ρ is separable, then the sum of all λ_k in its Schmidt decomposition in operator space is smaller than one:

$$\sum_k \lambda_k \leq 1.$$

Therefore, if $\sum_k \lambda_k > 1$ the state must be entangled.

The remarkable fact is that the CCNR criterion allows one to detect entanglement for many states where the PPT criterion fails. Combined with its simplicity, this makes it a useful tool for the analysis of entanglement. However, it does not detect all entangled states of two-qubits [31]. Therefore, one may view it as complementary to the PPT criterion.

Beside the PPT and CCNR criterion, there are many other approaches to derive separability criteria (as fully reviewed in the survey "Entanglement detection" by O.

Gühne and G. Tóth, [4]). Let us mention here some of them for the sake of completeness.

The range criterion was one of the first known criteria for the detection of states for which the PPT criterion fails [32]. It states that if a state ρ is separable, then there is a set of product vectors $|a_i b_i\rangle$ such that the set $\{|a_i b_i\rangle\}$ spans the range of ρ while the set $\{|a_i^* b_i\rangle\}$ spans the range of ρ^{T_A}. Let us note that this criterion, however, cannot be used if some state is affected by noise: then, the density matrix and its partial transpose will usually have full rank, hence the condition in the range criterion is automatically fulfilled.

Positive maps: other entanglement criteria similar to the PPT criterion can be formulated from other positive, but not *completely positive maps* (**CP-maps** for short). Let us briefly define a CP-map: let $\mathcal{H}_B$ and $\mathcal{H}_C$ be Hilbert spaces and let $\mathcal{B}(\mathcal{H}_i)$ denote the linear operators on it. A linear map $\Lambda : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_C)$ is called positive if it maps hermitian operators onto hermitian operators, fulfilling $\Lambda(X^\dagger) = \Lambda(X)^\dagger$, and it preserves the positivity, i.e. $X \geq 0 \Rightarrow \Lambda(X) \geq 0$. Note that the second condition implies that it maps valid density matrices onto density matrices, up to normalization. A positive map Λ is called *completely* positive when, for an arbitrary Hilbert space $\mathcal{H}_A$, the map $\mathbb{I}_A \otimes \Lambda$ is positive. Otherwise, Λ is positive, but not completely positive. Here, $\mathbb{I}_A$ denotes the identity on $\mathcal{B}(\mathcal{H}_A)$.

From this notions, other entanglement criteria similar to the PPT criterion can be formulated from other positive, but not completely positive maps: *for any separable state ρ and any positive map Λ we have*

$$(\mathbb{I}_A \otimes \Lambda)(\rho) \geq 0.$$

Furthermore, it has been shown in [27] that a state ρ is separable if and only if, for all positive maps Λ, the before relation holds. In this sense, the separability problem is equivalent to the classification of all positive maps.

The theory of positive maps is of fundamental importance for the fact that any entanglement witness gives rise to a positive, but not CP, map via the Choi-Jamiołkowski isomorphism. This problem was considered in the mathematical literature for a long time, however, it has not been yet solved. From the perspective of quantum information theory, the classification of positive maps has been under intensive research and has led to many new positive, but not completely positive maps, resulting in strong separability tests.

For instance, in [33], a reduction map is defined, which is an example of positive (but not CP) map, which gives the *reduction criterion for separability*; in the work of Størmer [34] positive linear maps are used; and in [35] a new family of indecomposable positive linear maps based on entangled quantum states is introduced.

Moreover, we mention the *majorization criterion* that relates the eigenvalues of the global state to reduced states [36].

Furthermore *algorithmic approaches* have been adopted for the separability problem, which formulate separability in terms of optimization problems or semidefinite programs, with the goal of deriving numerical algormithms for separability testing.

Another approach uses criteria based on *covariance matrices*, and a similar approach makes use of the *expectation value matrix*.

Finally, another kind of separability criterion was developed in [27] using *linear contractions and permutations*.

The list is very long and cannot be fully detailed here. Other approaches for detecting entanglement make use of special observables, namely Bell inequalities and entanglement witnesses. We are going to briefly review the latter to complete this introduction, and then move on to our work.

1.2.3 Witnesses

All the criteria listed above have something in common: at first sight they all assume that the density matrix is already known. They all require applying certain operations to a density matrix, to discern between entangled and separable states. There is, however, a necessary and sufficient entanglement criterion in terms of directly measurable observables. These are the so-called *entanglement witnesses* (see [37], [27], [21] and [38]).

The entanglement witnessing process consists in verifying that a source is producing entangled particles. Commonly this process requires, in the first place, to prove a certain mathematical identity that all separable states must satisfy. Subsequently, one can demonstrate experimentally whether the source under analysis violates or fulfills the identity. Such identity is usually referred to as an entanglement witness and its violation in experiment guarantees that the source is generating entangled particles.

Definition: An observable $\mathcal{W}$ is called an **"entanglement witness"** if the following conditions hold:

$$\begin{cases} \mathrm{Tr}(\mathcal{W}\rho_s) & \geq 0 \quad \text{for all separable } \rho_s \\ \mathrm{Tr}(\mathcal{W}\rho_e) & < 0 \quad \text{for at least one entangled } \rho_e. \end{cases}$$

Thus, if one measures $\mathrm{Tr}(\mathcal{W}\rho) < 0$ one knows for sure that the state ρ is entangled. We than say that such state is *detected* by $\mathcal{W}$.

The fact that entanglement witnesses are directly measurable quantities makes them a very useful tool for the analysis of entanglement in experiment, and this makes them one of the main methods used to detect entanglement experimentally.

For further understanding, it is crucial to note that entanglement witnesses have a clear geometrical meaning. The expectation value of an observable depends linearly on

the state. Thus, the set of states where $\mathrm{Tr}(\mathcal{W}\rho) = 0$ holds is a hyperplane in the set of all states, cutting this set into two parts. In the portion of hyperplane with $\mathrm{Tr}(\mathcal{W}\rho) > 0$ lies the set of all separable states, while the other portion, with $\mathrm{Tr}(\mathcal{W}\rho) < 0$ is the set of states detected by $\mathcal{W}$.

From this geometrical interpretation it follows that all entangled states can be detected by witnesses, as stated in the following theorem:

Completeness of witnesses: for each entangled state ρ_e there exist an entanglement witness detecting it.

Although this theorem ensures that any entangled state can in principle be detected with an entanglement witness, the task remains to construct witnesses. This is not an easy problem, since solving this problem would also solve the separability problem. A large part of the review by Gühne and Tòth [4], to which most of this introductions refers, is concerned with the proper construction and evaluation of witnesses for the multipartite case.

Chapter 2

Entanglement witness based on classical correlations

2.1 Introduction

Let us start by describing the protocol introduced by Maccone et al. in [1] to detect entanglement by measuring classical correlations. The protocol is schematically depicted in Figure (2.1).

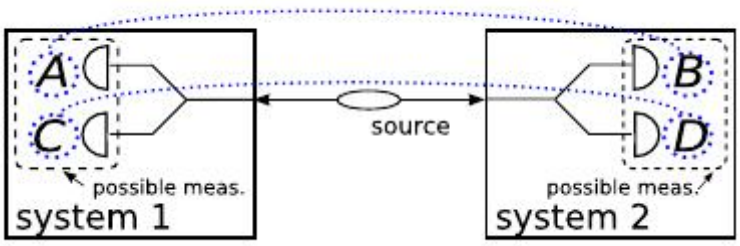

Figure 2.1: Schematic illustration of the protocol to detect entanglement by measuring classical correlations. A and C are observables for system 1 and B and D are observables on system 2: each of the two systems is subject to the measurement of one of the two respective observables. Classical correlations are evaluated between the results of A and B and those of C and D (dashed lines). A and C are complementary on the first system, B and D are complementary on the second. $A \otimes B$ and $C \otimes D$ are bipartite complementary observables. Figure taken from [1].

Let us consider two systems of finite dimension d and let us identify two bipartite complementary observables: $A \otimes B$ and $C \otimes D$, where A and C are complementary on

system 1 and B and D on system 2 i.e.:

$$\begin{array}{l} |\langle a_i|c_j\rangle|^2 = 1/d \\ |\langle b_i|d_j\rangle|^2 = 1/d \end{array} \tag{2.1}$$

for all eigenstates of A and C and of B and D, respectively. We will be concerned in measuring the classical correlations between A and B and between C and D, which are indicated by the dashed line in figure (2.1). The central idea is that only quantum correlated states can exhibit strong classical correlations in the measurement outcomes of local complementary observables [24]. Let us then measure the statistics of the outcomes of the two observables. We can see that local measurements on the two systems will suffice [39].

These measurements return the joint probabilities $p(a_0, b_0)$ of obtaining the outcome a_0 on system 1 for the observable A and the outcome b_0 on system 2 for the observable B, and $p(c_0, d_0)$ of obtaining the outcome c_0 on system 1 for the observable C and the outcome d_0 on system 2 for the observable D. These probabilities are then used to calculate the *mutual information* I_{AB} among measurements results for $A \otimes B$ and I_{CD} among results for $C \otimes D$.

One has then:

$$I_{AB} \equiv \sum_{a,b} p(a,b) \log_2 \frac{p(a,b)}{(\sum_a p(a,b) \sum_b p(a,b))},$$

and analogously for I_{CD}:

$$I_{CD} \equiv \sum_{c,d} p(c,d) \log_2 \frac{p(c,d)}{(\sum_c p(c,d) \sum_d p(c,d))}.$$

The criterium to certify entanglement is then the following [39], as outlined in figure (2.2):

- if $I_{AB} + I_{CD} > \log_2 d$, the two systems are *entangled*;
- if $I_{AB} + I_{CD} = 2 \log_2 d$, the two systems are *maximally entangled*;
- if $I_{AB} + I_{CD} < \log_2 d$, we would have all separable and some entangled states.

The proof of this statements is based on Maassen and Uffink's entropic uncertainty relation (EUR) [40] and is given in the original paper's Supplemental material.

In the next section we are going to briefly review the steps that led, from the pioneering work by Deutsch [41] and a conjecture of Kraus [42], to the derivation by Maassen and Uffink [40] of a new class of uncertainty relations for measurements of pairs of observables in a finite-dimensional Hilbert space which do not have any common eigenvector (i.e. they are complementary).

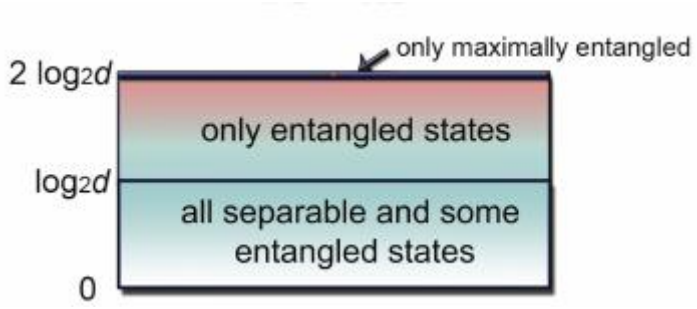

Figure 2.2: Taken two complementary observables $A \otimes B$ and $C \otimes D$, we calculate the mutual information I_{AB} and I_{CD} of their outcomes. Being d the dimension of the system, if the value of the sum of the mutual informations $I_{AB} + I_{CD}$ is larger than $\log_2 d$ bits, then the two systems are certified to be entangled. Moreover, they are certified to be entangled if the value of the sum is $2 \log_2 d$ bits. Figure taken from [39].

2.1.1 Entropic uncertainty relations (EUR)

In general quantum theory, the uncertainty principle states that non-commuting observables cannot *simultaneously* have precisely defined values. This means that a measure on an observable A will necessarily influence the outcome of a following measure on an observable B, when A and B are incompatible; i.e. the act of acquiring information on a physical information unavoidably disturbs (perturbs) the state of the system. That is, if we let A and B denote two Hermitian operators representing physical observables in an N-dimensional Hilbert space, and let $\{|a_j\rangle\}$ and $\{|b_j\rangle\}$ with $j = 1, ..., N$ be the corresponding complete sets of normalized eigenvectors, the general uncertainty principle states that, for any quantum state (normalized) vector $|\phi\rangle$, the two probability distributions $p = (p_1, ..., p_N)$ and $q = (q_1, ..., q_N)$, defined by

$$p_j = |\langle a_j|\phi\rangle|^2 \text{ and } q_j = |\langle b_j|\phi\rangle|^2$$

cannot be arbitrarily peaked, provided that A and B are sufficiently non-commuting.

Originally, this principle was expressed by the, following, *Robertson relation* [43]:

$$\Delta_\phi A \Delta_\phi B \geq \frac{1}{2} |\langle [A, B] \rangle_\phi|, \tag{2.2}$$

where $\Delta_\phi A$ and $\Delta_\phi B$ denote the standard deviations of the probability distributions p and q:

$$(\Delta_\phi A)^2 = \langle A^2 \rangle_\phi - (\langle A \rangle_\phi)^2$$

and analogously for $\Delta_\phi B$.

When the expectation value of the commutator $[A, B]$ does not vanish, (2.2) expresses the purely quantum mechanical property of limitating the possibility of preparing a

quantum ensemble with arbitrarily narrow variances for the relative observables. Hence it generally gives physically useful information about the considered observables for the pure case associated to the state.

However, this formulation of the uncertainty principle has been soon replaced by the so-called *"entropic uncertainty relations"*. The problem with the Robertson formulation was that the right-hand side of the inequality does not constitute a fixed lower bound, but depends on the particular state $|\phi\rangle$. For example, when $|\phi\rangle$ is an eigenstate of A, one obtains the trivial solution of 0 on both sides, and it follows that a restriction on $\Delta_\phi B$ cannot be imposed.

The **entropic uncertainty relations** (*EUT*, for short), rely on the *Shannon entropy* H as a measure of uncertainty, which is naturally associated to the measurement process of a pair of observables.

Definition - Shannon Entropy: We define the Shannon entropy for a general probability distribution $P = (P_1, ..., P_N)$, with $P_i \geq 0$ and $\sum_i P_i = 1$, on a set of N possible outcomes as follows:

$$H(P) = -\sum_j P_j \log P_j,$$

where P_j are the probabilities of getting the eigenvalue a_j in a measurement of the observable A in the given state $|\phi\rangle$.

Applying this notion to the probability distributions p and q, as defined above, the following improved EURs have been introduced:

- **Deutsch:**

$$H(p) + H(q) \geq -2 \log \tfrac{1}{2}(1 + c);$$

- **Kraus:**

$$H(p) + H(q) \geq -2 \log c.$$

They basically state that, for any state, the sum of Shannon entropies has a lower bound (that represents the measure of incompatibility of the two observables) which is a function of the maximum *overlap* of the two measurements, namely:

$$c = max_{j,k} |\langle a_j | b_k \rangle|.$$

As one can see, these relations have a bound which is independent of the state $|\phi\rangle$, unlike the one defined by Robertson. It follows that they yield non-trivial information on the probability distributions (i.e. the non-trivial lower bound is strictly positive), concerning the sum of the uncertainties associated to the measurement outcomes, as long as the observables A and B do not share any common eigenvector.

