

Lectures on Quantum Information

Chapter 1: The separability versus entanglement problem

Aditi Sen(De)¹, Ujjwal Sen¹, Maciej Lewenstein^{1,*}, and Anna Sanpera^{2,*}

¹*ICFO-Institut de Ciències Fotòniques, E-08034 Barcelona, Spain*

²*Grup de Física Teòrica, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain*

Contents

I. Introduction	1
II. Bipartite pure states: Schmidt decomposition	2
III. Bipartite mixed states: Separable and entangled states	2
IV. Operational entanglement criteria	3
A. Partial Transposition	3
B. Majorization	4
V. Non-operational entanglement criteria	4
1. Technical Preface	5
A. Entanglement Witnesses	5
1. Entanglement Witness from the Hahn-Banach theorem	5
2. Examples	6
B. Positive maps	7
1. Introduction and definitions	7
2. Positive maps and entangled states	8
3. Jamiołkowski Isomorphism	9
VI. Bell inequalities	10
A. Detection of entanglement by Bell inequality	11
VII. Classification of bipartite states with respect to quantum dense coding	12
A. The Holevo bound	13
B. Capacity of quantum dense coding	13
VIII. Further reading: Multipartite states	15
IX. Problems	15
References	15

I. INTRODUCTION

Quantum theory, formalised in the first few decades of the 20th century, contains elements that are radically different from the classical description of Nature. An important aspect in these fundamental differences is the existence of quantum correlations in the quantum formalism. In the classical description of Nature, if a system is formed by different subsystems, complete knowledge of the whole system implies that the sum of the information of the subsystems makes up the complete information for the whole system. This is no longer true in the quantum formalism. In the quantum world, there exists states of composite systems for which we might have the complete information, while our knowledge about the subsystems might be completely random. One may reach some paradoxical conclusions if one applies a classical description to states which have characteristic quantum signatures.

During the last decade, it was realized that these fundamentally nonclassical states, also denoted as “entangled states”, can provide us with something else than just paradoxes. They may be *used* to perform tasks that cannot be achieved with classical states. As benchmarks of this turning point in our view of such nonclassical states, one might mention the spectacular discoveries

of (entanglement-based) quantum cryptography (1991) [1], quantum dense coding (1992) [2], and quantum teleportation (1993) [3].

In this chapter, we will focus on bipartite composite systems. We will define formally what entangled states are, present some important criteria to discriminate entangled states from separable ones, and show how they can be classified according to their capability to perform some precisely defined tasks. Our knowledge in the subject of entanglement is still far from complete, although significant progress has been made in the recent years and very active research is currently underway.

II. BIPARTITE PURE STATES: SCHMIDT DECOMPOSITION

In this chapter, we will primarily consider bipartite systems, which are traditionally supposed to be in possession of Alice (A) and Bob (B), who can be located in distant regions. Let Alice's physical system be described by the Hilbert space \mathcal{H}_A and that of Bob by \mathcal{H}_B . Then the joint physical system of Alice and Bob is described by the tensor product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$.

Def. 1 *Product and entangled pure states:*

A pure state, i.e. a projector $|\psi_{AB}\rangle\langle\psi_{AB}|$ on a vector $|\psi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, is a product state if the states of local subsystems are also pure states, that is, if $|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$. However, there are states that cannot be written in this form. These states are called entangled states.

An example of entangled state is the well-known singlet state $(|01\rangle - |10\rangle)/\sqrt{2}$, where $|0\rangle$ and $|1\rangle$ are two orthonormal states. Operationally, product states correspond to those states, that can be locally prepared by Alice and Bob at two separate locations. Entangled states can, however, be prepared only after the particles of Alice and Bob have interacted either directly or by means of an ancillary system. A very useful representation, only valid for pure bipartite states, is the, so-called, Schmidt representation:

Theorem 1 *Schmidt decomposition:*

Every $|\psi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ can be represented in an appropriately chosen basis as

$$|\psi_{AB}\rangle = \sum_{i=1}^M a_i |e_i\rangle \otimes |f_i\rangle, \quad (1)$$

where $|e_i\rangle$ ($|f_i\rangle$) form a part of an orthonormal basis in \mathcal{H}_A (\mathcal{H}_B), $a_i > 0$, and $\sum_{i=1}^M a_i^2 = 1$, where $M \leq \dim \mathcal{H}_A, \dim \mathcal{H}_B$. The positive numbers a_i are known as the Schmidt coefficients of $|\psi_{AB}\rangle$. Note that product pure states correspond to those states, whose Schmidt decomposition has one and only one Schmidt coefficient. If the decomposition has more than one Schmidt coefficient, the state is entangled. Notice that the squares of the Schmidt coefficients of a pure bipartite state $|\psi_{AB}\rangle$ are the eigenvalues of either of the reduced density matrices $\rho_A (= \text{tr}_B \rho_{AB})$ and $\rho_B (= \text{tr}_A \rho_{AB})$ of $|\psi_{AB}\rangle$.

III. BIPARTITE MIXED STATES: SEPARABLE AND ENTANGLED STATES

As discussed in the last section, the question whether a given pure bipartite state is separable or entangled is straightforward. One has just to check if the reduced density matrices are pure. This condition is equivalent to the fact that a bipartite pure state has a single Schmidt coefficient. The determination of separability for mixed states is much harder, and currently lacks a complete answer, even in composite systems of dimension as low as $\mathcal{C}^2 \otimes \mathcal{C}^4$.

To reach a formal definition of separable and entangled states, consider the following preparation procedure of a bipartite state between Alice and Bob. Suppose that Alice prepares her physical system in the state $|e_i\rangle$ and Bob prepares his physical system in the state $|f_i\rangle$. Then, the combined state of their joint physical system is given by:

$$|e_i\rangle\langle e_i| \otimes |f_i\rangle\langle f_i|. \quad (2)$$

We now assume that they can communicate over a classical channel (a phone line, for example). Then, whenever Alice prepares the state $|e_i\rangle$ ($i = 1, 2, \dots, K$), which she does with probability p_i , she communicates that to Bob, and correspondingly Bob prepares his system in the state $|f_i\rangle$ ($i = 1, 2, \dots, K$). Of course, $\sum_i p_i = 1$. The state that they prepare is then

$$\rho_{AB} = \sum_{i=1}^K p_i |e_i\rangle\langle e_i| \otimes |f_i\rangle\langle f_i|. \quad (3)$$

The important point to note here is that the state displayed in Eq. (3) is the most general state that Alice and Bob will be able to prepare by local quantum operations and classical communication (LOCC) [4].

Def. 2 *Separable and entangled mixed states:*

A mixed state ρ_{AB} is separable if and only if it can be represented as a convex combination of the product of projectors on local states as stated in Eq. (3). Otherwise, the mixed state is said to be entangled.

Entangled states, therefore, cannot be prepared locally by two parties even after communicating over a classical channel. To prepare such states, the physical systems must be brought together to interact[65]. Mathematically, a nonlocal unitary operator[66] must *necessarily* act on the physical system described by $\mathcal{H}_A \otimes \mathcal{H}_B$, to produce an entangled state from an initial separable state.

The question whether a given bipartite state is separable or not turns out to be quite complicated. Among the difficulties, we notice that for an arbitrary state ρ_{AB} , there is no stringent bound on the value of K in Eq. (3), which is only limited by the Caratheodory theorem to be $K \leq (\dim \mathcal{H})^2$ with $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ (see [6]). Although the general answer to the separability problem still eludes us, there has been significant progress in recent years, and we will review some such directions in the following sections.

IV. OPERATIONAL ENTANGLEMENT CRITERIA

In this section, we will introduce some operational entanglement criteria. In particular, we will discuss the partial transposition criterion [7, 8], and the majorization criterion [9]. There exist several other criteria (see e.g. Refs. [10, 11, 12]), which will not be discussed here. However note that, up to now, a necessary and sufficient criterion for detecting entanglement of an arbitrary given mixed state is still lacking.

