Politecnico di Torino

1859

Master's Degree in Physics of Complex Systems

Quantifying Entanglement

Supervisor:
Prof. Davide Girolami

Candidate:
Vincenzo Maria Pianese

July 2024

Acknowledgements

I ringraziamenti riportati in questa pagina servono a pagare il tributo di gratitudine per la maggior parte del debito che riconosco verso i miei compagni di università e generalmente verso persone che ho ritenuto importanti.Tuttavia, nulla di ciò che dirò riuscirà a dare più che una pallida idea del numero e della natura dei debiti personali che ho verso molti individui con cui mi sono confrontato: l'elenco di tutti coloro che possono trovare segni precisi della loro influenza in questa pagina avrebbe quasi la stessa estensione dell'elenco dei miei amici e conoscenti. In circostanze come queste, devo limitarmi a ricordare quelle più significative influenze che anche anche una labile memoria non riuscirà mai a dimenticare completamente.

Mi duole che la mancanza di spazio mi tolga la soddisfazione di ringraziare tutti singolarmente. Tuttavia non posso lasciarmi sfuggire l'occasione di rendere atto del grande debito che ho nei confronti del mio relatore Davide Girolami che, negli ultimi 7 mesi, mediante suggeriementi e critiche, ha sorretto e orientato lo sviluppo di questa tesi, collaborando con me in tutti i modi possibili con il suo vasto tesoro di cultura ed il suo alto senso critico.

Ringrazio tutti i ragazzi incontrati nel biennio al Camplus: insieme a voi ho condiviso molte delle fatiche e delle soddisfazioni che ha regalato questo percorso universitario: la pertinace ostinazione nel concluderlo a tutti i costi si è alimentata anche della vostra amicizia, che mi auguro si protragga ancora per tanto tempo.

Grazie agli amici superstiti alla decade formata dagli anni del liceo e dell'università. Benché ciascuno abbia preso la propria strada, continuano a legarci frammenti di esistenza che, sono certo, lasceranno ricordi vivi e pregnanti in futuro.

Saluto i ragazzi di Interno 23, a cui attribuisco l'indiscusso merito di avermi forgiato, incentivandomi a sviluppare la mia resilienza, rendendola tale da sembrare una sorta di scudo adamantino contro le avversità. Questo loro intento non dichiarato ha costituito per me una palestra di vita. Desidero fare una menzione speciale a Gabriele, che probabilmente è l'unica persona con la quale io sia riuscito ad esplorare le mie idee nella forma di frasi incomplete. Questa forma di comunicazione prova un intendimento reciproco, grazie al quale egli è stato in grado di indicarmi la via per superare o aggirare alcuni grossi ostacoli che ho incontrato nel preparare alcuni esami.

Dedico un pensiero agli amici e, più generalmente, alle persone che mi hanno accompagnato per un tratto breve di questo percorso. A voi, improbabili destinatari di un giorno di giubilo, dico che ho imparato e sto imparando parecchio. Sto cercando di migliorare l'attitudine nella disciplina dello slalom della sollecitazione emozionale e soprattutto in quella pratica filosofica del "lasciare andare".

Devo infine pagare un tributo di speciale riconoscimento ai miei cari e, soprattutto, a mio fratello Fernando. Sebbene non abbiano contribuito intellettualmente al mio lavoro, ciascuno di loro ha fatto qualcosa di più importante. Lo hanno, cioè, lasciato continuare ed addirittura hanno incoraggiato la mia devozione ad esso. Chiunque abbia lottato con un progetto come il mio comprenderà quanto esso abbia potuto costare loro in parecchie occasioni. Non so come ringraziarli.