Hence, an entropic uncertainty relation shows that for any input state $|\phi\rangle$ there is some uncertainty in at least one of the two observables, and it is quantified by their relative Shannon entropies. In general, if we donote the lower bound with b_c one has $1/d \leq c \leq 1$ and therefore $0 \leq b_c \leq \log d$.

In the paper by Maassen and Uffink, they demonstrate the inequality by Kraus. Moreover, they show that it represents just one member of a general class of inequalities, all of which are seen to express the uncertainty principle in the sense that they put bounds, on the extent to which the distributions p and q can be simultaneously peaked.

However, the bounds given by Deutsch and Kraus are in general not optimal. The optimal lower bound for two given observables can only be found by calculating explicitly the minimum of the entropy sum over all the normalized state vectors $|\phi\rangle \in \mathcal{H}$.

Optimal EURs have then been extensively studied, expecially by Sánches in his paper from 1998 [44], in the particular case of a two-dimensional Hilbert space, which we are going to shortly discuss in the next section.

2.1.2 Optimal EUR in two-dimensional Hilbert spaces

For the sake of semplicity, let us now assume that both observables A and B have non-degenerate spectra. Therefore, the probability distributions $\{p_i(A)\}$ and $\{p_i(B)\}$, with $i = 1, ..., N$, for the N possible outcomes of measurements of A and B when the quantum system is described by $|\phi\rangle$ are given by:

$$p_i(A) = |\langle\phi|a_i\rangle|^2 \quad \text{and} \quad p_i(B) = |\langle\phi|b_i\rangle|^2,$$

or, more generally, if the state of the system is described by a density matrix ρ, by:

$$p_i(A) = \langle a_i|\rho|a_i\rangle \quad \text{and} \quad p_i(B) = \langle b_i|\rho|b_i\rangle$$

Then, an *entropic uncertainty relation* for A and B is an inequality of the form:

$$H(A) + H(B) \geq \inf_{|\phi\rangle}(H(A) + H(B)) \equiv H_{AB} > 0, \tag{2.3}$$

where $H(A)$ and $H(B)$ are the information (Shannon) entropies corresponding to the probability distributions $\{p_i(A)\}$ and $\{p_i(B)\}$, respectively:

$$\begin{aligned} H(A) &= -\sum_{i=1}^{N} p_i(A) \log p_i(A) \\ H(B) &= -\sum_{i=1}^{N} p_i(B) \log p_i(B) \end{aligned} \tag{2.4}$$

Moreover, one has:

$$\langle \phi | \phi \rangle = 1 \Rightarrow \sum_i p_i = 1$$

According to Shannon's point of view of information theory [45], entropy is showed to be the only rigorous quantitative measure of the uncertainty, or lack of information, associated to a random variable.

Then, the Maassen and Uffink's entropic uncertainty relation places a non-trivial lower bound on the joint (information-theoretical) uncertainty about the outcomes of simoultaneous measurements of A and B in any quantum state.

Therefore, it properly expresses the physical contents of the uncertainty principle for our two observables, i.e., the impossibility of simoultaneously having complete *information* about the values of a pair of incompatible observables, unlike the original Heisenberg inequality for standard deviations (and also the Robertson form we discussed before).

The concavity property of entropy implies that the state-independent bound H_{AB} is attained for pure states, although stronger bounds may be achieved if only restricted classes of (mixed) states are considered.

The already improved bound proposed by Maassen and Uffink (2.3) is generally not optimal, and this means that in most cases there is no quantum state for which it is attained.

The problem of obtaining the optimal lower bound on the Shannon entropy sum for two arbitrary observables is not an easy one. A. J. M. Garrett and S. F. Gull [46] were able to find the exact lower bound on the entropy sum, in the simplest particular case when the Hilbert space is two dimensional and the transformation matrix relating the two observables A and B is real, i.e.:

$$(\langle a_i | b_j \rangle) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{2.5}$$

In this particular case, Garrett and Gull were able to find the exact lower bound on the Shannon entropy sum over the set of pure states whose representation in the basis $\{|a_1\rangle, |a_2\rangle\}$ is a real matrix.

Next, Sánches [44] proved that the bound found by Garrett and Gull is actually the exact lower bound on the entropy sum, and, moreover, not only for the particular case of a real transformation matrix, but also for an arbitrary pair of observables in two-dimensional Hilbert space, whose transformation matrix is generally complex:

$$(\langle a_i|b_j\rangle) = \begin{bmatrix} e^{i\eta}\cos\theta & -e^{-i\mu}\sin\theta \\ e^{i\mu}\sin\theta & e^{-i\eta}\cos\theta \end{bmatrix}$$

2.1.3 Garrett and Gull approach to EUR

We are now going to outline the approach followed by Garrett and Gull in their paper from 1990 [46].

An arbitrary pure state $|\phi\rangle$ is represented in the basis $\{|a_1\rangle, |a_2\rangle\}$, up to an irrelevant costant phase factor, by a column matrix of the form:

$$\begin{bmatrix} e^{i\alpha}\sin\gamma \\ \cos\gamma \end{bmatrix} \tag{2.6}$$

One can then define the probabilities of obtaining the eigenvalues a_1 and a_2 in a measurement of observable A, respectively, as:

$$\begin{cases} p_1(A) = \sin^2\gamma \\ p_2(A) = \cos^2\gamma \end{cases}$$

In the particular case when the transformation matrix is real, the probabilities of obtaining the eigenvalues b_1 and b_2 for the observable B are calculated form the representation of the state vector $|\phi\rangle$ in the basis $\{|b_1\rangle, |b_2\rangle\}$, which is obtained by multiplying the transformation matrix (2.5) and the column matrix (2.6):

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} e^{i\alpha\sin\gamma} \\ \cos\gamma \end{bmatrix} = \begin{bmatrix} e^{i\alpha}\cos\theta\sin\gamma - \sin\theta\cos\gamma \\ e^{i\alpha}\sin\theta\sin\gamma + \cos\theta\cos\gamma \end{bmatrix}$$

One than obtains:

$$\begin{aligned} p_1(B) &= \sin^2\gamma\cos^2\theta + \sin^2\theta\cos^2\gamma - 2\cos\alpha\sin\gamma\cos\gamma\sin\theta\cos\theta \\ &= \sin^2(\gamma-\theta) + \sin^2(\frac{1}{2}\alpha)\sin 2\gamma\sin\theta \end{aligned} \tag{2.7}$$

$$\begin{aligned} p_2(B) &= \sin^2\gamma\sin^2\theta + \cos^2\theta\cos^2\gamma + 2\cos\alpha\sin\gamma\cos\gamma\sin\theta\cos\theta \\ &= \cos^2(\gamma-\theta) - \sin^2(\frac{1}{2}\alpha)\sin 2\gamma\sin\theta \end{aligned} \tag{2.8}$$

One then sees that when $\alpha = 0$, i.e., when the state vector $|\phi\rangle$ is represented in the basis $\{|a_1\rangle, |a_2\rangle\}$ by a real matrix, the probabilities referring to observable B turn out to be in a way simpler form:

$$\begin{cases} p_1(B) = \sin^2(\gamma - \theta) \\ p_2(B) = \cos^2(\gamma - \theta) \end{cases}$$

It follows that, if we denote for brevity the Shannon entropy sum as $H(A) + H(B) \equiv \Sigma(\gamma, \theta)$, we have for the case $\alpha = 0$ the following expression, considered by Garrett and Gull:

$$\begin{aligned} H(A) + H(B) = & -\cos^2\gamma \log\cos^2\gamma - \sin^2\gamma \log\sin^2\gamma \\ & -\cos^2(\gamma-\theta)\log\cos^2(\gamma-\theta) - \sin^2(\gamma-\theta)\log\sin^2(\gamma-\theta) \\ \equiv & \ \Sigma(\gamma, \theta) \end{aligned} \tag{2.9}$$

There is a minimum for Σ, over all θ and γ, of value zero, if A and B have at least one common eigenvector and the wavefunction $|\phi\rangle$ is directed along it.

$\Sigma(\gamma, \theta)$ achieves a maximum when all eigenvectors of A and B coincide, that is, when they commute, and the vector state is choosen to have equal components along each eigenvector. In our case, this occurs when $\theta = 0$ or π, and $p_1 = p_2 = \frac{1}{2}$, at $\gamma = \frac{1}{4}\pi$ or $\frac{5}{4}\pi$, and the maximum has value $2\log 2$.

One is now concerned to find the minimum on the entropy sum, i.e., to find stationary points over γ at fixed θ, and clearly these are bounded by the stationary points over both variables.

The function $\Sigma(\gamma, \theta)$ follows the important periodicity property, namely:

$$\Sigma(\gamma, \theta) = \Sigma(\gamma + \frac{1}{2}\pi m, \theta + \frac{1}{2}\pi n),$$

where m and n are arbitrary integers.

Therefore, one can see that it is only necessary to examine $\Sigma(\gamma, \theta)$ in intervals of $\pi/2$ in γ and in θ.

Closer examination reveals that we need only to study it in a smaller square of side $1/4\pi$ and then be able to reconstruct it everywhere; although, for the sake of clarity, we shall restrict our tabulations to the intervals $0 \leq \gamma \leq \frac{\pi}{2}$, $0 \leq \theta \leq \frac{\pi}{2}$ without any loss of generality.

Two more symmetry properties of $\Sigma(\gamma, \theta)$, the following:

$$\Sigma(\frac{1}{4}\pi + \frac{1}{2}\theta + \gamma', \theta) = \Sigma(\frac{1}{4}\pi + \frac{1}{2}\theta - \gamma', \theta) \tag{2.10}$$

$$\Sigma(\frac{1}{2}\theta + \gamma', \theta) = \Sigma(\frac{1}{2}\theta - \gamma', \theta) \tag{2.11}$$

imply that $\Sigma(\gamma, \theta)$ has stationary points in γ at $\gamma = 1/2(1/2\pi + \theta)$ and $\gamma = 1/2\,\theta$.

For our knowledge of the bound over γ and θ jointly, it is obvious that, at least for small values of θ, these stationary points are respectively a maximum and a minimum.

Let us now investigate what happens at larger values of θ. The first two derivatives of the entropy sum with respect to γ are:

$$\frac{\partial\Sigma}{\partial\gamma}(\gamma, \theta) = \sin 2\gamma \log\left(\frac{1+\cos 2\gamma}{1-\cos 2\gamma}\right) + [\text{same with } \gamma \to (\gamma - \theta)] \tag{2.12}$$

$$\frac{\partial^2\Sigma}{\partial\gamma^2}(\gamma, \theta) = 2\cos 2\gamma \log\left(\frac{1+\cos 2\gamma}{1-\cos 2\gamma}\right) + [\text{same with } \gamma \to (\gamma - \theta)] - 8 \tag{2.13}$$

Here we have assumed that logarithms are natural. The curvature of $\Sigma(\gamma, \theta)$ with respect to γ is infinite at $\gamma = 0$ and $\gamma = \theta$. The stationary point at $\gamma = 1/2\,\theta$ has value: $\Sigma(\gamma = \theta/2, \theta)$, as we will see later in this chapter.

The extrema of $\Sigma(\gamma, \theta)$ with respect to γ for fixed θ were than studied by Garrett and Gull, who obtained the explicit expression of the minimum, which we are going to denote by:

$$\Sigma_{INF}(\gamma) = \Sigma(\gamma_{INF}, \theta),$$

as a function of γ_{INF}, given by:

$$\gamma_{INF} = \begin{cases} \frac{1}{2}\theta, & \text{for} \quad 0 \le \theta \le \theta^* \\ f(\theta), & \text{for} \quad \theta^* \le \theta \le \frac{1}{2}\pi - \theta^* \\ \frac{1}{2}\theta + \frac{1}{4}\pi, & \text{for} \quad \frac{1}{2}\pi - \theta^* \le \theta \le \frac{1}{2}\pi, \end{cases} \tag{2.14}$$

where θ^* is defined as the solution of the following trascendental equation:

$$\cos x \log\left(\frac{1+\cos x}{1-\cos x}\right) = 2 \tag{2.15}$$

satisfying the condition $0 < x < \frac{\pi}{4}$. It is found that its approximate numerical value is $\theta^* \approx 0.585\ rad = 33.5°$.

We could say that θ^* represents a critical value of θ, in correspondence of which the number of absolute minima of the entropy sum in the interval $0 < \theta < \pi/2$, changes from one to two, creating a phenomenon of "bifurcation". [47]

Moreover, the bivaluated function $f(\theta)$ is implicitly defined by the conditions:

$$\left.\frac{\partial\Sigma(\gamma,\theta)}{\partial\gamma}\right|_{\gamma=f(\theta)} = 0, \tag{2.16}$$

$$f(\theta) \neq \frac{1}{2}\theta, \frac{1}{2}\theta + \frac{1}{4}\pi. \tag{2.17}$$

We are given a simple analytical expression (depending only on the angle θ) for the lower bound on the entropy sum in the particular case when θ does not belong to the interval $[\theta^*, \frac{1}{2}\pi - \theta^*]$, as follows:

$$\begin{aligned}\Sigma_{INF}(\theta) &= -(1+\cos\theta)\log\left(\frac{1+\cos\theta}{2}\right) - (1-\cos\theta)\log\left(\frac{1-\cos\theta}{2}\right)\\ &= -2\cos^2\left(\frac{\theta}{2}\right)\log\cos^2\left(\frac{\theta}{2}\right) - 2\sin^2\left(\frac{\theta}{2}\right)\log\sin^2\left(\frac{\theta}{2}\right)\\ &= -\theta^2\log\theta + O(\theta^2), \quad \theta \ll 1\end{aligned}$$

while otherwise it has to be calculated numerically.

In figure (2.3), Σ_{INF} is plotted in green against θ (obtained by minimizing the expression for $\Sigma(\gamma,\theta)$ in (2.9) with respect to γ for $0 < \gamma < \pi/2$), for $0 < \theta < \pi/2$, where the vertical black lines show the interval of definition of the function $f(\theta)$, i.e. $\theta^* \leq \theta \leq \pi/2 - \theta^*$, while the orange and blue curves are the entropy sum $H(A) + H(B) \equiv \Sigma(\gamma,\theta)$, for $\gamma = \theta/2$ and $\gamma = \theta/2 + \pi/4$, respectively.

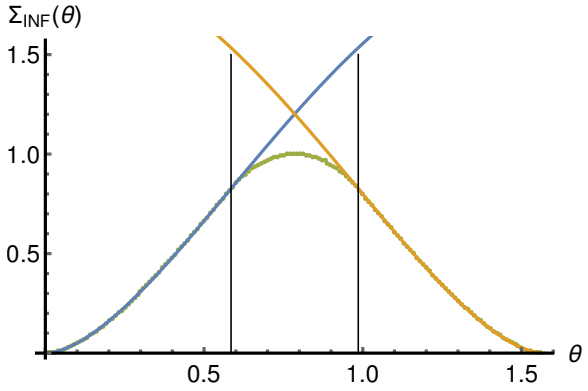

Figure 2.3: The bound Σ_{INF} is plotted as a function of the angle θ (green), for $0 < \theta < \pi/2$ and $0 < \gamma < \pi/2$. The vertical black lines show the interval of definition of the function $f(\theta)$, i.e. $\theta^* \leq \theta \leq \pi/2 - \theta^*$. In orange and blue is shown the entropy sum $H(A) + H(B) \equiv \Sigma(\gamma,\theta)$, for the stationary points $\gamma = \theta/2$ and $\gamma = \theta/2 + \pi/4$, respectively.