A. Partial Transposition

Def. 3 Let ρ_{AB} be a bipartite density matrix, and let us express it as

$$\rho_{AB} = \sum_{i,j=1}^{N_A} \sum_{\mu,\nu=1}^{N_B} a_{ij}^{\mu\nu} (|i\rangle\langle j|)_A \otimes (|\mu\rangle\langle\nu|)_B, \quad (4)$$

where $\{|i\rangle\}$ ($i = 1, 2, \dots, N_A$; $N_A \leq \dim \mathcal{H}_A$) ($\{|\mu\rangle\}$ ($\mu = 1, 2, \dots, N_B$; $N_B \leq \dim \mathcal{H}_B$)) is a set of orthonormal vectors in \mathcal{H}_A (\mathcal{H}_B). The partial transposition, $\rho_{AB}^{T_A}$, of ρ_{AB} with respect to subsystem A, is defined as

$$\rho_{AB}^{T_A} = \sum_{i,j=1}^{N_A} \sum_{\mu,\nu=1}^{N_B} a_{ij}^{\mu\nu} (|j\rangle\langle i|)_A \otimes (|\mu\rangle\langle\nu|)_B. \quad (5)$$

A similar definition exists for the partial transposition of ρ_{AB} with respect to Bob's subsystem. Notice that $\rho_{AB}^{T_B} = (\rho_{AB}^{T_A})^T$. Although the partial transposition depends upon the choice of the basis in which ρ_{AB} is written, its eigenvalues are basis independent. We say that a state has Positive Partial Transposition (PPT), whenever $\rho_{AB}^{T_A} \geq 0$, i.e. the eigenvalues of $\rho_{AB}^{T_A}$ are non-negative. Otherwise, the state is said to be Non-positive under Partial Transposition (NPT). Note here that transposition is equivalent to time reversal.

Theorem 2 [7]

If a state ρ_{AB} is separable, then $\rho_{AB}^{T_A} \geq 0$ and $\rho_{AB}^{T_B} = (\rho_{AB}^{T_A})^T \geq 0$.

Proof:

Since ρ_{AB} is separable, it can be written as

$$\rho_{AB} = \sum_{i=1}^K p_i |e_i\rangle\langle e_i| \otimes |f_i\rangle\langle f_i| \geq 0. \quad (6)$$

Now performing the partial transposition w.r.t. A, we have

$$\begin{aligned} \rho_{AB}^{T_A} &= \sum_{i=1}^K p_i (|e_i\rangle\langle e_i|)^{T_A} \otimes |f_i\rangle\langle f_i| \\ &= \sum_{i=1}^K p_i |e_i^*\rangle\langle e_i^*| \otimes |f_i\rangle\langle f_i| \geq 0. \end{aligned} \quad (7)$$

Note that in the second line, we have used the fact that $A^\dagger = (A^*)^T$. \square

The *partial transposition criterion*, for detecting entanglement is simple: Given a bipartite state ρ_{AB} , find the eigenvalues of any of its partial transpositions. A negative eigenvalue immediately implies that the state is entangled. Examples of states for which the partial transposition has negative eigenvalues include the singlet state.

The partial transposition criterion allows to detect in a straightforward manner all entangled states that are NPT states. This is a huge class of states. However, it turns out that there exist PPT states which are not separable, as pointed out in Ref. [13] (see also [14]). Moreover, the set of PPT entangled states is not a set of measure zero [15]. It is, therefore, important to have further independent criteria of entanglement detection which permits to detect entangled PPT states. It is worth mentioning that PPT states which are entangled, form the only known examples of the “bound entangled states” (see Refs. [14, 16] for details). Note also that both separable as well as PPT states form convex sets.

Theorem 2 is a necessary condition of separability in any arbitrary dimension. However, for some special cases, the partial transposition criterion is both, a necessary and sufficient condition for separability:

Theorem 3 [8]

In $\mathcal{C}^2 \otimes \mathcal{C}^2$ or $\mathcal{C}^2 \otimes \mathcal{C}^3$, a state ρ is separable if and only if $\rho^{TA} \geq 0$.

B. Majorization

The partial transposition criterion, although powerful, is not able to detect entanglement in a finite volume of states. It is, therefore, interesting to discuss other independent criteria. The majorization criterion, to be discussed in this subsection, has been recently shown to be *not* more powerful in detecting entanglement. We choose to discuss it here, mainly because it has independent roots. Moreover, it reveals a very interesting thermodynamical property of entanglement.

Before presenting the criterion, we must first give the definition of majorization [17].

Def. 3 *Let $x = (x_1, x_2, \dots, x_d)$, and $y = (y_1, y_2, \dots, y_d)$ be two probability distributions, arranged in decreasing order, i.e. $x_1 \geq x_2 \geq \dots \geq x_d$ and $y_1 \geq y_2 \geq \dots \geq y_d$. Then we define “ x majorized by y ”, denoted as $x \prec y$, as*

$$\sum_{i=1}^l x_i \leq \sum_{i=1}^l y_i, \quad (8)$$

where $l = 1, 2, \dots, d-1$, and equality holds when $l = d$.

Theorem 4 [9]

If a state ρ_{AB} is separable, then

$$\lambda(\rho_{AB}) \prec \lambda(\rho_A), \quad \text{and} \quad \lambda(\rho_{AB}) \prec \lambda(\rho_B), \quad (9)$$

where $\lambda(\rho_{AB})$ is the set of eigenvalues of ρ_{AB} , and $\lambda(\rho_A)$ and $\lambda(\rho_B)$ are the sets of eigenvalues of the corresponding reduced density matrix of the state ρ_{AB} , and where all the sets are arranged in decreasing order.

The majorization criterion: Given a bipartite state, it is entangled if Eq. (9) is violated. However, it was recently shown in Ref. [18], that a state that is not detected by the positive partial transposition criterion, will not be detected by the majorization criterion also. Nevertheless, the criterion has other important implications. We will now discuss one such.

Let us reiterate an interesting fact about the singlet state: The global state is pure, while the local states are completely mixed. In particular, this implies that the von Neumann entropy[67] of the global state is lower than either of the von Neumann entropies of the local states. The von Neumann entropy can however be used to quantify disorder in a quantum state. This implies that there exist bipartite quantum states for which the global disorder can be more than either of the local disorders. This is a nonclassical fact as for two classical random variables, the Shannon entropy[68] of the joint distribution cannot be smaller than that of either. In Ref. [19], it was shown that a similar fact is true for separable states:

Theorem 5

If a state ρ_{AB} is separable,

$$S(\rho_{AB}) \geq S(\rho_A), \quad \text{and} \quad S(\rho_{AB}) \geq S(\rho_B). \quad (10)$$

Although the von Neumann entropy is an important notion for quantifying disorder, the theory of majorization is a more stringent quantifier [17]: For two probability distributions x and y , $x \prec y$ if and only if $x = Dy$, where D is a doubly stochastic matrix[69]. Moreover, $x \prec y$ implies that $H(\{x_i\}) \geq H(\{y_i\})$. Quantum mechanics therefore allows the existence of states for which global disorder is greater than local disorder even in the sense of majorization.

A density matrix that satisfies Eq. (9), automatically satisfies Eq. (10). In this sense, Theorem 4 is a generalization of Theorem 5.

V. NON-OPERATIONAL ENTANGLEMENT CRITERIA

In this section, we will discuss two further entanglement criteria. We will show how the Hahn-Banach theorem can be used to obtain “entanglement witnesses”. We will also introduce the notion of positive maps, and present the entanglement criterion

based on it. Both the criteria are “non-operational”, in the sense that they are not state-independent. Nevertheless, they provide important insight into the structure of the set of entangled states. Moreover, the concept of entanglement witnesses can be used to detect entanglement experimentally, by performing only a few *local* measurements, assuming some prior knowledge of the density matrix [20, 21].

1. Technical Preface

The following lemma and observation will be useful for later purposes.

Lemma 1

$$\text{tr}(\rho_{AB}^{T_A} \sigma_{AB}) = \text{tr}(\rho_{AB} \sigma_{AB}^{T_A}).$$

Observation:

The space of linear operators acting on \mathcal{H} (denoted by $\mathcal{B}(\mathcal{H})$) is itself a Hilbert space, with the (Euclidean) scalar product

$$\langle A|B \rangle = \text{tr}(A^\dagger B) \quad A, B \in \mathcal{B}(\mathcal{H}). \quad (11)$$

This scalar product is equivalent to writing A and B row-wise as vectors, and scalar-multiplying them:

$$\text{tr}(A^\dagger B) = \sum_{ij} A_{ij}^* B_{ij} = \sum_{k=1}^{(\dim \mathcal{H})^2} a_k^* b_k. \quad (12)$$

A. Entanglement Witnesses

1. Entanglement Witness from the Hahn-Banach theorem

Central to the concept of entanglement witnesses, is the Hahn-Banach theorem, which we will present here limited to our situation and without proof (see e.g. [22] for a proof of the more general theorem):

Theorem 6

Let S be a convex compact set in a finite dimensional Banach space. Let ρ be a point in the space with $\rho \notin S$. Then there exists a hyperplane [70] that separates ρ from S .