In figure (2.4) Garrett and Gull compared the infimum Σ_{INF}, the Maassen-Uffink bound Σ_{MU} and the Deutsch bound Σ_D plotted together against the angle θ.

It is evident from this figure that the Maassen and Uffink's bound is clearly better than the Deutsch bound (in fact, it has been found that $\sum_{MU} \geq \sum_{D}$ [46]), however it is clear that there is still prospect of further improving.

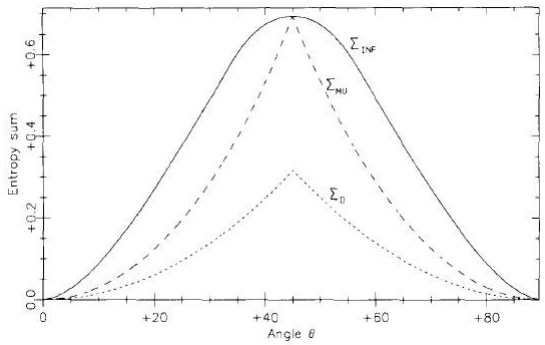

Figure 2.4: Comparison of the bound Σ_{INF} with the Maassen-Uffink bound Σ_{MU} and the Deutsch bound Σ_D, all plotted together against the angle θ. Figure taken from [46].

One asks then if this minima are global, i.e. if they are infima.

In figure (2.5) we show a plot of γ_{INF} versus θ (obtained by plotting the values of γ for which $\Sigma(\gamma, \theta)$ is minimized), for $0 < \theta < \pi/2$. The blue and orange lines represent the two values for γ, $\gamma = \theta/2$ and $\gamma = 1/2(\pi/2 + \theta)$ respectively. The vertical black lines, again, mark the interval $\theta^* \leq \theta \leq \pi/2 - \theta^*$.

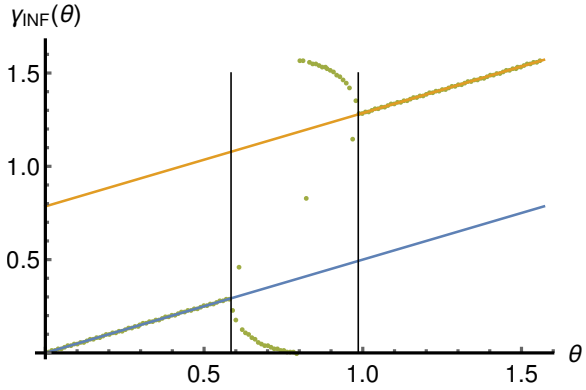

Figure 2.5: γ_{INF} is plotted against θ, for $0 < \theta < \pi/2$, showing an interesting bifurcation, as θ increases, the minimum at $\gamma = \theta/2$ bends to create a maximum between two minima, that, because of the symmetry properties of $\Sigma(\gamma, \theta)$, appear to be symmetrical about the curvature and thus have equal value. Contrariwise, the maximum at $\gamma = 1/2(\pi/2 + \theta)$, and its side minima, merge into a single minimum as θ increases. The blue and orange lines represent $\gamma = \theta/2$ and $\gamma = 1/2(\pi/2 + \theta)$ respectively. The vertical black lines, again, mark the interval $\theta^* \leq \theta \leq \pi/2 - \theta^*$.

As we can see in fig. 2.5, when $\theta^* \leq \theta \leq \pi/2 - \theta^*$ the bifurcation phenomenon occurs. Thus, as θ increases, the minimum at $\gamma = \theta/2$ bends to create a maximum between two minima. However, expressions for these side minima have not yet been found. Still, from the symmetry properties of $\Sigma(\gamma, \theta)$ we can say that they are symmetrical about the curvature and thus have equal value. In the converse process, the maximum at $\gamma = 1/2(\pi/2 + \theta)$, and its side minima, merge into a single minimum as θ increases.

From the first formulation we gave of the two probabilities $p_1(B)$ (2.7) and $p_2(B)$ (2.8) we see that, when one considers arbitrary pure states (i.e. in general with $\alpha \neq 0$), in the case of real transformation matrices, $H(A) + H(B)$ is in fact a function of three variables (namely θ, γ and α) and we are faced with the problem of finding the minimum of this function with respect to γ and α for fixed θ. On the other hand, in the most general case of a complex transformation matrix, the entropy sum becomes a function of five variables (namely $\theta, \gamma, \alpha, \eta$ and μ), and we should find the minimum of this function for fixed θ with respect to the other four variables, by means of a direct variational calculation, generalizing the approach by Garrett and Gull. However, Sánches [44] solved the problem in a simpler way, using purities or *Hilbert-Schmidt norms*, but this is not the right place to deal with this subject.

2.1.4 Choice of the observables

Having now discussed the topic of EURs, let us go back to the scope of our work. Before concentrating on classical correlations as a resource for the detection of entanglement, let us briefly make a clarification regarding the choice of the observables.

We observe that in the case of pure states the obvious choice would be the Schmidt bases and their respective Fourier bases. On the other hand, for mixed states, one can diagonalise the density matrix, identify the eigenvector with the largest weight and use the Schmidt basis, and its respective Fourier bases. Otherwise, one can choose the basis that diagonalises the reduced density matrices. While there it is not guaranteed that these choises will allow one to implement the procedure, they are the ones that may uncover the most correlations.
We note that this method is simple to implement for systems of arbitrary dimensions, since it requires only indipendent measurements of two local observables on the two systems, as well as robust, since it guarantees that the systems are entangled if the conditions given at the beginning of section 2.1 are satisfied for *any* couple of complementary observables [39].

In regards to the necessary measurements, the observable $A \otimes B$ corresponds to σ_z, namely the computational basis $\{|0\rangle, |1\rangle\}$ for each qubit.

The Fourier basis can be expressed as tensor products of single-qubit states. The Fourier basis in arbitrary dimension identifies an observable complementary to the com-

putational basis.

As for the observable $C \otimes D$, however, in the case considered here, there are complementary bases that are simpler to access experimentally, namely the bases where one measures σ_x or σ_y on each qubit.

Let us now consider for the moment the basis σ_x, which is given by $|c_k\rangle$, with $k = 0, ..., d-1$, obtained by expressing the binary digits of k in the $\{|+\rangle, |-\rangle\}$ basis, i.e.:

$$|c_k\rangle = \frac{[(|0\rangle + (-1)^{\gamma_1} |1\rangle)(|0\rangle + (-1)^{\gamma_2} |1\rangle)...]}{\sqrt{2^n}}$$

where γ_l are the bits of the number k. The σ_x basis is complementary to the computational basis, since $|\langle j|c_k\rangle|^2 = 1/2^n = 1/d, \forall j, k$.

2.2 Classical correlations

In the paper by Maccone et al. [1], they provide an interpretation of entanglement based on classical correlations between measurement outcomes of complementary properties, i.e., states that have correlations beyond a certain threshold are entangled. However, they would point out that the reverse is not true. They also show that, surprisingly, all separable nonclassical states exhibit smaller correlations for complementary observables than some strictly classical states. They use various measures of classical correlations, such as the Pearson correlation coefficient, the sum of conditional probabilities or the mutual information. In what follows, we are going to concentrate on the latter.

Definition: Two properties of a quantum state are called *complementary* if they are such that, if one knows the value of one property, all possible values of the other property are equiprobable. More rigorously, let $|a_i\rangle$ represent the eigenstates corresponding to possible values of a nondegenerate property $A = \sum_i f(a_i) |a_i\rangle \langle a_i|$, and $|c_i\rangle$ the eigenstates of a nondegenerate property $C = \sum_j g(c_j) |c_j\rangle \langle c_j|$ (with f and g arbitrary bijective functions). Then $\mathcal{A}$ and $\mathcal{B}$ are said to be complementary properties if for all *i, j* we have

$$|\langle a_i|c_j\rangle|^2 = 1/d$$

where d is the Hilbert space dimension. Clearly, complementary properties with this definition identify two mutually unbiased bases (for short, **MUBs**) [48]. In this article they study what classical correlations in the measurements of these complementary properties tell us about the quantum correlations of the state of the system.

2.2.1 Complementary correlations

Complementary correlations can reveal the genuine quantum correlations present in a composite quantum system.
Consider two systems of finite dimension d and two observables $A \otimes B$ and $C \otimes D$ where A and C are complementary on the first system and B and D on the second (as shown in figure (2.1)), i.e.:

$$\begin{aligned} |\langle a_i|c_j\rangle|^2 &= 1/d \\ |\langle b_i|d_j\rangle|^2 &= 1/d \end{aligned} \tag{2.18}$$

for all eigenstates of A and C and of B and D, respectively.

Therefore, the correlations between the measurement results of A and B can be quantified by some correlation measure, that we can denote with χ_{AB}, and in the same way the correlations relative to the observables C and D will be indicated with χ_{CD}.

Hence, generally, a measure of the overall correlation of the initial state, which we name the *"complementary correlations"*, can be quantified as the sum of the absolute value of the two measures or as their product, although the latter is typically a weaker measure than the former, since an upper bound for the sum implies an upper bound for the product, i.e.:

$$\left(|\chi_{AB}|^{1/2} - |\chi_{CD}|^{1/2}\right)^2 \geq 0 \Rightarrow |\chi_{AB}| - 2|\chi_{AB}|^{1/2}|\chi_{CD}|^{1/2} + |\chi_{CD}| \geq 0$$
$$\Rightarrow 2\sqrt{|\chi_{AB}\chi_{CD}|} \leq |\chi_{AB}| + |\chi_{CD}|$$

Hence, we are going to consider the sum of the correlations for complementary observables $|\chi_{AB}| + |\chi_{CD}|$ as an evaluation of the complementary correlations.

We can quantify the correlations between the results of the measurements of A and B with the mutual information, which we are now going to define.

2.2.2 Mutual Information

As we said in subsection 2.1.2, entropy is seen as a means to quantify lack of information, hence one could think that measures that can quantify the *presence* of information or correlation would also be useful. In this sense, we define the *mutual information* I_{AB}, which quantifies the correlation existing between two random variables A and B, and is given by [49]:

$$I_{AB} \coloneqq H(A) + H(B) - H(A, B)$$
$$\equiv H(A) - H(A|B)$$

where $H(A)$ is the Shannon entropy (as defined in equation (2.4)) of the probabilities of the measurement outcomes of the first system and $H(A|B)$ is the conditional entropy of the outcomes of the first system conditioned on the second. The complementary correlations for our system are then $I_{AB} + I_{CD}$.

Mutual information is indeed a relavant quantity in information theory [45], since it has the central role of quantifying the information *gained*, or equivalently the "loss of ignorance", about a certain observable A when access to B is given.

If we think to Shannon entropy and mutual information in terms of messages communication, we can think that observable A belongs to a transmitter, Alice, and observable B to a receiver, Bob.

The Shannon entropy H(A) quantifies how much information is transmitted, on the average, by a unit of information (the "message" a) drawn from the ensemble A, since it

represents how many bits are required to encode that information. On the other hand, the mutual information I_{AB} quantifies the correlation between the two messages. Suppose Bob receives the message sent by Alice, but the channel of communication is noisy. Thus the message sent from Alice (a) might be different from the message received by Bob (b). [8]

We can characterize the noisy channel with the conditional probability $p(a|b)$, namely the probability that b is received when a has been sent. Let $p(a)$ be the *a priori* probability that message a is sent. Then we ask how much information we gain (i.e. Bob gains) about the original message a when we receive b.

As we said before, the Shannon entropy $H(A)$ quantifies the a priori ignorance about the message a, *before* it is received. But after reading the received value of b, we can use Bayes' rule to write the (conditional) probability distribution:

$$p(a|b) = \frac{p(b|a)p(a)}{p(b)},$$

where $p(b|a)$ is known from the properties of the channel and thus we can express $p(b)$ as $p(b) = \sum_a p(b|a)p(x)$.

Therefore, given the message Bob has received, the quantity of information (bits) required to specify a particular string of information is expressed by the *conditional entropy*, i.e.:

$$H(A|B) = \langle -\log p(a|b) \rangle . \tag{2.19}$$

Since $p(a|b) = p(a,b)/p(b)$, we can rewrite (2.19) as:

$$H(A|B) = \langle -\log p(a,b) + \log p(b) \rangle = H(A,B) - H(B),$$

and in the same way we have:

$$H(B|A) = \langle -\log p(b|a) \rangle = \langle -\log p(a,b) + \log p(x) \rangle = H(A,B) - H(A).$$

Then, one can interpret $H(A|B)$ as the quantity of additional information needed to specify both a and b when the latter is known. Hence, it is clear that it cannot be negative.

Follows logically that the information about A that we gain when we receive B might be represented by the *decrease* of information needed to specify A when B is known, or more rigorously:

$$\begin{aligned} I_{AB} &\equiv H(A) - H(A|B) \\ &= H(A) + H(B) - H(A,B) \\ &= H(B) - H(B|A) \end{aligned}$$

This is the definition of *mutual information* and we can see that it is symmetric under interchange of A and B, that is, we learn as much about A by learning B as about B by learning A.

Moreover, learning A can never reduce our knowledge of B, and vice versa, therefore I_{AB} must be obviously non-negative.

Finally, if A and B are completely uncorrelated, one has $p(a,b) = p(a)p(b)$ and naturally

$$I_{AB} \equiv \langle \log \left(p(a,b)/p(a)p(b) \right) \rangle = 0,$$

that is, we obviously gain no information about A by receiving B if they are not correlated.

2.2.3 Sufficient condition for entanglement using mutual information

Maccone et al. [1] proved that if the following relation holds:

$$I_{AB} + I_{CD} > \log_2 d, \tag{2.20}$$

then the state of the two systems is *entangled*. This theorem can be stated in an equivalent form saying that if $I_{AB} + I_{CD} \leq \log_2 d$ then the state is *separable*.

We have already defined mutual information as:

$$I_{AB} \equiv H(A) - H(A|B),$$

where $H(A)$ is the (Shannon) entropy of the A measurement outcomes and $H(A|B)$ is the socalled *conditional entropy* of the A outcomes (given B), which can also be written as

$$H(A|B) = -\sum_{a,b} p(a|b)p(b) \log_2 p(a|b) = \sum_b p(b) H(A|B=b),$$

where

$$H(A|B=b) = -\sum_a p(a|b) \log_2 p(a|b)$$

is the entropy of the probability distribution $p(a|b)$ for fixed b.

Now, a separable state is defined as:

$$\rho = \sum_l p_l \rho_l \otimes \sigma_l.$$

The *conditional state* $\rho^{(b)}$ one obtains when the result b is obtained from a measurement of B on the second subsystem is given by:

$$\rho^{(b)} = \sum_l \beta_l^{(b)} \rho_l,$$

with

$$\beta_l^{(b)} = \frac{p_l \langle b|\sigma_l|b\rangle}{\sum_{l'} p_{l'} \langle b|\sigma_{l'}|b\rangle}$$

We note that the term in the denominator in the above expression for $\beta_l^{(b)}$ is actually $p(b)$, namely the probabilty of getting the outcome b when measuring B on the second subsystem.

Taking into account the concavity propriety of entropy, one obtains:

$$H(A|B=b) = H(A)_{\rho^{(b)}} \geq \sum_l \beta_l^{(b)} H(A)_{\rho_l}$$
$$\Rightarrow H(A|B) = \sum_b p(b) H(A|B=b) \geq \sum_l p_l H(A)_{\rho_l}$$

where $H(X)_\rho$ denotes the Shannon entropy of a measurement of X on the state ρ.

With an analougus reasonment for the observables C and D one obtains:

$$H(C|D) \geq \sum_l p_l H(C)_{\rho_l}. \tag{2.21}$$

Let us now consider Maassen-Uffink EUR (2.3), according to which for any state ρ we have:

$$H(A)_\rho + H(C)_\rho \geq -2 \log c,$$

with $c = max_{j,k} |\langle a_j|c_k\rangle|$.

In the particular case of complementary observables (i.e. for MUBs such that $|\langle a_j|c_k\rangle|^2 = 1/d$), we will have that the overlap matrix is flat (namely $c_{jk} = |\langle a_j|c_k\rangle| = 1/\sqrt{d}, \forall j,k$). Hence:

$$-2\log c = -2\log|\langle a_j|c_k\rangle| = -\log|\langle a_j|c_k\rangle|^2 = -\log 1/d = \log_2 d$$

Therefore the lower bound on the uncertainty will become maximal:

$$H(A) + H(C) \geq \log d.$$

We note that this is a necessary and sufficient condition: indeed, we have that $c = 1/d$ if and only if the two bases are mutually unbiased, and therefore MUBs uniquely give the

strongest uncertainty bound in this case. For general observables the overlap matrix is not necessarily flat and the asymmetry of the overlap matrix elements can be quantified by taking the maximum over all the elements of the matrix itself. This is clear from the fact that, if the maximum entry of the overlap matrix is $1/d$, then all entries in the matrix must be $1/d$.

This means that :

$$H(A|B) + H(C|D) \geq \sum_l p_l [H(A)_{\rho_l} + H(C)_{\rho_l}] \geq \log_2 d.$$

where the first inequality is due to the concavity of the entropy and the second to (2.3). The above chain of inequalities and the fact that

$$\begin{cases} H(A) \leq \log_2 d \\ H(C) \leq \log_2 d \end{cases}$$

imply that

$$\begin{aligned} I_{AB} + I_{CD} &= H(A) - H(A|B) + H(C) - H(C|D) \\ &\leq 2\log_2 d - \log_d = \log_2 d \end{aligned}$$

In our case $d = 2$, thus $|\langle a_j|c_k\rangle| = 1/\sqrt{d} = 1/\sqrt{2}$, so we have $\log_2 2 = 1$, and then if:

$$I_{AB} + I_{CD} \leq 1$$

we say that our state ρ is separable.

Then we also have that a state is *maximally entangled* if:

$$I_{AB} + I_{CD} = 2;$$

and, moreover, a state is *entangled* if:

$$I_{AB} + I_{CD} > 1. \tag{2.22}$$

2.3 Entanglement witness based on classical correlations

Joining this last result from the paper by Maccone et al. [1] with that from Sánches [44] for the minimum of entropy sum, namely:

$$H(A) + H(C) \geq \Sigma_{INF}(\theta)$$

for our case of a two-qubit system, we obtain:

$$\begin{aligned} I_{AB} + I_{CD} &= H(A) - H(A|B) + H(C) - H(C|D) \\ &\leq 2\log_2 d - \Sigma_{INF} \end{aligned}$$

In detail, when A and C are complementary, we have $\Sigma_{INF} = 1$ (and as already noticed: $\log_2 2 = 1$), hence:

$$I_{AB} + I_{CD} \leq 2 - 1 = 1 \tag{2.23}$$

otherwise, we would have $\Sigma < 1$, and therefore:

$$I_{AB} + I_{CD} \leq 2 - \Sigma(\theta_1) \quad \text{or} \quad I_{AB} + I_{CD} \leq 2 - \Sigma(\theta_2) \tag{2.24}$$

or equivalently

$$I_{AB} + I_{CD} \leq 1 + \epsilon_{AC} \quad \text{or} \quad I_{AB} + I_{CD} \leq 1 + \epsilon_{BD} \tag{2.25}$$

depending on which pair of observables we are evaluating. We would finally consider the smaller of them in order to obtain the strongest possible bound, and therefore have in general:

$$I_{AB} + I_{CD} \leq 1 + \epsilon. \tag{2.26}$$

2.4 Extension to more than two observables

As mentioned in [1], the studies and the results discussed to this point could be immediately extended to more than two complementary observables. As proved in [48], all systems have at least three complementary observables, and it is known that there are $d + 1$ MUBs for d-dimensional systems if d is a power of a prime. The generalisation is apparently straightforward, and it could be performed by calculating the correlations of all the known complementary observables and considering the sum of the two largest ones.

For mutual information, in particular, we can extend the condition (2.20), found for the case of two qubits, to the more general one, as follows:

$$\max(I_{AB}, I_{CD}, I_{EF}, ...) + \max_2(I_{AB}, I_{CD}, I_{EF}, ...) > \log_2 d$$

where $\max_2$ denotes the second largest term in the list, and where $A \otimes B$, $C \otimes D$, $E \otimes F$, etc., are all observables complementary to each other.

In the Supplemental Material of reference [1], it is proved, moreover, at least in the case of qubits, that the bound in (2.20) can be made stronger by adding correlations for a

third complementary observable, resulting in a significative improvement in the efficiency of the present entanglement detection method.

In fact, we see that the condition (2.22) for the entanglement of pair of qubits in $d = 2$ can be made stronger by adding a third MUB for each qubit, say E and F. For a separable two-qubit state the argument given in subsection 2.2.3 until equation (2.21) can be actually applied also for the additional pair of bases EF, and therefore we can write:

$$H(A|B) + H(C|D) + H(E|F) \geq \sum_l p_l[H(A)_{\rho_l} + H(C)_{\rho_l} + H(E)_{\rho_l}]$$

In reference [50], moreover, a generalization of the inequality by Maassen and Uffink (2.3) to sets of many (more than two) observables has been outlined. Specifically, for a set of mutually non-commuting Hermitian operators $\{A_k\}, k = 1, ..., M$, with $M > 2$, one has:

$$\sum_{k=1}^{M} H(A_k) \geq \inf_{|\psi\rangle} \left(\sum_{k=1}^{M} H(A_k) \right) > 0,$$

where $H(A_k)$ is the entropy corresponding to the probability distribution $p_i(A_k)$.

The latter can be obtained from the $\frac{1}{2}M(M-1)$ uncertainty relations of the form (2.3) for each pair of operators of the set, i.e.:

$$H(A_i) + H(A_j) \geq \inf_{|\psi\rangle}[H(A_i) + H(A_j)] \equiv H_{ij}, \quad (\text{for } i < j).$$

In the case we are considering here, we can say that for A, C, and E, i.e. the complementary observables for system 1, we will obtain, for any qubit state ρ, the following:

$$H(A)_\rho + H(C)_\rho + H(E)_\rho \geq 2,$$

and then we conclude that separable states fulfill the following condition:

$$I_{AB} + I_{CD} + I_{EF} \leq 1.$$

Numerical tests from reference [1] show striking improvements in the power of the criterion for the detection of entanglement.

More recently, in reference [51] they have developed the analysis on complementary correlations, concentrating on the case of three MUBs. Specifically, regarding mutual information, they found that the complementary correlations in the global system ($I_{AB} + I_{CD} + I_{EF}$) are such that the following cases hold:

- the state of a bipartite quantum system is *maximally entangled* if and only if there exist three MUBs such that $I_{AB} + I_{CD} + I_{EF} = 3 \log d$;
- if $I_{AB} + I_{CD} + I_{EF} > \log d$, the state of the system is *entangled*;
- if $I_{AB} + I_{CD} + I_{EF} \leq \log d$, then the state is *separable*.

Chapter 3

Detecting entanglement in qubit systems

In this chapter we are going to present our results, in the form of numerical tests for our criterion for the detection of entanglement.

The results presented here were obtained through numerical simulations using the *Wolfram Mathematica* software.

We started by considering a considerable number (approximately 10^5) of random states, obtained by generating 4×4 random unitary matrices.

Then, we naturally proceeded by putting our criterion to the test by quantifying the goodness of its performance in entanglement detection in terms of percentage of detected states, as a function of the angle θ.

Successively, we repeated our simulations in the case of a noisy environment, by applying quantum noise channels to the original randomly generated states.
Finally, we analysed its robustness to noise and checked its invariance under basis change.

Let us note that, as a guide for comparison for the performances of our method, we made use of the PPT criterion, as defined in section 1.2.2 to evaluate whether each state is entangled or not, since in this case it represents a necessary and sufficient condition [20].

In this way, our results make sense when they are compared to those obtained with the PPT criterion, that is guaranteed to detect 100% of entangled states. Therefore we will obtain results in terms of the "relative" rates of detection of entangled states, compared to the "full detection" (1/1), guaranteed by the PPT criterion, to have a measure of the "goodness" of our criterion. Otherwise we would only have had absolute numbers, with little meaning.

Examples of correlations with mutual information

We are going to show some examples of the correlation represented by the sum of mutual information $I_{AB} + I_{CD}$ plotted as a function of the parameter p for the following families of p-dependent two-qubit states:

$$\rho_1 = \frac{p}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|) + \frac{1-p}{2}(|{+}{+}\rangle\langle {+}{+}| + |{-}{-}\rangle\langle {-}{-}|) \tag{3.1}$$

$$\rho_2 = p\,|\Phi^+\rangle\langle\Phi^+| + (1-p)\frac{\mathbf{I}}{4} \tag{3.2}$$

$$\rho_3 = p\,|\Phi^+\rangle\langle\Phi^+| + (1-p)\,|\Phi^-\rangle\langle\Phi^-| \tag{3.3}$$

We certify the presence of entanglement for the three families of states listed above, by measuring the complementary observables $A \otimes B$ and $C \otimes D$, emphasizing in every example the threshold above which all states are entangled.

In particular, we notice that the family of states ρ_1 "mixes" a maximally entangled state with a mixed state and are always separable for $p \neq 0, 1$. The ρ_2 states are the so-called Werner states, which are an example of mixed entangled states.

Let us define the *"correlation sum"* C_s as the sum of mutual information for the two systems, namely $C_s = I_{AB} + I_{CD}$.

In figure (3.1) we start with the original case of correlations in the case of complementary observables. As we said in section 2.3, when A and C are complementary we have $\Sigma_{INF} = 1$ and therefore we have, from (2.23), that in this case our bound will be:

$$C_s = I_{AB} + I_{CD} \leq 1.$$

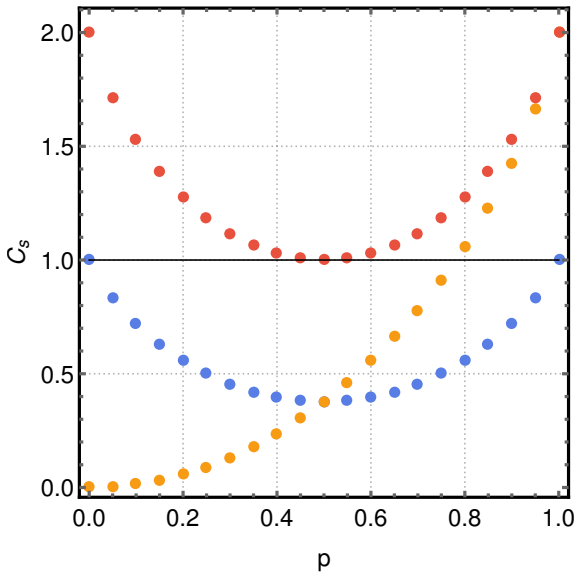

Figure 3.1: Plot of the mutual information sum $C_s = I_{AB} + I_{CD}$ for the three different families of states ρ_1, ρ_2 and ρ_3, in the case of mutually unbiased bases $\theta = \pi/4$. Above the threshold $C_s \leq 1$ (black line), the states are certainly entangled. Indeed, we see that ρ_1 states (in blue) are separable for $p \neq 0, 1$; the ρ_2 (in orange) states are entangled for $p > 1/3$; the ρ_3 states (in red) are entangled for $p \neq 1/2$.

For mutually unbiased bases $\theta = \pi/4$, therefore, we correctly recover the same figure found by Maccone et al. in [1]: indeed, we see that ρ_1 states (blue) are always separable for $p \neq 0, 1$; the ρ_2 (orange) states (Werner states) are entangled for $p > 1/3$; the ρ_3 states (red) are entangled for $p \neq 1/2$. Above the threshold $C_s \leq 1$ (black line), the states are certainly entangled.

We now proceeded to plot the same families of states by relaxing the condition of complementarity of the observables, hence trying different values of the angle θ and considering the relative bound as defined in (2.26).

For example in figure (3.2) we see that for $\theta = \pi/8$, the bound increases to $C_s = I_{AB} + I_{CD} \leq 2 - \Sigma_{INF} = 1.53335$: we see that the state ρ_1, in blue, stays as before under the threshold, thus it is always separable. On the other hand, fewer of the ρ_2 and especially of the ρ_3 states (which in the complementary case were entangled except for $p = 1/2$) are found above the bound.

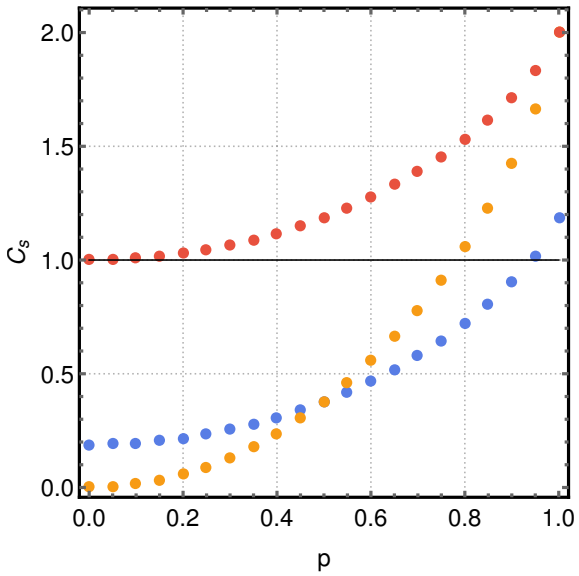

Figure 3.2: Plot of the mutual information sum $C_s = I_{AB} + I_{CD}$ for the three different families of states ρ_1, ρ_2 and ρ_3, in the case of non-complementary observables, namely with $\theta = \pi/8$. Above the threshold $C_s \leq 1.53335$ (black line), the states are certainly entangled. We see that in this case the ρ_1 states (in blue) are always separable but fewer of the ρ_2 (in orange) and the ρ_3 states (in red) are entangled, compared to the complementary case.