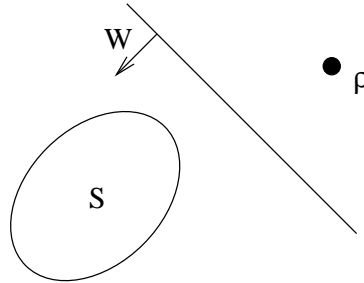


FIG. 1: Schematic picture of the Hahn-Banach theorem. The (unique) unit vector orthonormal to the hyperplane can be used to define *right* and *left* in respect to the hyperplane by using the signum of the scalar product.

The statement of the theorem is illustrated in figure 1. The figure motivates the introduction of a new coordinate system located within the hyperplane (supplemented by an orthogonal vector W which is chosen such that it points away from S). Using this coordinate system, every state ρ can be characterized by its distance from the plane, by projecting ρ onto the chosen orthonormal vector and using the trace as scalar product, i.e. $\text{tr}(W\rho)$. This measure is either positive, zero, or negative. We now suppose that S is the convex compact set of all separable states. According to our choice of basis in figure 1, every separable state has a positive distance while there are some entangled states with a negative distance. More formally this can be phrased as:

Def. 4 A hermitian operator (an observable) W is called an entanglement witness (EW) if and only if

$$\exists \rho \quad \text{such that} \quad \text{tr}(W\rho) < 0, \quad \text{while} \quad \forall \sigma \in S, \quad \text{tr}(W\sigma) \geq 0. \quad (13)$$

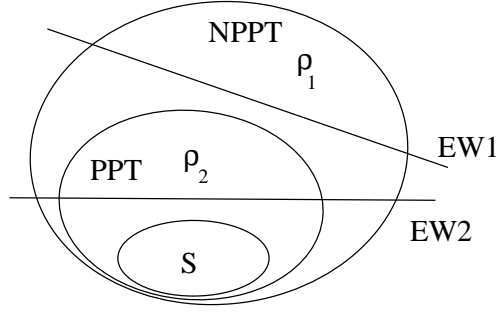


FIG. 2: Schematic view of the Hilbert-space with two states ρ_1 and ρ_2 and two witnesses $EW1$ and $EW2$. $EW1$ is a decomposable EW, and it detects only NPT states like ρ_1 . $EW2$ is an nd-EW, and it detects also some PPT states like ρ_2 . Note that neither witness detects *all* entangled states.

Def. 5 An EW is decomposable if and only if there exists operators P, Q with

$$W = P + Q^{T_A}, \quad P, Q \geq 0. \quad (14)$$

Lemma 2

Decomposable EW cannot detect PPT entangled states.

Proof:

Let δ be a PPT entangled state and W be a decomposable EW. Then

$$\text{tr}(W\delta) = \text{tr}(P\delta) + \text{tr}(Q^{T_A}\delta) = \text{tr}(P\delta) + \text{tr}(Q\delta^{T_A}) \geq 0. \quad (15)$$

Here we used Lemma 1. \square

Def. 6 An EW is called non-decomposable entanglement witness (nd-EW) if and only if there exists at least one PPT entangled state which is detected by the witness.

Using these definitions, we can restate the consequences of the Hahn-Banach theorem in several ways:

Theorem 7 [8, 23, 24, 25]

1. ρ is entangled if and only if, \exists a witness W , such that $\text{tr}(\rho W) < 0$.
2. ρ is a PPT entangled state if and only if \exists a nd-EW, W , such that $\text{tr}(\rho W) < 0$.
3. σ is separable if and only if \forall EW, $\text{tr}(W\sigma) \geq 0$.

From a theoretical point of view, the theorem is quite powerful. However, it does not give any insight of how to construct for a given state ρ , the appropriate witness operator.

2. Examples

For a decomposable witness

$$W' = P + Q^{T_A}, \quad (16)$$

$$\text{tr}(W'\sigma) \geq 0, \quad (17)$$

for all separable states σ .

Proof:

If σ is separable, then it can be written as a convex sum of product vectors. So if Eq. (17) holds for any product vector $|e, f\rangle$, any separable state will also satisfy the same.

$$\begin{aligned} \text{tr}(W'|e, f\rangle\langle e, f|) &= \langle e, f|W'|e, f\rangle \\ &= \underbrace{\langle e, f|P|e, f\rangle}_{\geq 0} + \underbrace{\langle e, f|Q^{T_A}|e, f\rangle}_{\geq 0}, \end{aligned} \quad (18)$$

because

$$\langle e, f | Q^{TA} | e, f \rangle = \text{tr}(Q^{TA} | e, f \rangle \langle e, f |) = \text{tr}(Q | e^*, f \rangle \langle e^*, f |) \geq 0. \quad (19)$$

Here we used Lemma 1, and $P, Q \geq 0$. \square This argumentation shows that $W = Q^{TA}$ is a suitable witness also.

Let us consider the simplest case of $\mathcal{C}^2 \otimes \mathcal{C}^2$. We can use

$$|\phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle), \quad (20)$$

to write the density matrix

$$Q = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}. \quad \text{Then} \quad Q^{TA} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}. \quad (21)$$

One can quickly verify that indeed $W = Q^{TA}$ fulfills the witness requirements. Using

$$|\psi^-\rangle = \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle), \quad (22)$$

we can rewrite the witness:

$$W = Q^{TA} = \frac{1}{2} (I - 2|\psi^-\rangle \langle \psi^-|). \quad (23)$$

This witness now detects $|\psi^-\rangle$:

$$\text{tr}(W |\psi^-\rangle \langle \psi^-|) = -\frac{1}{2}. \quad (24)$$

B. Positive maps

1. Introduction and definitions

So far we have only considered states belonging to a Hilbert space \mathcal{H} , and operators acting on the Hilbert space. However, the space of operators $\mathcal{B}(\mathcal{H})$ has also a Hilbert space structure. We now look at transformations of operators, the so-called maps which can be regarded as *superoperators*. As we will see, this will lead us to an important characterization of entangled and separable states. We start by defining linear maps:

Def. 7 A linear, self-adjoint map ϵ is a transformation

$$\epsilon : \mathcal{B}(\mathcal{H}_B) \rightarrow \mathcal{B}(\mathcal{H}_C), \quad (25)$$

which

- is linear, i.e.

$$\epsilon(\alpha O_1 + \beta O_2) = \alpha \epsilon(O_1) + \beta \epsilon(O_2) \quad \forall O_1, O_2 \in \mathcal{B}(\mathcal{H}_B), \quad (26)$$

where α, β are complex numbers,

- and maps hermitian operators onto hermitian operators, i.e.

$$\epsilon(O^\dagger) = (\epsilon(O))^\dagger \quad \forall O \in \mathcal{B}(\mathcal{H}_B). \quad (27)$$

For brevity, we will only write “linear map”, instead of “linear self adjoint map”. The following definitions help to further characterize linear maps.

Def. 8 A linear map ϵ is called trace preserving if

$$\text{tr}(\epsilon(O)) = \text{tr}(O) \quad \forall O \in \mathcal{B}(\mathcal{H}_B). \quad (28)$$

Def. 9 *Positive map:*

A linear, self-adjoint map ϵ is called positive if

$$\forall \rho \in \mathcal{B}(\mathcal{H}_B), \quad \rho \geq 0 \quad \Rightarrow \quad \epsilon(\rho) \geq 0. \quad (29)$$

Positive maps have, therefore, the property of mapping positive operators onto positive operators. It turns out that by considering maps that are a tensor product of a positive operator acting on subsystem A, and the identity acting on subsystem B, one can learn about the properties of the composite system.

Def. 10 *Completely positive map:*

A positive linear map ϵ is completely positive if for any tensor extension of the form

$$\epsilon' = \mathcal{I}_A \otimes \epsilon, \quad (30)$$

where

$$\epsilon' : \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B) \rightarrow \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_C), \quad (31)$$

ϵ' is positive. Here \mathcal{I}_A is the identity map on $\mathcal{B}(\mathcal{H}_A)$.