As a last example, we consider the rather "unfortunate" case of $\theta = \pi/20$. For such a value of the angle θ, the bound grows to the value $C_s = I_{AB} + I_{CD} \leq 2 - \Sigma_{INF} = 1.89188$. Thus, we see from figure (3.3) that only a small fraction of the ρ_3 (red) states are found above the bound, while all of the ρ_1 (blue) and ρ_2 (orange) are found below the black line.

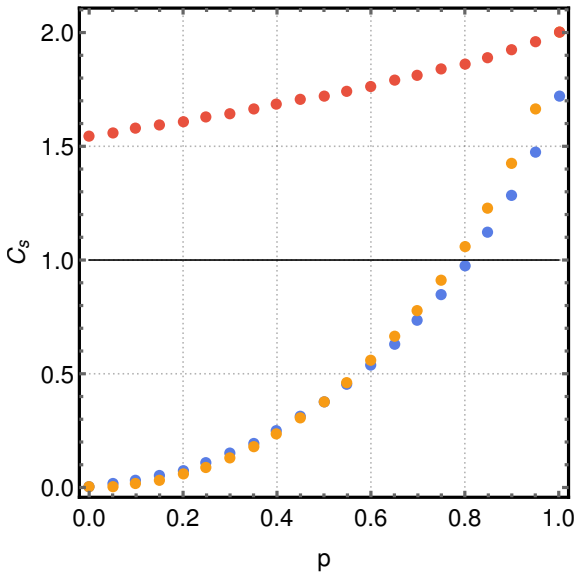

Figure 3.3: Plot of the mutual information sum $C_s = I_{AB} + I_{CD}$ for the three different families of states ρ_1, ρ_2 and ρ_3, in the case of $\theta = \pi/20$, i.e. of non-complementary observables. Above the threshold $C_s \leq 1.89188$ (black line), we know that the states must be certainly entangled. In this case, we see that all the ρ_1 sates (in blue) and the ρ_2 states (in orange) are found below the bound, and a very small fraction of the ρ_3 states (in red) are found above the black line.

3.1 Random maximally entangled states

Testing our witness on generical entangled states, however, resulted in little efficiency. Indeed, as we said at the end of subsection 1.2.1, the optimal resource for bipartite entanglement is the maximally entangled state. Therefore we are going to present here first the results on maximally entangled states to benchmark the results observed in noisy environments.

We can write maximally entangled states of a two-qubit system as:

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}} \left(|\alpha_0\rangle |\beta_0\rangle + |\alpha_1\rangle |\beta_1\rangle \right),$$

where $\{|\alpha_0\rangle, |\alpha_1\rangle\}$ and $\{|\beta_0\rangle, |\beta_1\rangle\}$ are the bases for the two qubits.

Here, we tested our entanglement witness on randomly generated maximally entangled states.

In the same context of generalization of the postulates of quantum mechanics deligned at the beginning of the first chapter, a further generalization regarding the dynamical

evolution of a system was introduced, through the formalism of *quantum operations*, which we are going to adopt in what follows. [22]

By quantum operation, we mean a map $\rho \to \varepsilon(\rho)$ which transforms a quantum state ρ into another quantum state $\varepsilon(\rho)$. This map ε must respect certain requirements to describe physically admissible operations, which are a generalization of the request of unitarity of the standard folmulation of the evolution postulate, namely:

- to preserve the Hermitian character of density operators, that is, to be positive and trace-preserving: $\varepsilon(\rho) \geq 0$ and $\mathrm{Tr}[\varepsilon(\rho)] = \mathrm{Tr}[\rho] = 1$, or at least trace non-increasing: $0 \leq \mathrm{Tr}[\varepsilon(\rho)] \leq 1$;
- to be linear: $\varepsilon(\sum_k p_k \rho_k) = \sum_k p_k \varepsilon(\rho_k)$;
- to be *completely positive (CP)*, i.e. besides being positive, to be such that if we introduce an additional system, any map of the form $\varepsilon \otimes \mathbb{I}$ acting on the extended Hilbert space is also positive. This means that the map must be physically meaningful also when acting on a portion of a larger, composite system.

A quantum operation captures the dynamic change to a state which occurs as the result of a physical process, that is, if ρ is the initial state before the process than $\varepsilon(\rho)$ will be the final state after the process took place, possibly up to some normalization factor.

We are going to use a very powerful mathematical representation for quantum operations known as the *operator-sum representation*. This method is rather abstract, but is very useful for calculations and theoretical work. We can provide a theorem stating that a map is a quantum operation if and only if it is the partial trace of a unitary evolution in a larger Hilbert space. This theorem, moreover, provides a convenient form to express the action of a quantum operation on a quantum state, namely the socalled *Kraus decomposition*, or operator-sum representation. This theorem is also known as *Kraus-Choi-Sudarshan Theorem*.

Theorem: A map ε is a quantum operation, i.e. it satisfies the requirements stated above, *if and only if* it is the partial trace of a unitary evolution on a larger Hilbert space or, equivalently, it possesses a **Kraus decomposition**, i.e. its action can be represented as $\varepsilon(\rho) = \sum_k M_k \rho M_k^\dagger$ with $\sum_k M_k^\dagger M_k = \mathbb{I}$.

The Kraus decomposition of a quantum operation, and the Kraus theorem, also allow us to have a unified picture of quantum evolution, either due to an interaction or to a measurement, being the *modification* of the state in both processes described by a set of operators M_k, satisfying $\sum_k M_k^\dagger M_k = \mathbb{I}$. In this framework, the Kraus operators of a measurement are what we have referred to as *detection operators* in a POVM.

In quantum information theory, the relevant concept of channel-state duality is extremely useful and fruitful: it refers to the correspondence between quantum channels and bipartite systems (described by density matrixes). It is often called the *Jamiołkowski-Choi isomorphism* [52].

The channel-state duality, usually manifested in the form of the Jamiołkowski-Choi isomorphism, refers to the statement that any *channel* (i.e., quantum operation, or equivalently, any linear, completely positive, trace-preserving map) from the state space of an input quantum system to that of an output system corresponds to a *bipartite state* of the tensor product of the two relevant systems. This correspondence links dynamics to kinematics, and is not merely mathematical but also has a fundamental physical meaning, profound consequences, and many applications.

It should be emphasized, however, that although the Jamiołkowski-Choi isomorphism is an injection in the sense that a channel corresponds to a unique bipartite state, the converse is not true: it is not a surjection. There are many bipartite states which cannot be represented as a channel.

Apart from the above mathematical consideration of the Jamiołkowski-Choi isomorphism, there is also a fundamental physical interpretation. Indeed, one of its formulations, known as the *Choi isomorphism*, states the equivalence between an operator on the Hilbert space $\mathcal{H}$ and a vector in $\mathcal{H} \otimes \mathcal{H}$, that is, we have $\mathcal{L}(\mathcal{H}) \equiv \mathcal{H} \otimes \mathcal{H}$, where $\mathcal{L}(\mathcal{H})$ is the space of linear operators on $\mathcal{H}$.

Physically, this means that a unitary operator, which represents the dynamics of a closed system, can be thought as *equivalent* to a vector on a tensor product Hilbert space, which represents a pure state of a bipartite system. Therefore, it can be noted that, up to normalization, through the isomorphism we could map unitary operators to maximally entangled states. In our case, we can generate 2×2 unitary matrices ($U(2)$) that, through the isomorphism, are "linearized" to 4-dimensional vectors, i.e. pure states of a 2-qubit system, which are maximally entangled.

Rigorously, one generates random unitary matrices from the classical compact group $U(N)$ with a probability distribution given by its invariant measure. The algorithm could be quite straightforwardly implemented using standard linear algebra packages, but we used here the *"QI"* package for Mathematica computer algebra system which contains the implemented function `RandomSpecialUnitary[d]` that returns a random special unitary matrix of size d.

The *"QI"* package implements many functions used in the analysis of quantum states, focusing on geometrical aspects of quantum information theory [53].

An $N \times N$ unitary matrix $U = (u_{jk})$ is by definition a matrix such that $U^\dagger U = UU^\dagger = \mathbb{I}$, where $U^\dagger = (u_{jk}^\dagger)$ is the conjugate transpose of U. Thus, in terms of the matrix elements,

we will have the constraints:

$$\sum_{k=1}^{N} u_{jk}^{\dagger} u_{kl} = \delta_{jl}$$
$$\sum_{k=1}^{N} u_{jk} u_{kl}^{\dagger} = \delta_{jl}$$

This constraints state simply that the columns (or the rows) of a unitary matrix form an orthonormal basis in $\mathbb{C}^N$ and moreover we know that the set $U(N)$ of unitary matrices forms a compact Lie group of dimension N^2. This is then made into a probability space by assigning as a distribution the unique measure invariant under group multiplication, which is known as *Haar measure*. [54]

Writing a correct and numerically stable algorithm for generating random unitary matrices may present some obstacles, since the constraints listed above imply that the matrix elements are not independent and thus are statistically correlated. For this reason, writing an explicit formula for the infinitesimal volume element of $U(N)$ can be rather complicated.

Even though one could write down, in theory, an explicit expression for the Haar measure in terms of local coordinates, it is found that regarding the generation of random matrices one would only need to know that the Haar measure is invariant and unique [54]. Moreover, the Haar measure normalized to one represents a natural choice for a probability measure on a compact group since, being invariant under group multiplication, any region of $U(N)$ carries the same weight in a group average.

On this ground, we generated 10^5 random special unitary matrices, and we are now going to show the performances of our criterion in the form of histograms representing the fraction of states as a function of the correlation sum $C_s = I_{AB} + I_{CD}$. As said before, the states which are found above the bound given by (2.26) are entangled.

In figure (3.4), we first plot the case of complementary observables,. i.e. with $\theta = \pi/4$. The bound, as we already know, is given in this case by (2.23), therefore all the states found above the bound $C_s = 1$ are entangled. We found through our numerical calculations that in this case our method detects 16.099% of the entangled states.

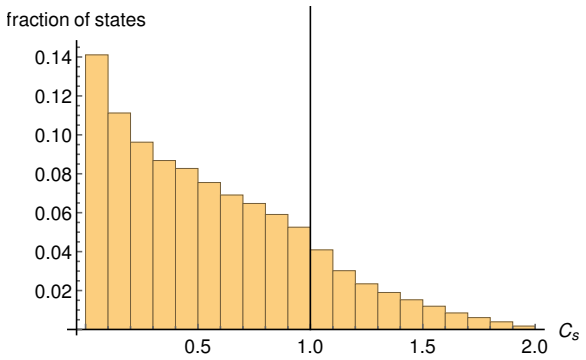

Figure 3.4: Fraction of states versus the correlation sum C_s for the case of complementary observables, i.e. $\theta = \pi/4$. The states above the bound $C_s = 1$ (that is, at the right of the black line) are entangled. In this case we detected 16.099% of the entangled states.

As we did in the last section, let us now consider non-complementary observables, and see how the results change from the complementary case.

In figure (3.5) we consider the case for $\theta = \pi/8$. In this case, the bound moves towards right according to 2.26, rising to the value $C_s = 1.53335$. Therefore, for $C_s \geq 1.53335$ we find a smaller fraction of detected entangled states, In particular, we found from our computation that only 2.057% of entangled states were detected.

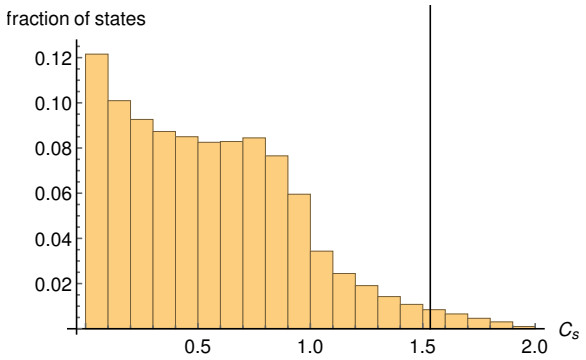

Figure 3.5: Fraction of states versus the correlation sum C_s for $\theta = \pi/8$. In this case, the bound reaches the value $C_s = 1.53335$ (black vertical line). Therefore, for $C_s \geq 1.53335$ we find a smaller fraction of detected entangled states. In particular, we found that only 2.057% of entangled states were detected.

Finally, we show in figure (3.6) an histogram for the value of the angle $\theta = \pi/20$. In this case, the bound is given by $C_s = 1.89188$ and we found that only 0.288% of the entangled states have been detected.

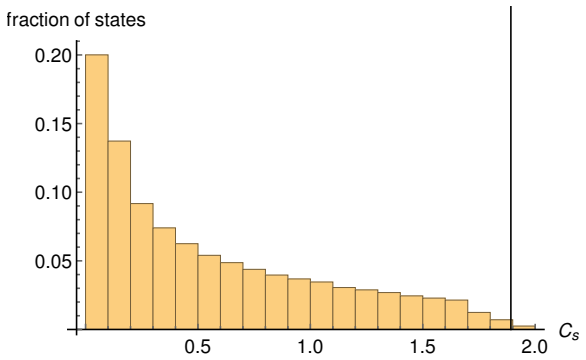

Figure 3.6: Fraction of states versus the correlation sum C_s for $\theta = \pi/20$. In this case, the bound reaches the value $C_s = 1.89188$ (black vertical line). Therefore, for $C_s \geq 1.89188$ we find that only 0.288% of entangled states are detected.

Before moving on to the next chapter, and introducing noisy environments in our analysis, let us define a different notation for our density matrix ρ in a generalization of the Bloch notation for a two-qubit system:

$$\rho = \frac{1}{4}\left(\mathbb{I}\otimes\mathbb{I} + r\cdot\sigma\otimes\mathbb{I} + \mathbb{I}\otimes s\cdot\sigma + \sum_{n,m=1}^{3} t_{nm}\sigma_n\otimes\sigma_m\right)$$

which allows us to write our code for the correlation sum in a more compact way, acting directly on the matrix elements.

From this notation it is also clear that all the correlation properties are enclosed in the 3×3 matrix (t_{nm}), but when given an unknown state we would need full tomography to completely characterize it. Thus, as we said in the introduction, we would need to measure 16 observables. With our criterion, on the other hand, the measures needed are considerably fewer than those needed for full tomography.

3.2 Detection of entanglement in a noisy environment

We proceeded to examine what happened to our states if we let them evolve through various quantum channels. We are going to state the following theorem, that we are going to use in the next sections for our applications on quantum channels.

Theorem: All quantum operations ε on a system of Hilbert space dimension d can be generated by an operator-sum representation containing at most d^2 elements, namely:

$$\varepsilon(\rho) = \sum_{k=1}^{M} E_k \rho E_k^{\dagger},$$

where $0 \leq M \leq d^2$.