Example: Hamiltonian evolution of a quantum system. Let $O \in \mathcal{B}(\mathcal{H}_B)$ and U an unitary matrix and let us define ϵ by

$$\begin{aligned} \epsilon : \mathcal{B}(\mathcal{H}_A) &\rightarrow \mathcal{B}(\mathcal{H}_A) \\ \epsilon(O) &= UOU^\dagger. \end{aligned} \quad (32)$$

As an example for this map, consider the time-evolution of a density matrix. It can be written as $\rho(t) = U(t)\rho(0)U^\dagger(t)$, i.e. in the form given above. Clearly this map is linear, self-adjoint, positive and trace-preserving. It is also completely positive, because for $0 \leq w \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_A)$,

$$(\mathcal{I}_A \otimes \epsilon)w = (I_A \otimes U)w(I_A \otimes U^\dagger) = \tilde{U}w\tilde{U}^\dagger, \quad (33)$$

where \tilde{U} is unitary. But then $\langle \psi | \tilde{U}w\tilde{U}^\dagger | \psi \rangle \geq 0$, if and only if $\langle \psi | w | \psi \rangle \geq 0$ (since positivity is not changed by unitary evolution).

Example: Transposition. An example of a positive but not completely positive map is the transposition T defined as:

$$\begin{aligned} T : \mathcal{B}(\mathcal{H}_B) &\rightarrow \mathcal{B}(\mathcal{H}_B) \\ T(\rho) &= \rho^T. \end{aligned} \quad (34)$$

Of course this map is positive, but it is not completely positive, because

$$(\mathcal{I}_A \otimes T)w = w^{T_B}, \quad (35)$$

and we know that there exist states for which $\rho \geq 0$, but $\rho^{T_B} \not\geq 0$.

Def. 11 *A positive map is called decomposable if and only if it can be written as*

$$\epsilon = \epsilon_1 + \epsilon_2 T \quad (36)$$

where ϵ_1, ϵ_2 are completely positive maps and T is the operation of transposition.

2. Positive maps and entangled states

Partial transposition can be regarded as a particular case of a map that is positive but not completely positive. We have already seen that this particular positive but not completely positive map gives us a way to discriminate entangled states from separable states. The theory of positive maps provides with stronger conditions for separability, as shown in Ref. [8].

Theorem 8

A state $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is separable if and only if for all positive maps

$$\epsilon : \mathcal{B}(\mathcal{H}_B) \rightarrow \mathcal{B}(\mathcal{H}_C), \quad (37)$$

we have

$$(\mathcal{I}_A \otimes \epsilon)\rho \geq 0. \quad (38)$$

Proof:

[\Rightarrow] As ρ is separable, we can write it as

$$\rho = \sum_{k=1}^P p_k |e_k\rangle\langle e_k| \otimes |f_k\rangle\langle f_k|, \quad (39)$$

for some $P > 0$. On this state, $(\mathcal{I}_A \otimes \epsilon)$ acts as

$$(\mathcal{I}_A \otimes \epsilon)\rho = \sum_{k=1}^P p_k |e_k\rangle\langle e_k| \otimes \epsilon(|f_k\rangle\langle f_k|) \geq 0, \quad (40)$$

where the last \geq follows because $|f_k\rangle\langle f_k| \geq 0$, and ϵ is positive.

[\Leftarrow] The proof in this direction is not as easy as the only if direction. We shall prove it at the end of this section.

Theorem 8 can also be recasted into the following form:

Theorem 8 [8]

A state $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is entangled if and only if there exists a positive map $\epsilon : \mathcal{B}(\mathcal{H}_B) \rightarrow \mathcal{B}(\mathcal{H}_C)$, such that

$$(\mathcal{I}_A \otimes \epsilon)\rho \not\geq 0. \quad (41)$$

Note that Eq. (41) can never hold for maps, ϵ , that are completely positive, and for non-positive maps, it may hold even for separable states. Hence, any positive but not completely positive map can be used to detect entanglement.

3. Jamiołkowski Isomorphism

In order to complete the proof of Theorem 8, we introduce first the Jamiołkowski isomorphism [26] between operators and maps. Given an operator $E \in \mathcal{B}(\mathcal{H}_B \otimes \mathcal{H}_C)$, and an orthonormal product basis $|k, l\rangle$, we define a map by

$$\begin{aligned} \epsilon : \mathcal{B}(\mathcal{H}_B) &\rightarrow \mathcal{B}(\mathcal{H}_C) \\ \epsilon(\rho) &= \sum_{k_1, l_1, k_2, l_2} {}_{BC} \langle k_1 l_1 | E | k_2 l_2 \rangle_{BC} \quad |l_1\rangle_{CB} \langle k_1 | \rho | k_2 \rangle_{BC} \langle l_2|, \end{aligned} \quad (42)$$

or in short form,

$$\epsilon(\rho) = \text{tr}_B(E\rho^{TB}). \quad (43)$$

This shows how to construct the map ϵ from a given operator E . To construct an operator from a given map we use the state

$$|\psi^+\rangle = \frac{1}{\sqrt{M}} \sum_{i=1}^M |i\rangle_{B'} |i\rangle_B \quad (44)$$

(where $M = \dim \mathcal{H}_B$) to get

$$M (I_{B'} \otimes \epsilon) (|\psi^+\rangle\langle\psi^+|) = E. \quad (45)$$

This isomorphism between maps and operators results in the following properties:

Theorem 10 [8, 23, 24, 25, 26]

1. $E \geq 0$ if and only if ϵ is a completely positive map.
2. E is an entanglement witness if and only if ϵ is a positive map.
3. E is a decomposable entanglement witness if and only if ϵ is decomposable.
4. E is a non-decomposable entanglement witness if and only if ϵ is non-decomposable and positive.

To indicate further how this equivalence between maps and operators works, we develop here a proof for the “only if” direction of the second statement. Let $E \in \mathcal{B}(\mathcal{H}_B \otimes \mathcal{H}_C)$ be an entanglement witness, then $\langle e, f | E | e, f \rangle \geq 0$. By the Jamiołkowski isomorphism, the corresponding map is defined as $\epsilon(\rho) = \text{tr}_B(E\rho^{T_B})$ where $\rho \in \mathcal{B}(\mathcal{H}_B)$. We have to show that

$${}_C\langle\phi|\epsilon(\rho)|\phi\rangle_C = {}_C\langle\phi|\text{tr}(E\rho^{T_B})|\phi\rangle_C \geq 0 \quad \forall |\phi\rangle_C \in \mathcal{H}_C. \quad (46)$$

Since ρ acts on Bob’s space, using the spectral decomposition of ρ , $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$, leads to

$$\rho^{T_B} = \sum_i \lambda_i |\psi_i^*\rangle\langle\psi_i^*|, \quad (47)$$

where all $\lambda_i \geq 0$. Then

$$\begin{aligned} {}_C\langle\phi|\epsilon(\rho)|\phi\rangle_C &= {}_C\langle\phi|\sum_i \text{tr}_B(E\lambda_i|\psi_i^*\rangle_{BB}\langle\psi_i^*|)|\phi\rangle_C \\ &= \sum_i \lambda_i {}_{BC}\langle\psi_i^*, \phi|E|\psi_i^*, \phi\rangle_{BC} \geq 0. \end{aligned} \quad (48)$$

□

We can now proof the \Leftarrow direction of Theorem 8 or, equivalently, the \Rightarrow direction of Theorem 9. We thus have to show that if ρ_{AB} is entangled, there exists a positive map $\epsilon : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_C)$, such that $(\epsilon \otimes \mathcal{I}_B)\rho$ is not positive definite. If ρ is entangled, then there exists an entanglement witness W_{AB} such that

$$\begin{aligned} \text{tr}(W_{AB}\rho_{AB}) &< 0, \quad \text{and} \\ \text{tr}(W_{AB}\sigma_{AB}) &\geq 0, \end{aligned} \quad (49)$$

for all separable σ_{AB} . W_{AB} is an entanglement witness (which detects ρ_{AB}) if and only if W_{AB}^T (note the complete transposition!) is also an entanglement witness (which detects ρ_{AB}^T). We define a map by

$$\epsilon : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{H}_C), \quad (50)$$

$$\epsilon(\rho) = \text{tr}_A(W_{AC}^T \rho_{AB}^{T_A}), \quad (51)$$

where $\dim \mathcal{H}_C = \dim \mathcal{H}_B = M$. Then

$$(\epsilon \otimes \mathcal{I}_B)(\rho_{AB}) = \text{tr}_A(W_{AC}^T \rho_{AB}^{T_A}) = \text{tr}_A(W_{AC}^{T_C} \rho_{AB}) = \tilde{\rho}_{CB}, \quad (52)$$

where we have used Lemma 1, and that $T = T_A \circ T_C$. To complete the proof, one has to show that $\tilde{\rho}_{CB} \not\geq 0$, which can be done by showing that ${}_{CB}\langle\psi^+|\tilde{\rho}_{CB}|\psi^+\rangle_{CB} < 0$, where $|\psi^+\rangle_{CB} = \frac{1}{\sqrt{M}} \sum_i |ii\rangle_{CB}$, with $\{|i\rangle\}$ being an orthonormal basis. □

VI. BELL INEQUALITIES

The first criterion used to detect entanglement was Bell inequalities, which we briefly discuss in this section. As we shall see, Bell inequalities are essentially a special type of entanglement witness. An additional property of Bell inequalities is that any entangled state detected by them is nonclassical in a particular way: It violates “local realism”.

The assumptions of “locality” and “realism” were already present in the famous argument of Einstein, Podolsky, and Rosen [27], that questions the completeness of quantum mechanics. Bell [28] made these assumptions more precise, and more importantly, showed that the assumptions are actually *testable* in experiments. He derived an inequality that must be satisfied by any physical theory of nature, that is “local” as well as “realistic”, the precise meanings of which will be described below. The inequality is actually a constraint on a linear function of results of certain experiments. He then went on to show that there exist states in quantum theory that violate this inequality in experiments. Modulo some so-called loopholes (see e.g. [29]), these inequalities have been shown to be actually violated in experiments (see e.g. [30] and references therein). In this section, we will first derive a Bell inequality [71] and then show how this inequality is violated by the singlet state.

Consider a two spin-1/2 particle state where the two particles are far apart. Let the particles be called A and B . Let projection valued measurements in the directions a and b be done on A and B respectively. The outcomes of the measurements performed on the particles A and B in the directions a and b , are respectively A_a and B_b . The measurement result A_a (B_b), whose values can be ± 1 , may depend on the direction a (b) and some other uncontrolled parameter λ which may depend on anything, that is, may depend upon system or measuring device or both. Therefore we assume that A_a (B_b) has a definite pre-measurement value $A_a(\lambda)$ ($B_b(\lambda)$). Measurement merely uncovers this value. This is the *assumption of reality*. λ is usually called a hidden

variable and this assumption is also termed as the hidden variable assumption. Moreover, the measurement result at A (B) does not depend on what measurements are performed at B (A). That is, for example $A_a(\lambda)$ does not depend upon b . This is the *assumption of locality*, also called the Einstein's locality assumption. The parameter λ is assumed to have a probability distribution, say $\rho(\lambda)$. Therefore $\rho(\lambda)$ satisfies the following:

$$\int \rho(\lambda) d\lambda = 1, \quad \rho(\lambda) \geq 0. \quad (53)$$

The correlation function of the two spin-1/2 particle state for a measurement in a fixed direction a for particle A and b for particle B , is then given by (provided the hidden variables exist)

$$E(a, b) = \int A_a(\lambda) B_b(\lambda) \rho(\lambda) d\lambda. \quad (54)$$

Here

$$A_a(\lambda) = \pm 1, \quad \text{and} \quad B_b(\lambda) = \pm 1, \quad (55)$$

because the measurement values were assumed to be ± 1 .

Let us now suppose that the observers at the two particles A and B can choose their measurements from two observables a , a' and b , b' respectively, and the corresponding outcomes are A_a , $A_{a'}$ and B_b , $B_{b'}$ respectively. Then

$$\begin{aligned} & E(a, b) + E(a, b') + E(a', b) - E(a', b') \\ &= \int [A_a(\lambda)(B_b(\lambda) + B_{b'}(\lambda)) + A_{a'}(\lambda)(B_b(\lambda) - B_{b'}(\lambda))] \rho(\lambda) d\lambda. \end{aligned} \quad (56)$$

Now $B_b(\lambda) + B_{b'}(\lambda)$ and $B_b(\lambda) - B_{b'}(\lambda)$ can only be ± 2 and 0, or 0 and ± 2 respectively. Consequently,

$$-2 \leq E(a, b) + E(a, b') + E(a', b) - E(a', b') \leq 2. \quad (57)$$

This is the well-known CHSH inequality. Note here that in obtaining the above inequality, we have never used quantum mechanics. We have only assumed Einstein's locality principle and an underlying hidden variable model. Consequently, a Bell inequality is a constraint that any physical theory that is both, local and realistic, has to satisfy. Below, we will show that this inequality can be violated by a quantum state. Hence quantum mechanics is incompatible with an underlying local realistic model.

A. Detection of entanglement by Bell inequality

Let us now show how the singlet state can be detected by a Bell inequality. This additionally will indicate that quantum theory is incompatible with local realism. For the singlet state, the quantum mechanical prediction of the correlation function $E(a, b)$ is given by

$$E(a, b) = \langle \psi^- | \sigma_a \cdot \sigma_b | \psi^- \rangle = -\cos(\theta_{ab}), \quad (58)$$

where $\sigma_a = \vec{\sigma} \cdot \vec{a}$ and similarly for σ_b . $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, where σ_x , σ_y and σ_z are the Pauli spin matrices. And θ_{ab} is the angle between the two measurement directions a and b .

So for the singlet state, one has

$$\begin{aligned} B_{CHSH} &= E(a, b) + E(a, b') + E(a', b) - E(a', b') \\ &= -\cos \theta_{ab} - \cos \theta_{ab'} - \cos \theta_{a'b} + \cos \theta_{a'b'}. \end{aligned} \quad (59)$$

The maximum value of this function is attained for the directions a, b, a', b' on a plane, as given in Fig. 3, and in that case

$$|B_{CHSH}| = 2\sqrt{2}. \quad (60)$$

This clearly violates the inequality in Eq. (57). But Eq. (57) was a constraint for any physical theory which has an underlying local hidden variable model. As the singlet state, a state allowed by the quantum mechanical description of nature, violates the constraint (57), quantum mechanics cannot have an underlying local hidden variable model. In other words, quantum mechanics is not local realistic. This is the statement of the celebrated Bell theorem.

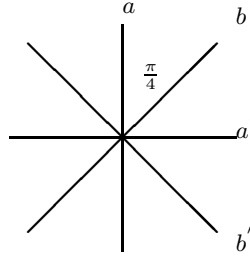


FIG. 3: Schematic diagram showing the direction of a, b, a', b' for obtaining maximal violation of Bell inequality by the singlet state.

Moreover, it is easy to convince oneself that any separable state does have a local realistic description, so that such a state cannot violate a Bell inequality. Consequently, the violation of Bell inequality by the singlet state indicates that the singlet state is an entangled state. Further, the operator (cf. Eqs. (58) and (59))

$$\tilde{B}_{CHSH} = \sigma_a \cdot \sigma_b + \sigma_a \cdot \sigma_{b'} + \sigma_{a'} \cdot \sigma_b - \sigma_{a'} \cdot \sigma_{b'} \quad (61)$$

can, by suitable scaling and change of origin, be considered as an entanglement witness for the singlet state, for a, b, a', b' chosen as in figure 3 (cf. [33]).

VII. CLASSIFICATION OF BIPARTITE STATES WITH RESPECT TO QUANTUM DENSE CODING

Up to now, we have been interested in splitting the set of all bipartite quantum states into separable and entangled states. However, one of the main motivations behind the study of entangled states is that some of them can be used to perform certain tasks, which are not possible if one uses states without entanglement. It is, therefore, important to find out which entangled states are useful for a given task. We discuss here the particular example of quantum dense coding [2].

Suppose that Alice wants to send two bits of classical information to Bob. Then a general result known as the Holevo bound (to be discussed below), shows that Alice must send two qubits (i.e. 2 two-dimensional quantum systems) to Bob, if only a noiseless quantum channel is available. However, if additionally Alice and Bob have previously shared entanglement, then Alice may have to send less than two qubits to Bob. It was shown by Bennett and Wiesner [2], that by using a previously shared singlet (between Alice and Bob), Alice will be able to send two bits to Bob, by sending just a single qubit.