In the next sections we are going to see what happens when we introduce various kinds of noise, in particular considering the depolarization, amplitude damping and dephasing channels.

We are going to present our results in terms of a *detection rate* η, namely the fraction of the number of detected entangled states over the number of total entangled states in function of the angle θ. Moreover, every quantum noise channel depends on a certain parameter p, which represents the probability that the original state is left untouched.

In addition, for every quantum channel considered, we are going to show some results about the robustness to noise of our criterion. The robustness to noise is given by the minimal value of the parameter p that has to be reached to be able to detect the state. In an equivalent way, we could define the parameter $\gamma = 1 - p$ that quantifies the noise that could be "tolerated" to detect the state. This is often a good indicator to compare different entanglement detection criteria.

In particular, we are going to show the percentage of detected entangled states as a function of the parameter p, for various values of the angle θ, to emphasize the role of the noise parameter in the detection of entangled states in a noisy environment. Finally, we will also present, for the sake of completeness, the plots for the threshold value of p, that is the value of p above which the states are detected, as a function of the angle θ.

3.2.1 Depolarization

If we consider a qubit system (for instance, the polarization of a photon), on which, according to a suitable coding procedure, we have encoded binary information that is travelling from a sender to a receiver, then the propagation needs a physical support (like an optical fiber) and this unavoidably can lead to possible perturbations to our qubit, due to the interaction with the environment. For a qubit in a noisy environment, a quite general description of the detrimental effects of the environment is the so-called *depolarizing channel*.

The depolarizing channel represents an important kind of quantum noise. Let us consider a single qubit, and with probability $\gamma = 1 - p$ that qubit is depolarized. That is, it is replaced by the completely mixed state, $\mathbb{I}/2$, while with probability $\gamma - 1 = p$ the qubit is left untouched. Of course $0 \leq p \leq 1$. The state of the quantum system after this noise is therefore given by: [22]

$$\varepsilon(\rho) = p\rho + (1-p)\frac{\mathbb{I}}{2}.$$

In other words, we have that the original state ρ is sent to a linear combination of

itself and the maximally mixed state $\mathbb{I}/2$, which is also referred to as the *depolarized state*.

In our case, we first act with the channel on the global two-qubit state, and hence we considered a 4×4 random maximally entangled vector. Therefore we would have a probability $\gamma = 1 - p$ that our state is depolarized, i.e., it is replaced by completely mixed state $\mathbb{I}_4/4$, and a probability $\gamma - 1 = p$ that it is left untouched. The state of the quantum system after this noise (perturbation) is, therefore:

$$\varepsilon(\rho) = p\rho + (1-p)\frac{\mathbb{I}_4}{4}.$$

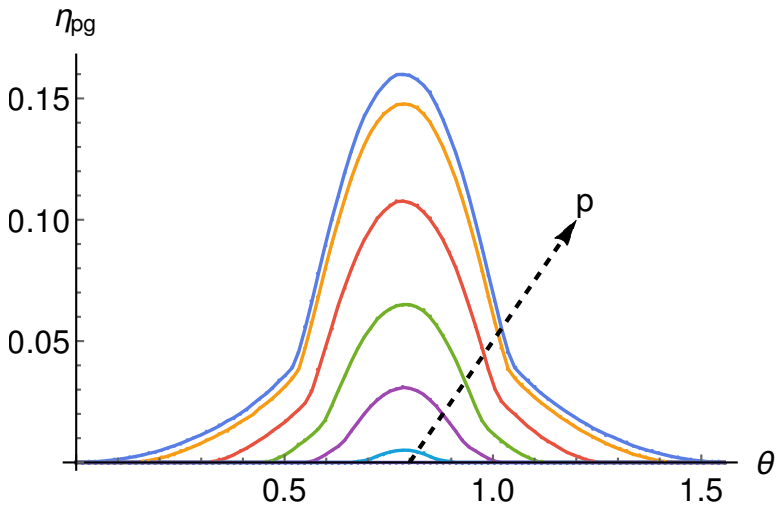

Figure 3.7: Detection rate for the global depolarization channel, η_{pg}, plotted as a function of the angle θ. The various curves represent different values of the parameter p, from minimum to maximum, as indicated by the arrow. The maximum for η_{pg} is obtained for $p = 1$, that is, when the state is left untouched, while the detection rate is null for $p = 0.75$.

In figure (3.7) we show a plot of the detection rate for the case of the global depolarization channel, η_{pg}, as a function of the angle θ for various values of the parameter p. In particular, for the maximum value of the parameter, that is $p = 1$, the state is left untouched, while for decreasing p the detection rate is always smaller, as clear from the dofferent curves in the plot. In this case of global depolarization channel, we see that for $p = 0.75$ the detection rate η_{pg} is null.

In figure (3.8) we plot the rate of detected entangled states η_{pg} as a function of the noise parameter $\gamma = 1 - p$: we can see that, as γ increases, less states are detected. We present the results for various curves relative to different values of the angle θ. We see, as we could expect, that the maximum rate of detected states is obtained for the case of complementary observables, namely for $\theta = \pi/4$, and decreases as it tends to 0. In figure

(3.9), moreover, the threshold value of the channel parameter p_{th} is plotted, above which the states are detected, as a function of the angle θ.

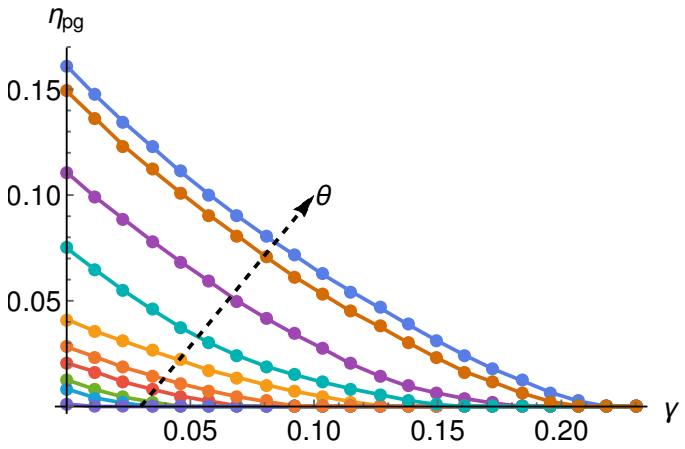

Figure 3.8: Detection rate for the global depolarization channel η_{pg} as a function of the parameter $\gamma = p - 1$. The various curves represent different values of the angle θ, from the minimum to the maximum represented by complementary observables (i.e. $\theta = \pi/4$), as indicated by the arrow. As we can see, the percentage of states decreases as γ grows, thus being maximum for $p = 1$.

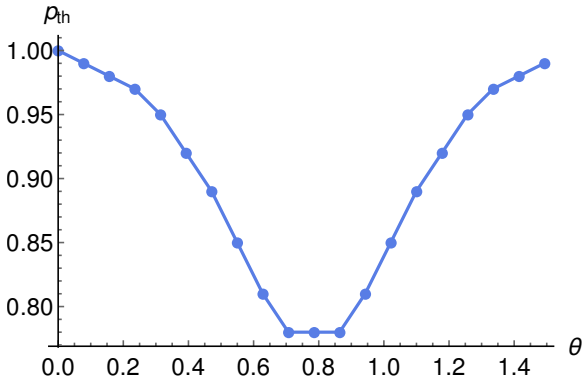

Figure 3.9: Plot of threshold value p_{th} of the parameter, above which the states are detected, as a function of the angle θ

On the other hand, if we want to act on the single qubit, we perform a local quantum operation, using the operator-sum representation. The operation elements are $m_0 = \sqrt{1 - 3/4(1-p)}\,\mathbb{I}$ and $m_k = \sqrt{(1-p)/4}\,\sigma_k$, with $k = 1, 2, 3,$, namely:

$$m_0 = \sqrt{1 - \frac{3}{4}(1-p)} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{3.4}$$

$$m_1 = \sqrt{\frac{1-p}{4}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{3.5}$$

$$m_2 = \sqrt{\frac{1-p}{4}} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{3.6}$$

$$m_3 = \sqrt{\frac{1-p}{4}} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{3.7}$$

and therefore the Kraus operator can be written as:

$$\varepsilon(\rho) = p\rho + \frac{1-p}{3} \sum_k \sigma_k \rho \sigma_k.$$

Anyway, since our system is a two - qubit system, we consider as our operation elements the $E_k = \mathbb{I} \otimes m_k, k = 0, 1, 2, 3$, namely:

$$E_0 = \sqrt{1 - \frac{3}{4}(1-p)} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.8}$$

$$E_1 = \sqrt{\frac{1-p}{4}} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{3.9}$$

$$E_2 = \sqrt{\frac{1-p}{4}} \begin{bmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix} \tag{3.10}$$

$$E_3 = \sqrt{\frac{1-p}{4}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{3.11}$$

and therefore can write the output state (with "noise") in the usual operator-sum representation:

$$\varepsilon(\rho) = \sum_{k=1}^{M} E_k \rho E_k^\dagger.$$

In figure (3.10) we plotted the detection rate for the case of the local depolarization channel, η_{pl}, as a function of the angle θ for various values of the parameter p. We see

here that, again, the maximum value of η_{pl} is obtained for $p = 1$, that is when the state is left untouched. In this case the detection rate is turned off for $p = 0.5$.

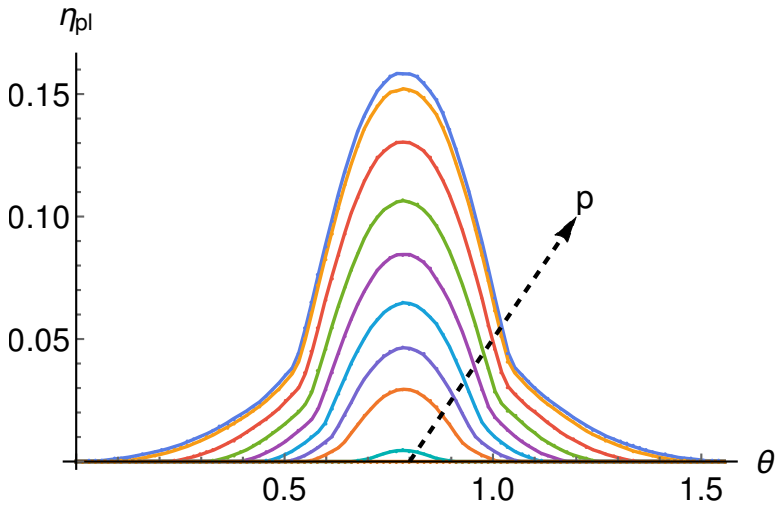

Figure 3.10: Detection rate for the local depolarization channel, η_{pl}, plotted as a function of the angle θ. The various curves represent different values of the parameter p, from minimum to maximum, as indicated by the arrow. The maximum for η_{pl} is obtained for $p = 1$, that is, when the state is left untouched, while the detection rate is turned off for $p = 0.5$.

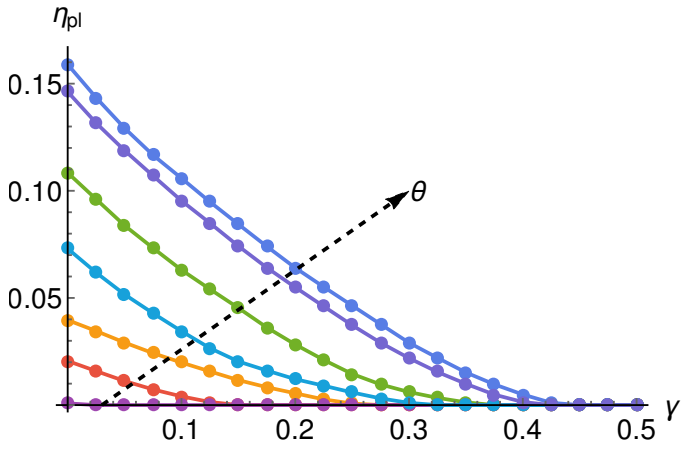

Figure 3.11: Detection rate for the local depolarization channel η_{pl} plotted as a function of the parameter $\gamma = p - 1$. The various curves represent different values of the angle θ, from the minimum to the maximum represented by complementary observables (i.e. $\theta = \pi/4$), as indicated by the arrow. As we can see, the detection rate of states decreases as γ grows, thus being maximum for $p = 1$.

In figure (3.11) we show the detection rate of entangled states for the local depolarization channel η_{pl} as a function of the noise parameter $\gamma = 1 - p$: we see again that, as γ increases, less states are detected. The results for various curves relative to different values of the angle θ are showed. We see that the maximum rate of detected states is obtained for the case of complementary observables, namely for $\theta = \pi/4$, and decreases as the angle tends to 0. In figure (3.12), moreover, the threshold value p_{th} of the parameter p is plotted, above which the states are detected, as a function of the angle θ.

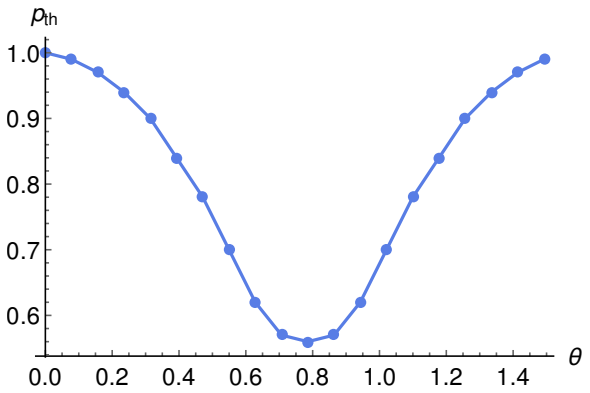

Figure 3.12: Plot of threshold value p_{th} of the parameter p, above which the states are detected, as a function of the angle θ, for the local depolarizing channel.

3.2.2 Amplitude damping

Amplitude and phase damping (the latter of which we are going to analize in the next section) are ideal models of noise that capture many of the most important features of the noise occurring in quantum mechanical systems.

In particular, the description of *energy dissipation*, that is, effects due to loss of energy from a quantum system, follows a general behaviour that can be well characterized by the quantum operation of *amplitude damping*.

Suppose we have a single optical mode containing the quantum state $a|0\rangle + b|1\rangle$, i.e. a superposition of zero or one photons. The scattering of a photon from this mode can be modeled by thinking of inserting a partially silvered mirror, a beamsplitter, in the path of the photon, that allows the photon to couple to another single optical mode (representing the environment), according to the unitary transformation $B = \exp[\theta(a^\dagger b - ab^\dagger)]$, where $a, a^\dagger$ and $b, b^\dagger$ are annihilation and creation operators for photons in the two modes. The output after the beamsplitter, assuming the environment starts out with no photons, is simply $B|0\rangle(a|0\rangle + b|1\rangle) = a|00\rangle + b(\cos\theta|01\rangle + \sin\theta|10\rangle)$.

Again, acting on single qubits, we perform a local quantum operation, where the operation elements are:

$$m_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{p} \end{bmatrix} \tag{3.12}$$

$$m_1 = \begin{bmatrix} 0 & \sqrt{1-p} \\ 1 & 0 \end{bmatrix} \tag{3.13}$$

We point out that $p = \cos^2\theta$ can be thought of as the probability of not losing a photon.