The protocol of dense coding [2] works as follows. Assume that Alice and Bob share a singlet state

$$|\psi^-\rangle = \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle). \quad (62)$$

The crucial observation is that this entangled two-qubit state can be transformed into four orthogonal states of the two-qubit Hilbert space by performing unitary operations on just a single qubit. For instance, Alice can apply a rotation (the Pauli operations) or do nothing to her part of the singlet, while Bob does nothing, to obtain the three triplets (or the singlet):

$$\begin{aligned} \sigma_x \otimes I |\psi^-\rangle &= -|\phi^-\rangle, & \sigma_y \otimes I |\psi^-\rangle &= i|\phi^+\rangle, \\ \sigma_z \otimes I |\psi^-\rangle &= |\psi^+\rangle, & I \otimes I |\psi^-\rangle &= |\psi^-\rangle, \end{aligned} \quad (63)$$

where

$$\begin{aligned} |\psi^\pm\rangle &= \frac{1}{\sqrt{2}} (|01\rangle \pm |10\rangle), \\ |\phi^\pm\rangle &= \frac{1}{\sqrt{2}} (|00\rangle \pm |11\rangle), \end{aligned} \quad (64)$$

and I is the qubit identity operator. Suppose that the classical information that Alice wants to send to Bob is i , where $i = 0, 1, 2, 3$. Alice and Bob previously agree on the following correspondence between the operations applied at Alice's end and the information i that she wants to send:

$$\begin{aligned} \sigma_x &\Rightarrow i = 0, & \sigma_y &\Rightarrow i = 1, \\ \sigma_z &\Rightarrow i = 2, & I &\Rightarrow i = 3. \end{aligned} \quad (65)$$

Depending on the classical information she wishes to send, Alice applies the appropriate rotation on her part of the shared singlet, according to the above correspondence. Afterwards, Alice sends her part of the shared state to Bob, via the noiseless quantum

channel. Bob now has in his possession, the entire two-qubit state, which is in any of the four Bell states $\{|\psi^\pm\rangle, |\phi^\pm\rangle\}$. Since these states are mutually orthogonal, he will be able to distinguish between them and hence find out the classical information sent by Alice.

To consider a more realistic scenario, usually two avenues are taken. One approach is to consider a *noisy* quantum channel, while the additional resource is an arbitrary amount of shared bipartite *pure* state entanglement (see e.g. [34, 35], see also [36, 37]). The other approach is to consider a *noiseless* quantum channel, while the assistance is by a given bipartite *mixed* entangled state (see e.g. [36, 37, 38, 39, 40, 41]).

Here, we consider the second approach, and derive the capacity of dense coding in this scenario, for a given state, where the capacity is defined as the number of classical bits that can be accessed by the receivers, per usage of the noiseless channel. This will lead to a classification of bipartite states according to their degree of ability to assist in dense coding. In the case where a noisy channel and an arbitrary amount of shared pure entanglement is considered, the capacity refers to the channel (see e.g. [34, 35]). However, in our case when a noiseless channel and a given shared (possibly mixed) state is considered, the capacity refers to the state. Note that the mixed shared state in our case can be thought of as an output of a noisy channel. A crucial element in finding the capacity of dense coding is the Holevo bound [42], which is a universal upper bound on classical information that can be decoded from a quantum ensemble. Below we discuss the bound, and subsequently derive the capacity of dense coding.

A. The Holevo bound

The Holevo bound is an upper bound on the amount of classical information that can be accessed from a quantum ensemble in which the information is encoded. Suppose therefore that Alice (*A*) obtains the classical message i that occurs with probability p_i , and she wants to send it to Bob (*B*). Alice encodes this information i in a quantum state ρ_i , and sends it to Bob. Bob receives the ensemble $\{p_i, \rho_i\}$, and wants to obtain as much information as possible about i . To do so, he performs a measurement, that gives the result m , with probability q_m . Let the corresponding post-measurement ensemble be $\{p_{i|m}, \rho_{i|m}\}$. The information gathered can be quantified by the mutual information between the message index i and the measurement outcome [43]:

$$I(i : m) = H(\{p_i\}) - \sum_m q_m H(\{p_{i|m}\}). \quad (66)$$

Note that the mutual information can be seen as the difference between the initial disorder and the (average) final disorder. Bob will be interested to obtain the maximal information, which is maximum of $I(i : m)$ for all measurement strategies. This quantity is called the accessible information:

$$I_{acc} = \max I(i : m), \quad (67)$$

where the maximization is over all measurement strategies.

The maximization involved in the definition of accessible information is usually hard to compute, and hence the importance of bounds [42, 44]. In particular, in Ref. [42], a universal upper bound, the Holevo bound, on I_{acc} is given:

$$I_{acc}(\{p_i, \rho_i\}) \leq \chi(\{p_i, \rho_i\}) \equiv S(\bar{\rho}) - \sum_i p_i S(\rho_i). \quad (68)$$

See also [45, 46, 47]. Here $\bar{\rho} = \sum_i p_i \rho_i$ is the average ensemble state, and $S(\varsigma) = -\text{tr}(\varsigma \log_2 \varsigma)$ is the von Neumann entropy of ς .

The Holevo bound is asymptotically achievable in the sense that if the sender Alice is able to wait long enough and send long strings of the input quantum states ρ_i , then there exists a particular encoding and a decoding scheme that asymptotically attains the bound. Moreover, the encoding consists in collecting certain long and “typical” strings of the input states, and sending them all at once [48, 49].

B. Capacity of quantum dense coding

Suppose that Alice and Bob share a quantum state ρ_{AB} . Alice performs the unitary operation U_i with probability p_i , on her part of the state ρ_{AB} . The classical information that she wants to send to Bob is i . Subsequent to her unitary rotation, she sends her part of the state ρ^{AB} to Bob. Bob then has the ensemble $\{p_i, \rho_i\}$, where

$$\rho_i = U_i \otimes I \rho_{AB} U_i^\dagger \otimes I.$$

The information that Bob is able to gather is $I_{acc}(\{p_i, \rho_i\})$. This quantity is bounded above by $\chi(\{p_i, \rho_i\})$, and is asymptotically achievable. The “one-capacity” $C^{(1)}$ of dense coding for the state ρ_{AB} is the Holevo bound for the best encoding by Alice:

$$C^{(1)}(\rho) = \max_{p_i, U_i} \chi(\{p_i, \rho_i\}) \equiv \max_{p_i, U_i} \left(S(\bar{\rho}) - \sum_i p_i S(\rho_i) \right). \quad (69)$$

The superscript (1) reflects the fact that Alice is using the shared state once at a time, during the asymptotic process. She is not using entangled unitaries on more than one copy of her parts of the shared states ρ_{AB} . As we will see below, encoding with entangled unitaries does not help her to send more information to Bob.

In performing the maximization in Eq. (69), first note that the second term in the right hand side (rhs) is $-S(\rho)$, for all choices of the unitaries and probabilities. Secondly, we have

$$S(\bar{\rho}) \leq S(\bar{\rho}_A) + S(\bar{\rho}_B) \leq \log_2 d_A + S(\bar{\rho}_B),$$

where d_A is the dimension of Alice’s part of the Hilbert space of ρ_{AB} , and $\bar{\rho}_A = \text{tr}_B \bar{\rho}$, $\bar{\rho}_B = \text{tr}_A \bar{\rho}$. Moreover, $S(\bar{\rho}_B) = S(\rho_B)$, as nothing was done at Bob’s end during the encoding procedure. (In any case, unitary operations does not change the spectrum, and hence the entropy, of a state.) Therefore, we have

$$\max_{p_i, U_i} S(\bar{\rho}) \leq \log_2 d_A + S(\rho_B).$$

But the bound is reached by any complete set of orthogonal unitary operators $\{W_j\}$, to be chosen with equal probabilities, which satisfy the *trace rule* $\frac{1}{d_A^2} \sum_j W_j^\dagger \Xi W_j = \text{tr}[\Xi] I$, for any operator Ξ . Therefore, we have

$$C^{(1)}(\rho) = \log_2 d_A + S(\rho_B) - S(\rho).$$

The optimization procedure above sketched essentially follows that in Ref. [41]. Several other lines of argument are possible for the maximization. One is given in Ref. [39] (see also [50]). Another way to proceed is to guess where the maximum is reached (maybe from examples or by taking the most symmetric option), and then perturb the guessed result. If the first order perturbations vanish, the guessed result is correct, as the von Neumann entropy is a concave function and the maximization is carried out over a continuous parameter space.

Without using the additional resource of entangled states, Alice will be able to reach a capacity of just $\log_2 d_A$ bits. Therefore, entanglement in a state ρ^{AB} is useful for dense coding if $S(\rho^B) - S(\rho) > 0$. Such states will be called dense-codeable (DC) states. Such states exist, an example being the singlet state.

Note here that if Alice is able to use entangled unitaries on two copies of the shared state ρ , the capacity is not enhanced (see Ref. [51]). Therefore, the one-capacity is really the asymptotic capacity, in this case. Note however that this additivity is known only in the case of encoding by unitary operations. A more general encoding may still have additivity problems (see e.g. [37]). Here, we have considered unitary encoding only. This case is both mathematically more accessible, and experimentally more viable.

A bipartite state ρ^{AB} is useful for dense coding if and only if $S(\rho^B) - S(\rho) > 0$. It can be shown that this relation cannot hold for PPT entangled states [36] (see also [50]). Therefore a DC state is always NPT. However, the converse is not true: There exist states which are NPT, but they are not useful for dense coding. Examples of such states can be obtained by the considering the Werner state $\rho_p = p|\psi^-\rangle\langle\psi^-| + \frac{1-p}{4}I \otimes I$ [4].

The discussions above leads to the following classification of bipartite quantum states:

1. Separable states: These states are of course not useful for dense coding. They can be prepared by LOCC.
2. PPT entangled states: These states, despite being entangled, cannot be used for dense coding. Moreover, their entanglement cannot be detected by the partial transposition criterion.
3. NPT non-DC states: These states are entangled, and their entanglement can be detected by the partial transposition criterion. However, they are not useful for dense coding.
4. DC states: These entangled states can be used for dense coding.

The above classification is illustrated in figure 4. A generalisation of this classification has been considered in Refs. [50, 51].

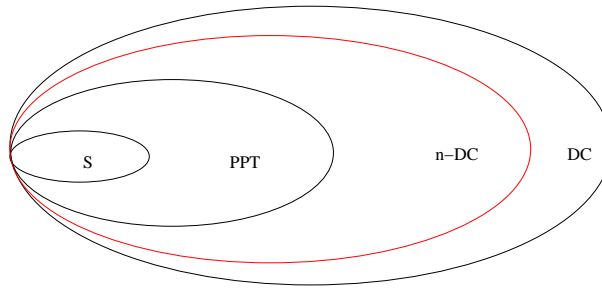


FIG. 4: Classification of bipartite quantum states according to their usefulness in dense coding. The convex innermost region, marked as S, consists of separable states. The shell surrounding it, marked as PPT, is the set of PPT entangled states. The next shell, marked as n-DC, is the set of all states that are NPT, but not useful for dense coding. The outermost shell is that of dense-codeable states.

VIII. FURTHER READING: MULTIPARTITE STATES

The discussion about detection of bipartite entanglement presented above is of course quite far from complete. For further reading, we have presented a small sample of references embedded in the text above. We prefer to conclude this chapter with a few remarks on multipartite states.

The case of detection of entanglement of pure states is again simple. One quickly realizes that a multipartite pure state is entangled if and only if it is entangled in at least one bipartite splitting. So, for example, the state $|\text{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$ [52] is entangled, because it is entangled in the A:BC bipartite splitting (as also in all others).

The case of mixed states is however quite formidable. In particular, the results obtained in the bipartite mixed state case, cannot be applied to the multipartite scenario. One way to see this is to notice the existence of states which are separable in any bipartite splitting, while the entire state is entangled. An example of such a state is given in Ref. [53]. For further results about entanglement criteria, detection, and classification of multipartite states, see e.g. [21, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64], and references therein.

IX. PROBLEMS

Problem 1 Show that the singlet state has nonpositive partial transposition.

Problem 2 Consider the Werner state $p|\psi^-\rangle\langle\psi^-| + (1-p)I/4$ in $2 \otimes 2$, where $0 \leq p \leq 1$ [4]. Find the values of the mixing parameter p , for which entanglement in the Werner state can be detected by the partial transposition criterion.

Problem 3 Show that in $\mathcal{C}^2 \otimes \mathcal{C}^2$, the partial transposition of a density matrix can have at most one negative eigenvalue.

Problem 4 Given two random variables X and Y , show that the Shannon entropy of the joint distribution cannot be smaller than that of either.

Problem 5 Prove Theorem 5.

Problem 6 Prove Lemma 1.

Problem 7 Prove Theorem 10.

Problem 8 Show that each of the shells depicted in figure 4 are nonempty, and of nonzero measure. Show also that all the boundaries are convex.

[*] Also at Institució Catalana de Recerca i Estudis Avançats.

[1] A.K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[2] C.H. Bennett and S.J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).

[3] C.H. Bennett, G. Brassard, C. Crepeau, R. Josza, A Peres, and W.K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).

[4] R.F. Werner, Phys. Rev. A **40**, 4277(1989).

[5] M. Żukowski, A. Zeilinger, M.A. Horne and A.K. Ekert, Phys. Rev. Lett. **71**, 4287 (1993); M. Żukowski, A. Zeilinger, and H. Weinfurter, Annals N.Y. Acad. Sci. **755**, 91 (1995); S. Bose, V. Vedral, and P.L. Knight, Phys. Rev. A **57**, 822 (1998); S. Bose, V. Vedral, and P.L. Knight, Phys. Rev. A **60**, 194 (1999).

[6] S. Karnas and M. Lewenstein, J. Phys. A **34**, 6919 (2001).

[7] A. Peres, Phys. Rev. Lett **77**, 1413 (1996).

[8] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).