We note that no linear combination can be made of m_0 and m_1 to give an operation element proportional to the identity. The m_1 operation changes a $|0\rangle$ state into a $|1\rangle$ state. On the other hand, m_0 leaves $|0\rangle$ unchanged, but reduces the amplitude of a $|1\rangle$ state; physically, this happens because a quantum of energy was not lost to the environment, and thus the environment now perceives it to be more likely that the system is in the $|0\rangle$ state, rather than the $|1\rangle$ state. [22]

Therefore, with similar reasoning as for the depolarizing channel we will consider as our operation elements the, following: $E_k = \mathbb{I} \otimes m_k, k = 0, 1$, namely:

$$E_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{p} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \sqrt{p} \end{bmatrix} \tag{3.14}$$

$$E_1 = \begin{bmatrix} 0 & \sqrt{1-p} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{1-p} \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.15}$$

and as usual our quantum operation will be represented by the map:

$$\varepsilon(\rho) = \sum_{k=1}^{M} E_k \rho E_k^\dagger.$$

In figure (3.13) we show a plot of the detection rate for the case of the amplitude damping channel, η_{ad}, as a function of the angle θ for various values of the parameter p. The maximum value of η_{ad} is obtained, as always, for $p = 1$, that is when the state is left untouched. For decreasing values of the parameter, the detection rate decreases as well, going to zero for the parameter value $p = 0.6$.

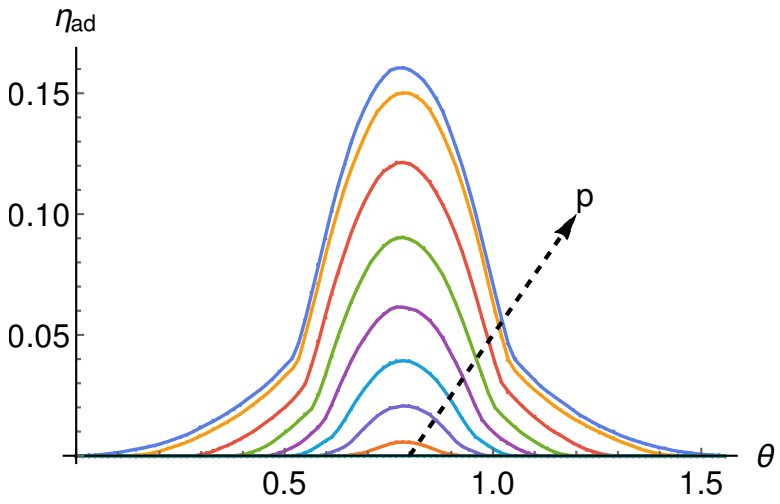

Figure 3.13: Detection rate for the amplitude damping channel, η_{ad}, plotted as a function of the angle θ. The various curves represent different values of the parameter p, from minimum to maximum, as indicated by the arrow. The maximum for η_{ad} is obtained for $p = 1$, that is, when the state is left untouched, while the detection rate is turned off in this case for the parameter value $p = 0.6$.

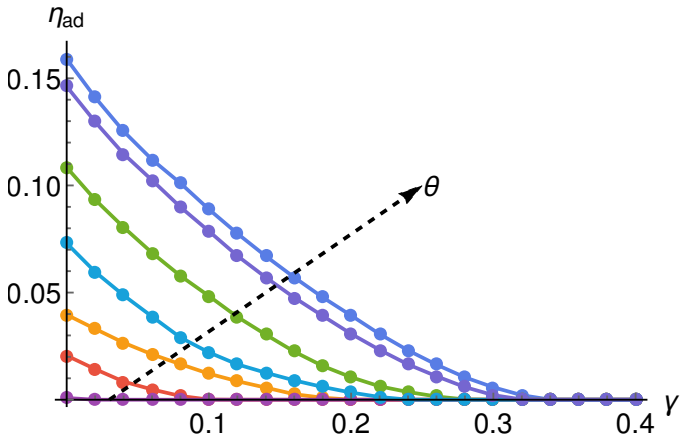

Figure 3.14: Detection rate for the amplitude damping channel η_{ad} plotted as a function of the parameter $\gamma = p - 1$. The various curves represent different values of the angle θ, from the minimum to the maximum represented by complementary observables (i.e. $\theta = \pi/4$), as indicated by the arrow. As we can see, the percentage of states decreases as γ grows, thus being maximum for $p = 1$.

In figure (3.14) we show the detection rate of entangled states for the amplitude damping channel η_{ad}, as a function of the noise parameter $\gamma = 1 - p$ in the case of

amplitude damping channel. We see here that, as γ increases, less states are detected. The results for various curves relative to different values of the angle θ are showed. We see that the maximum rate of detected states is obtained for the case of complementary observables, namely for $\theta = \pi/4$, and decreases as the angle tends to 0. In figure (3.15), moreover, the threshold value of p is plotted, above which the states are detected, as a function of the angle θ.

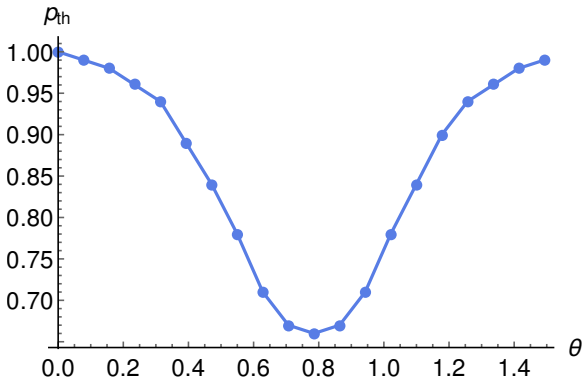

Figure 3.15: Plot of threshold value p_{th} of the parameter p, above which the states are detected, as a function of the angle θ, for the amplitude damping channel.

3.2.3 Dephasing

The phase damping, or *dephasing channel* represents a purely quantum mechanical noise process, which describes the loss of quantum information without loss of energy. Physically it can describe, for example, what happens when a photon scatters randomly as it travels through a waveguide, or how electronic states in an atom are perturbed upon interacting with distant electrical charges. The energy eigenstates of a quantum system do not change as a function of time, but do accumulate a phase which is proportional to the eigenvalue. When a system evolves for an amount of time which is not precisely known, partial information about this quantum phase, i.e. the relative phases between the energy eigenstates, is lost. [22]

We can derive the phase damping quantum operation by considering an interaction between two harmonic oscillators, in a manner similar to how amplitude damping was derived in the last section, and find the operation elements:

$$m_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{p} \end{bmatrix} \tag{3.16}$$

$$m_1 = \begin{bmatrix} 0 & 0 \\ 1 & \sqrt{1-p} \end{bmatrix} \tag{3.17}$$

Therefore we can define $\gamma = 1 - p$, that can be interpreted as the probability that a photon from the system has been scattered, although without any loss of energy.

Following the same reasoning as for the depolarizing channel, since our system is a two- qubit system, we consider as our operation elements the $E_k = \mathbb{I} \otimes m_k, k = 0, 1$, namely:

$$E_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{p} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \sqrt{p} \end{bmatrix} \tag{3.18}$$

$$E_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sqrt{1-p} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{1-p} \end{bmatrix} \tag{3.19}$$

and again we will obtain the operator-sum representation of the channel in the form:

$$\varepsilon(\rho) = \sum_{k=1}^{M} E_k \rho E_k^{\dagger}.$$

In figure (3.16) we show a plot of the detection rate for the case of the local dephasing channel, η_{dl}, as a function of the angle θ for various values of the parameter p. The maximum value of η_{dl} is obtained, as always, for $p = 1$, that is when the state is left untouched. For decreasing values of the parameter, in this case, we see that the detection rate decreases as well, but less dramatically compared to the previous channels. Indeed, we found that $\eta_{dl} = 0$ for the parameter value $p = 0.001$.

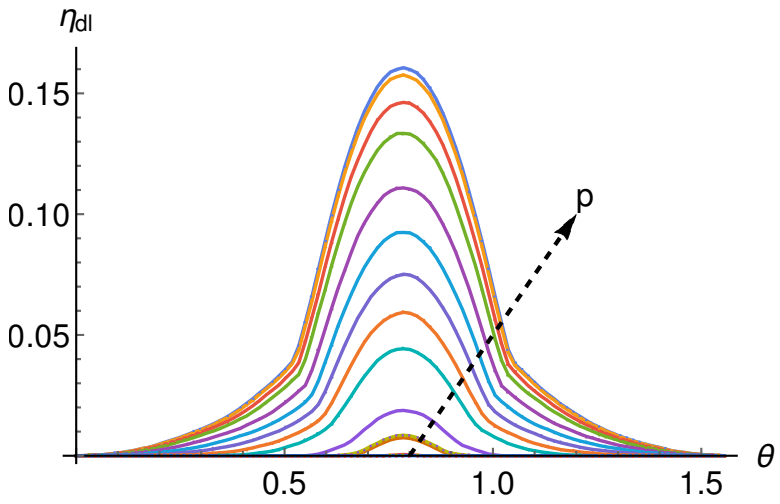

Figure 3.16: Detection rate for the local dephasing channel, η_{dl}, plotted as a function of the angle θ. The various curves represent different values of the parameter p, from minimum to maximum, as indicated by the arrow. The maximum for η_{dl} is obtained for $p = 1$, that is, when the state is left untouched, while in this case the detection rate is turned off for the minimum parameter value $p = 0.001$.

Figure (3.17) shows the detection rate for the local dephasing channel, η_{dl}, of detected entangled states, as a function of the noise parameter $\gamma = 1 - p$ in the case of the local dephasing channel. We see here that, as γ increases, less states are detected. The results for various curves relative to different values of the angle θ are showed. We see that the maximum rate of detected states is obtained for the case of complementary observables, namely for $\theta = \pi/4$, and decreases as the angle tends to 0. In figure (3.18), moreover, the threshold value p_{th} of p is plotted, above which the states are detected, as a function of the angle θ.

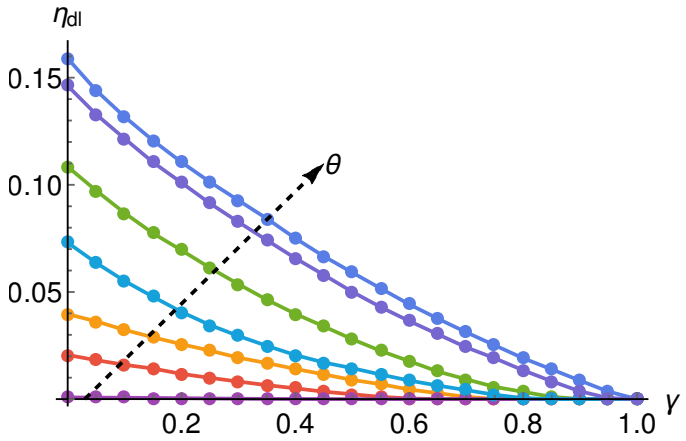

Figure 3.17: Detection rate for the local dephasing channel η_{dl} plotted as a function of the parameter $\gamma = p - 1$. The various curves represent the different values of the angle θ, from its minimum to maximum values. The maximum is naturally represented by the case of complementary observables (i.e. $\theta = \pi/4$), as indicated by the arrow. As we can see, the percentage of states decreases as γ grows, thus being maximum for $p = 1$.

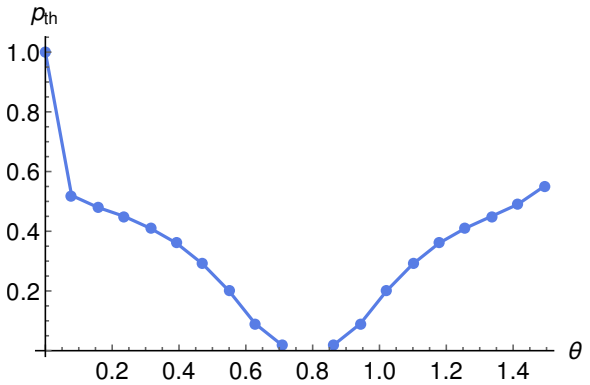

Figure 3.18: Plot of threshold value p_{th} of p above which the states are detected, as a function of the angle θ, for the case of the local dephasing channel.

On the other hand, if we want to consider the global two-qubit system, we consider the quantum operation that, with probability $\gamma = 1 - p$ replaces the ρ with the diagonalized original state, that is mimics the decoherence process. In a general two qubit density matrix, the diagonal real elements represent the probabilities of finding the system in a state of the choosen basis, while the off-diagonal elements, the so-called *quantum coherences*, have no classical equivalent. The dephasing channel induces a decay of those elements, thus representing the decoherence process.

In figure (3.19) we show a plot of the detection rate for the case of the global dephasing channel, η_{dg}, as always as a function of the angle θ for various values of the parameter p. The maximum value of η_{dg} is again obtained for $p = 1$, that is when the state is left untouched. Also in this global case the detection rate decreases with the parameter p, but not quickly. Indeed we found that $\eta_{dg} = 0$ for the parameter value $p = 0.01$.

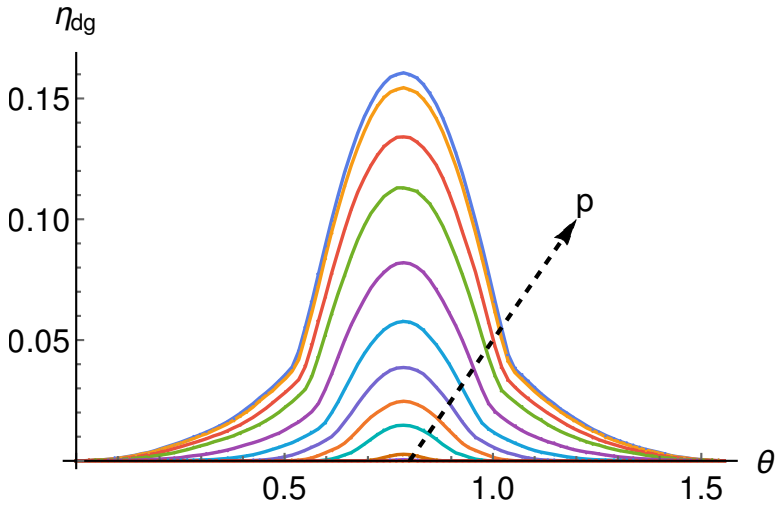

Figure 3.19: Detection rate for the local dephasing channel, η_{dg}, plotted as a function of the angle θ. The various curves represent different values of the parameter p, from minimum to maximum, as indicated by the arrow. The maximum for η_{dg} is obtained for $p = 1$, that is, when the state is left untouched, while in this case the detection rate is turned off for the minimum parameter value $p = 0.01$.

In figure (3.20) we show the detection rate for the global dephasing channel, η_{dg}, of detected entangled states, as a function of the noise parameter $\gamma = 1 - p$ in the case of the global dephasing channel. We see here that, as γ increases, less states are detected. The results for various curves relative to different values of the angle θ are showed. We see that the maximum rate of detected states is obtained for the case of complementary observables, namely for $\theta = \pi/4$, and decreases as the angle tends to 0.