- [9] M.A. Nielsen and J. Kempe, Phys. Rev. Lett. **86**, 5184 (2001).
- [10] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (2000).
- [11] O. Rudolph, *Further results on the cross norm criterion for separability*, quant-ph/0202121; K. Chen and L.-A. Wu, Quant. Inf. Comput., **3**, 193 (2003).
- [12] A. Doherty, P. Parillo, and F. Spedalieri, Phys. Rev. Lett. **88**, 187904 (2002); Phys. Rev. A **69**, 022308 (2004); F. Hulpke and D. Bruß, J. Phys. A: Math. Gen. **38**, 5573 (2005).
- [13] P. Horodecki, Phys. Lett. A **232**, 333 (1997).
- [14] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).
- [15] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, Phys. Rev. A, **58**, 883 (1998); K. Życzkowski, Phys. Rev. A **60**, 3496 (1999); S. Szarek, *The volume of separable states is super-doubly-exponentially small*, quant-ph/0310061 (Phys. Rev. A, in print), and references therein.
- [16] D.P. DiVincenzo, P.W. Shor, J.A. Smolin, B.M. Terhal, and A.V. Thapliyal, Phys. Rev. A **61**, 062312 (2000); W. Dür, J.I. Cirac, M. Lewenstein, and D. Bruß, *ibid.* **61**, 062313 (2000).
- [17] A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications* (Academic Press, New York, 1979); P.M. Alberti and A. Uhlmann, *Stochasticity and Partial order: Doubly Stochastic Maps and Unitary Mixing* (Dordrecht, Boston, 1982); R. Bhatia, *Matrix Analysis* (Springer, New York, 1997).
- [18] T. Hiroshima, Phys. Rev. Lett. **91**, 057902 (2003).
- [19] R. Horodecki, P. Horodecki, and M. Horodecki, Phys. Lett. A **210**, 377 (1996).
- [20] O. Gühne, P. Hyllus, D. Bruß, A. Ekert, M. Lewenstein, C. Macchiavello, and A. Sanpera, Phys. Rev. A **66**, 062305 (2002); O. Gühne, P. Hyllus, D. Bruß, A. Ekert, M. Lewenstein, C. Macchiavello, and A. Sanpera, J. Mod. Opt. **50**, 1079 (2003).
- [21] M. Bourennane, M. Eibl, C. Kurtsiefer, S. Gaertner, H. Weinfurter, O. Gühne, P. Hyllus, D. Bruß, M. Lewenstein, and A. Sanpera, Phys. Rev. Lett. **92**, 087902 (2004).
- [22] H.W. Alt, *Lineare Funktionalanalysis*, (Springer-Verlag, 1985).
- [23] S.L. Woronowicz, Commun. Math. Phys. **51**, 243 (1976); P. Kryszynski and S.L. Woronowicz, Lett. Math. Phys. **3**, 319 (1979); M.D. Choi, Proc. Sympos. Pure Math. **38**, 583 (1982).
- [24] B.M. Terhal, Lin. Alg. Appl. **323**, 61 (2001).
- [25] M. Lewenstein, B. Kraus, J.I. Cirac, and P. Horodecki, Phys. Rev. A **62**, 052310 (2000); D. Bruß, J.I. Cirac, P. Horodecki, F. Hulpke, B. Kraus, M. Lewenstein, and A. Sanpera, J. Mod. Opt. **49**, 1399 (2002).
- [26] A. Jamiołkowski, Rep. Math. Phys., **3**, 275 (1972).
- [27] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **57**, 777 (1935).
- [28] J.S. Bell, Physics **1**, 195 (1964).
- [29] E. Santos, Phys. Rev. Lett. **66**, 1388 (1991); E. Santos, Phys. Rev. A **46**, 3646 (1992); P.M. Pearle, Phys. Rev. D **2**, 1418 (1970); J.F. Clauser and M.A. Horne, Phys. Rev. D **10**, 526 (1974); P.G. Kwiat, P.H. Eberhard, A.M. Steinberg, and R.Y. Chiao, Phys. Rev. A **49**, 3209 (1994); N. Gisin and B. Gisin, Phys. Lett. A **260**, 323 (1999); S. Massar, S. Pironio, J. Roland, and B. Gisin, Phys. Rev. A **66**, 052112 (2002); R. Garcia-Patron, J. Fiurásek, N.J. Cerf, J. Wenger, R. Tualle-Brouiri, and Ph. Grangier, Phys. Rev. Lett. **93**, 130409 (2004).
- [30] S.J. Freedman and J.S. Clauser, Phys. Rev. Lett. **28**, 938 (1972); A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. **47**, 460 (1981); *ibid.* **49**, 91 (1982); A. Aspect, J. Dalibard, and G. Roger, *ibid.* **49**, 1804 (1982); P.G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A.V. Sergienko, and Y. Shih, Phys. Rev. Lett. **75**, 4337 (1995); G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, Phys. Rev. Lett. **81**, 5039 (1998); M.A. Rowe, D. Kielpinski, V. Meyer, C.A. Sackett, W.M. Itano, C. Monroe, and D.J. Wineland, Nature **409**, 791 (2001); W. Tittel, B. Gisin, T. Herzog, H. Zbinden, and N. Gisin, Phys. Rev. A **57**, 3229 (1998); A. Zeilinger Rev. Mod. Phys. **71**, S288 (1999).
- [31] J.F. Clauser, M.A. Horne, A. Shimony, and R.A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
- [32] J.S. Bell, in *Foundations of Quantum Mechanics*, ed. B. d'Espagnat (Academic, New York, 1971).
- [33] P. Hyllus, O. Gühne, D. Bruß, and M. Lewenstein Phys. Rev. A **72**, 012321 (2005).
- [34] C.H. Bennett, P.W. Shor, J.A. Smolin, and A.V. Thapliyal, Phys. Rev. Lett. **83**, 3081 (1999).
- [35] C.H. Bennett, P.W. Shor, J.A. Smolin, and A.V. Thapliyal, *Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem*, quant-ph/0106052.
- [36] M. Horodecki, P. Horodecki, R. Horodecki, D. Leung, and B. Terhal, Q. Inf. and Comput. **1**, 70 (2001).
- [37] A. Winter, J. Math. Phys. **43**, 4341 (2002).
- [38] S. Bose, M.B. Plenio, and V. Vedral, J. Mod. Opt. **47**, 291 (2000).
- [39] T. Hiroshima, J. Phys. A: Math. Gen. **34**, 6907 (2001).
- [40] X.S. Liu, G.L. Long, D.M. Tong, and F. Li, Phys. Rev. A **65**, 022304 (2002).
- [41] M. Ziman and V. Bužek, Phys. Rev. A **67**, 042321 (2003).
- [42] J.P. Gordon, in *Proc. Int. School Phys. "Enrico Fermi, Course XXXI"*, ed. P.A. Miles, p. 156 (Academic Press, NY 1964); L.B. Levitin, in *Proc. VI National Conf. Inf. Theory, Tashkent*, p. 111 (1969); A.S. Holevo, Probl. Pereda. Inf. **9**, 3 1973 [Probl. Inf. Transm. **9**, 110 (1973)].
- [43] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
- [44] R. Josza, D. Robb, and W.K. Wootters, Phys. Rev. A, **49**, 668 (1994).
- [45] B. Schumacher, M. Westmoreland, and W. K. Wootters, Phys. Rev. Lett. **76**, 3452 (1996).
- [46] P. Badziąg, M. Horodecki, A. Sen(De), and U. Sen, Phys. Rev. Lett. **91**, 117901 (2003).
- [47] M Horodecki, J. Oppenheim, A. Sen(De), and U. Sen, Phys. Rev. Lett. **93**, 170503 (2004).
- [48] B. Schumacher and M.D. Westmoreland, Phys. Rev. A **56**, 131 (1997).
- [49] A.S. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998).
- [50] D. Bruß, G.M. D'Ariano, M. Lewenstein, C. Macchiavello, A. Sen(De), and U. Sen, Phys. Rev. Lett. **93**, 210501 (2004).

- [51] D. Bruß, G.M. D'Ariano, M. Lewenstein, C. Macchiavello, A. Sen(De), and U. Sen, *Dense coding with multipartite quantum states*, quant-ph/0507146.
- [52] D.M. Greenberger, M.A. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, ed. M. Kafatos (Kluwer Academic, Dordrecht, 1989).
- [53] C.H. Bennett, D.P. DiVincenzo, T. Mor, P.W. Shor, J.A. Smolin, and B.M. Terhal, Phys. Rev. Lett. **82**, 5385 (1999).
- [54] W. Dür, J.I. Cirac, and R. Tarrach, Phys. Rev. Lett. **83**, 3562 (1999).
- [55] R.F. Werner and M.M. Wolf, Phys. Rev. A **64**, 032112 (2001); M. Żukowski and Č. Brukner, Phys. Rev. Lett. **88**, 210401 (2002)
- [56] G. Svetlichny, Phys. Rev. D **35**, 3066 (1987); D. Collins, N. Gisin, S. Popescu, D. Roberts, and V. Scarani Phys. Rev. Lett. **88**, 170405 (2002); M. Seevinck and G. Svetlichny, Phys. Rev. Lett. **89**, 060401 (2002); S.M. Roy, Phys. Rev. Lett. **94**, 010402 (2005).
- [57] A. Acín, D. Bruß, M. Lewenstein, and A. Sanpera, Phys. Rev. Lett. **87**, 040401 (2001).
- [58] M. Horodecki, P. Horodecki, and R. Horodecki, *Separability of mixed quantum states: linear contractions approach*, quant-ph/0206008.
- [59] C. Moura Alves and D. Jaksch Phys. Rev. Lett. **93**, 110501 (2004).
- [60] W. Laskowski, T. Paterek, M. Żukowski, and Č. Brukner, Phys. Rev. Lett. **93**, 200401 (2004).
- [61] F.G.S.L. Brandao and R.O. Vianna, Phys. Rev. Lett. **93**, 220503 (2004).
- [62] G. Toth and O. Gühne, Phys. Rev. Lett. **94**, 060501 (2005).
- [63] A.C. Doherty, P.A. Parrilo, and F.M. Spedalieri, Phys. Rev. A **71**, 032333 (2005).
- [64] O. Gühne, G. Toth, and H.J. Briegel, *Multipartite entanglement in spin chains*, quant-ph/0502160.
- [65] Due to the existence of the phenomenon of entanglement swapping [5], one must suitably enlarge the notion of preparation of entangled states. So, an entangled state between two particles can be prepared if and only if, either the two particles (call them A and B) themselves come together to interact at a time in the past, or two *other* particles (call them C and D) does the same, with C (D) having interacted beforehand with A (B).
- [66] A unitary operator on $\mathcal{H}_A \otimes \mathcal{H}_B$, is said to be “nonlocal”, if it is not of the form $U_A \otimes U_B$, with U_A (U_B) being a unitary operator acting on \mathcal{H}_A (\mathcal{H}_B).
- [67] The von Neumann entropy of a state ρ is $S(\rho) = -\text{tr} \rho \log_2 \rho$.
- [68] The Shannon entropy of a random variable X , taking up values X_i , with probabilities p_i , is given by $H(X) = H(\{p_i\}) = -\sum_i p_i \log_2 p_i$.
- [69] A matrix $D = (D_{ij})$ is said to be doubly stochastic, if $D_{ij} \geq 0$, and $\sum_i D_{ij} = \sum_j D_{ij} = 1$.
- [70] A hyperplane is a linear subspace with dimension one less than the dimension of the space itself.
- [71] We do not derive here the original Bell inequality, which Bell derived in 1964 [28]. Instead, we derive the stronger form of the Bell inequality which Clauser, Horne, Shimony, and Holt (CHSH) derived in 1969 [31]. A similar derivation was also given by Bell himself in 1971 [32].