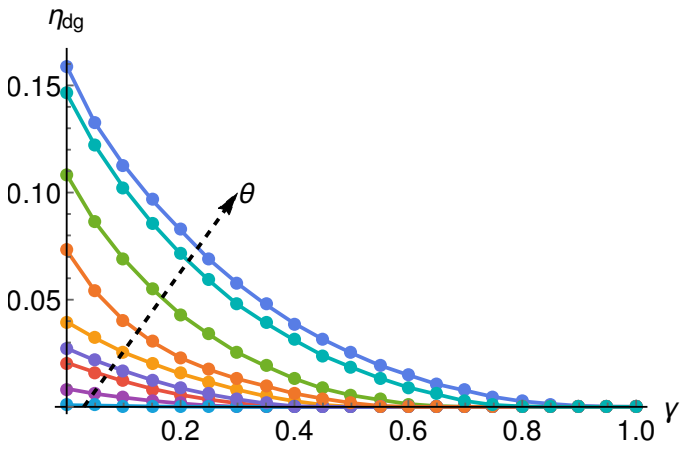

Figure 3.20: Detection rate for the global dephasing channel η_{dg} as a function of the parameter $\gamma = p - 1$. The various curves represent the different values of the angle θ, from its minimum to maximum values. The maximum is naturally represented by the case of complementary observables (i.e. $\theta = \pi/4$), as indicated by the arrow. As we can see, the percentage of states decreases as γ grows, thus being maximum for $p = 1$.

In figure (3.21), finally, the threshold value p_{th} of p is plotted, above which the states are detected, as a function of the angle θ.

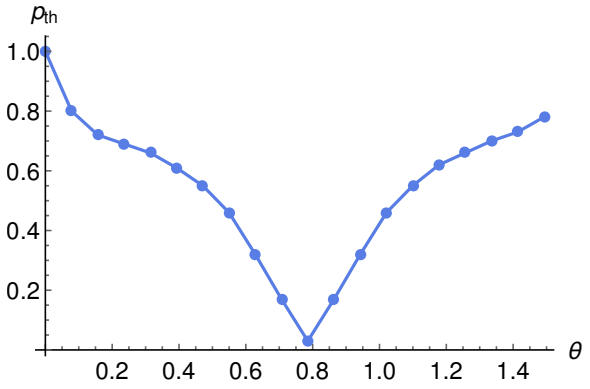

Figure 3.21: Plot of threshold value p_{th} of p above which the states are detected, as a function of the angle θ, for the global dephasing channel.

The numerical simulations showed in this chapter have been performed, as we said, on a sample of 10^5 randomly generated states. However, in the functions we plotted, two different types of statistical fluctuations could appear, due both to the number of states generated and to the fact that the input state we consider is generated randomly. Indeed, every time we generate an input random state, it is in fact a different state to that generated previously. We saw anyway that genereting a different random state for

every point or using the same state for all the calculation, leads to little difference in the final result. Now, for a bigger number of states, we would have obviously obtained a more precise statistics, but we believe that, for the scope of this work, our results were sufficiently accurate, as the curves appear rather "smooth".

3.3 Check of invariance

We performed the numerical calculations to this point considering measures on the basis of eigenstates of σ_z, the so-called *computational basis*, on the observable $A \otimes B$, that is, as defined in section 1.1, $\{|0\rangle, |1\rangle\}$ on each qubit.

Therefore, a general state can be written as

$$\alpha |0\rangle + \beta |1\rangle \tag{3.20}$$

and a measure will yield the result $|0\rangle$ with probability $|\alpha|^2$ or the result $|1\rangle$ with probability $|\beta|^2$.

Now we notice that we could as well consider the bases of the eigenstates of σ_y and σ_x. It would seem logic to think that the results, namely the percentage (i.e. the detection rate η) of entangled states revealed by our method should remain unvaried.

Let us start considering the basis for σ_x, given, on each qubit, by: $\{|+\rangle_x, |-\rangle_x\}$, where:

$$\begin{cases} |+\rangle_x & \equiv \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \\ |-\rangle_x & \equiv \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \end{cases}$$

Therefore the general state (3.20) can be written in this new basis as follows:

$$\begin{aligned} |\psi\rangle &= \alpha |0\rangle + \beta |1\rangle \\ &= \alpha \frac{|+\rangle_x + |-\rangle_x}{\sqrt{2}} + \beta \frac{|+\rangle_x - |-\rangle_x}{\sqrt{2}} \\ &= \frac{\alpha + \beta}{\sqrt{2}} |+\rangle_x + \frac{\alpha - \beta}{\sqrt{2}} |-\rangle_x \end{aligned}$$

Then, a measure on the state $|\psi\rangle$ will yield the result $|+\rangle_x$ with probability $\frac{|\alpha+\beta|^2}{2}$ or the result $|-\rangle_x$ with probability $\frac{|\alpha-\beta|^2}{2}$

For σ_y we have:

$$\begin{cases} |+\rangle_y & \equiv \frac{1}{\sqrt{2}} (|0\rangle + i |1\rangle) \\ |-\rangle_y & \equiv \frac{1}{\sqrt{2}} (|0\rangle - i |1\rangle) \end{cases}$$

Similarly, we can write the state (3.20) as:

$$
\begin{aligned}
|\psi\rangle &= \alpha\,|0\rangle + \beta\,|1\rangle \\
&= \alpha\frac{|+\rangle_y + |-\rangle_y}{\sqrt{2}} - i\beta\frac{|+\rangle_y - |-\rangle_y}{\sqrt{2}} \\
&= \frac{\alpha - i\beta}{\sqrt{2}}\,|+\rangle_y + \frac{\alpha + i\beta}{\sqrt{2}}\,|-\rangle_y\,.
\end{aligned}
$$

It is known that for any finite dimensional space there exist pairs of orthonormal bases that are mutually unbiased. In particular, a set of n orthonormal bases $\{X_j\}$ is said to be a set of n MUBs if each basis X_j is mutually unbiased to every other basis X_k, with $k \neq j$, in the set.

For a qubit, the eigenvectors of the Pauli operators

$$
\begin{aligned}
\sigma_x &= |0\rangle\langle 1| + |1\rangle\langle 0| \\
\sigma_y &= -i\,|0\rangle\langle 1| + i\,|1\rangle\langle 0| \\
\sigma_z &= |0\rangle\langle 0| - |1\rangle\langle 1|
\end{aligned}
$$

form a set of three MUBs.

The spin operators in the three coordinate directions, i.e. the Pauli operators σ_i, for $i = 1, 2, 3$ have the common eigenvalues $\lambda_\pm = \pm 1$ and their eigenvectors are, respectively:

$$
\begin{array}{cc}
\psi_{x_+} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} & \psi_{x_-} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \\
\psi_{y_+} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix} & \psi_{y_-} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix} \\
\psi_{z_+} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \psi_{z_-} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\end{array}
$$

Now, performing the same kinds of numerical simulations as those of section 3.2, we found that, starting by considering measures on the basis of eigenstates of both σ_x and σ_y on the observable $A \otimes B$, and then proceeding in considering non-complementary observables, the results (in terms of the detection rate η and the robustness measure γ), do not vary.

Therefore, we can safely say that our criterion is not dependent on the basis we choose for the observables, since, as we pointed out, the eigenvectors of the Pauli matrices form a set of MUBs, and its performances remain invariant under basis choice.

Chapter 4

Conclusions

We introduced a criterion for the characterization of what is considered the "quintessential phenomenon" of quantum theory, that is entanglement. Indeed, as we extensively discussed, one of the most interesting and stimulating open problems in the research in the quantum information theory field is that of the detection and characterization of entanglement.

Taking on the direction given by Maccone et al. in [1], who proposed a criterion for the detection of entanglement in a two-qubit system, based on the classical correlations existing between the measurement outcomes of the complementary observables of the system, we concentrated our study on the measure of classical correlations given by mutual information and asked ourselves what would happen if we relaxed the condition of complementarity for the observables.

The complementary case is obviously the optimal condition in this context, but through our analysis our criterion turned out to be greatly robust, since the detection rate of entangled states do not decrease dramatically when moving away from the complementarity condition, therefore representing a very useful resource in experimental context, where the laboratory conditions can deviate from the ideal ones.

As Maccone et al. point out in their article, if different entanglement detection methods are compared, it turns out that the mutual information does not have optimal performances in detecting entanglement, compared to the others, but it is interesting to note that most of the entangled states it succeeds to detect are distinct from the ones detected by other methods, like entanglement witnesses.

Moreover, the robustness of this criterion is also clear from the fact that, although it could be optimized through the choice of observables in order to maximize the sum of mutual information, the systems under analysis are guaranteed to be entangled if the conditions given in terms of the bound on the correlations sum are satisfied for any couple of complementary observables [39].

Lastly, as we already mentioned before, this criterion can be characterized with fewer and simpler measurements compared to those required for full tomography and for the

majority of commonly used entanglement witnesses.

As we pointed out in section 2.4, it has been demonstrated that, at least in the case of qubits, the bound we proposed could be made stronger by adding correlations for a third complementary observable, resulting in a significative improvement in the efficiency of the present entanglement detection method.

Up to date, many researches go further in this direction, in order to introduce new misures for the detection of entanglement based on classical correlations for multipartite quantum systems, as in [24], or for high dimensional systems, as in [39].

Moreover, another possible extensions of this criterion would be to relax the assumption of complementary observables to generalized observables, i.e. POVMs. This might represent an interesting resource, since POVMs can be used to model realistic measurement setups. Indeed, entropy uncertainty relations for POVMs do exist, however an optimized relation for qubits analogous to those discussed in this work is still lacking.

Bibliography

[1] L. Maccone, D. Bruß, and C. Macchiavello, Physical Review Letters **114**, 130401 (2015), arXiv:1408.6851.

[2] R. Feynman, *The Character of Physical Law,* Messenger lectures on the evolution of civilization (Cornell University, 1964).

[3] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Reviews of Modern Physics **81**, 865 (2009), quant-ph/0702225.

[4] O. Gühne and G. Tóth, Physics Reports **474**, 1 (2009), arXiv:0811.2803.

[5] S. Wu, Z. Ma, Z. Chen, and S. Yu, Scientific Reports **4**, 4036 (2014), arXiv:1301.6838.

[6] O. Gühne *et al.*, Phys. Rev. A **66**, 062305 (2002).

[7] M. G. A. Paris, European Physical Journal Special Topics **203**, 61 (2012), arXiv:1110.6815.

[8] John Preskill, lecture notes on quantum computation, `http://www.theory.caltech.edu/people/preskill/ph219/`, California Institute of Technology.

[9] A. Einstein, B. Podolsky, and N. Rosen, Physical Review **47**, 777 (1935).

[10] E. Schrödinger, Die Naturwissenschaften **23**, 807 (1935).

[11] J. S. Bell and I. b. A. Aspect, *Speakable and Unspeakable in Quantum Mechanics* (, 2004).

[12] J. S. Bell, Physics **1**, 195 (1964).

[13] N. J. Cerf and C. Adami, Phys. Rev. Lett. **79**, 5194 (1997).

[14] S. Lloyd, Phys. Rev. A **55**, 1613 (1997).

[15] B. Schumacher and M. A. Nielsen, Phys. Rev. A **54**, 2629 (1996).

[16] A. K. Ekert, Physical Review Letters **67**, 661 (1991).

[17] C. H. Bennett *et al.*, Phys. Rev. Lett. **70**, 1895 (1993).

[18] R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).

[19] R. F. Werner, Physical Review A **40**, 4277 (1989).

[20] A. Peres, Physical Review Letters **77**, 1413 (1996), quant-ph/9604005.

[21] B. M. Terhal, Physics Letters A **271**, 319 (2000), quant-ph/9911057.

[22] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (, 2010).

[23] C. Eltschka and J. Siewert, Journal of Physics A Mathematical General **47**, 424005 (2014), arXiv:1402.6710.

[24] D. Sauerwein, C. Macchiavello, L. Maccone, and B. Kraus, ArXiv e-prints (2017), arXiv:1701.07412.

[25] M. A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).

[26] J. I. de Vicente, C. Spee, and B. Kraus, Phys. Rev. Lett. **111**, 110502 (2013).

[27] M. Horodecki, P. Horodecki, and R. Horodecki, Physics Letters A **223**, 1 (1996), quant-ph/9605038.

[28] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, Physics Letters A **58**, 883 (1998), quant-ph/9804024.

[29] G. Vidal and R. F. Werner, Phys. Rev. A **65**, 032314 (2002).

[30] O. Rudolph, Journal of Physics A Mathematical General **36**, 5825 (2003), quant-ph/0202121.

[31] O. Rudolph, Phys. Rev. A **67**, 032312. (2002).

[32] P. Horodecki, Physics Letters A **232**, 333 (1997), quant-ph/9703004.

[33] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999).

[34] E. Størmer, Acta Math. **110**, 233 (1963).

[35] B. M. Terhal, eprint arXiv:quant-ph/9810091 (1998), quant-ph/9810091.

[36] M. A. Nielsen and J. Kempe, Phys. Rev. Lett. **86**, 5184 (2001).

[37] D. Bruß *et al.*, Journal of Modern Optics **49**, 1399 (2002), quant-ph/0110081.

[38] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki, Phys. Rev. A **62**, 052310 (2000).

[39] Z. Huang *et al.*, Scientific Reports **6**, 27637 (2016), arXiv:1604.05824.

[40] H. Maassen and J. B. M. Uffink, Physical Review Letters **60**, 1103 (1988).

[41] D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983).

[42] K. Kraus, Phys. Rev. D **35**, 3070 (1987).

[43] H. P. Robertson, Physical Review **34**, 163 (1929).

[44] J. Sánches-Ruiz, Physics Letters A **244**, 189 (1998).

[45] C. E. Shannon and W. Weaver, *The mathematical theory of communication* (, 1949).

[46] A. J. M. Garrett and S. F. Gull, Physics Letters A **151**, 453 (1990).

[47] G. Ghirardi, L. Marinatto, and R. Romano, Physics Letters A **317**, 32 (2003), quant-ph/0310120.

[48] T. Durt, B.-G. Englert, I. Bengtsson, and K. Zyczkowski, ArXiv e-prints (2010), arXiv:1004.3348.

[49] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, ArXiv e-prints (2015), arXiv:1511.04857.

[50] J. Sánchez, Physics Letters A **173**, 233 (1993).

[51] M.-M. Du, D. Wang, and L. Ye, Scientific Reports **7**, 40934 (2017).

[52] M. Jiang, S. Luo, and S. Fu, Phys. Rev. A **87**, 022310 (2013).

[53] J.A. Miszczak, Z. Puchała, P. Gawron, QI: quantum information package for Mathematica, `http://zksi.iitis.pl/wiki/projects:mathematica-qi`, (2010).

[54] F. Mezzadri, ArXiv Mathematical Physics e-prints (2006), math-ph/0609050.