Contents

Abstract

As a relevant concept in Physics, the role played by entanglement in quantum information processing is paramount, as it is the main resource enabling considerable technological achievements, such as quantum communication, quantum cryptography, quantum computational speed-up and so on. Mathematically, ***Quantum entanglement*** is the most unbelievable non-classical property of compound states that cannot be decomposed as a statistical mixtures of product states over subsystems, and it has a very complex structure, encasing many features described by just as much entanglement measures. Among all of them, the QUANTUM CONDITIONAL MUTUAL INFORMATION is a particularly interesting one and represents the object this thesis is devoted to. The exact quantification of many information measures is tipycally a daunting problem because it would involve a huge amount of computational resources that outstrip the capability on any existing computer : since the eigenvalues and relevant entropies of a density matrix operator and its subsystems are expected to be known, this task becomes quickly computationally demanding for large enough systems, setting itself as seemingly unsolvable. For this reason, it behooves us to understand how we can keep using them while endeavoring to devise some meaningful approximations. The upshot is that we try to handle with the aforementioned quantity in terms of **`lower`** and **`upper bounds`**, which can be obtained analytically in a simple yet effective way. This purpose can be generalized to other important information measures and foreshadows a topic of great concern in Quantum Information, i.e. the possibility to deal with operationally useful entanglement measures without spectrum reconstruction. In particular, we emphasize how our lower bound is tighter than the Carlen and Lieb's one, a well - known scientific result establishing a sharper refinement on the strong subadditivity inequality of the Von Neumann Entropy. Moreover, we run simulations over a suitable quantum state whose outcomes ostensibly validate our proposal. However, hitting the bullseye comes up after drawing a parallel between the classical and quantum worlds concerning some preliminary notions of *Information Theory*, to which we will dedicate the first and second chapter of this thesis, respectively. We will then proceed to the analytical derivation of the bounds and subsequently apply them to a mixed state in order to check their actual correctness.

Introduction

Among all the fundamental aspects of Quantum Mechanics, viz. INDETERMINISM, INTERFERENCE, UNCERTAINTY, SUPERPOSITION and ENTANGLEMENT, the latter is not only the most intriguing phenomenon, being inherently quantum and having no classical counterpart, but also the most involved in the definition of some physical quantities we will take into account in the present discussion.

Strictly speaking, entanglement arises in some composite quantum systems and refers to the property of their particles to exhibit quantum correlations stronger than any classical ones. Despite its seemingly strange features caused many scientists to be wary of a common misconception about the comprehension of the world, leading them to doubt whether reality and locality principles are consistent or not, Entanglement is a foundational concept that turned out to be very useful in quantum computation and quantum information.
Indeed, it has been widely recognized that quantum computers harnessing entanglement can outperform classical ones in solving problems efficiently. Moreover, it spawned a large area of research in Quantum Information: in the framework of Quantum Information science, many quantum protocols have been developed, e.g *Quantum teleportation*, that exploit the combination of entanglement and classical communication for the transmission of quantum information, or *Super dense coding*, merging entanglement and noiseless qubit channels for allowing the transmission of more classical information than would be possible with a noiseless qubit channel alone; both tasks are thus amazing applications of Quantum physics to the realm of Information theory which are far beyond the possibility of any classical implementation.

Driven by the desire to achieve a deeper comprehension of these mechanisms and a plethora of other ones we do not have time to mention, albeit they are a firm underpinning for Quantum Physics, this goal of this thesis is to quantify entanglement by taking into account relevant correlation measures that can be expressed in terms of the Von Neumann Entropy. Exploiting some well known results establishing lower and upper bounds for quantum entropy, our aim is to devise a methodological procedure for getting the same kind of bounds for an important informational measure, i.e. the **Quantum Conditional Mutual Information** (often referred to as **QCMI**). We confirm our intuition trying to obtain further refinements of the same results, in order to improve our computational accuracy and trying to achieve better analytical outcomes.

Chapter 1

Classical Information Theory: A Mixed Appetizer

You cannot evade quantity. You may fly to poetry and music, and quantity and numbers will face you in your rhytms and your octaves
Alfred North Whitehead

This chapter presents a simple introduction to *Classical Information Theory*, outlining theoretical notions and some useful concepts that provide the necessary background for a deeper level of understanding of the subsequent sections.
Indeed, the development of *Quantum Information Theory*, whom we will refer in the next chapter, has been hastened by the combination of the features of Quantum mechanics and the aforementioned key concepts. Neither mathematical rigor is a matter of concern nor we cover all aspects of this field: the purpose is solely to provide a systematic and not overly heavy guide for developing the essential toolkit, thus acting as a stepping stone for further investigation.

1.1 The Shannon Entropy

Classical Information Theory hinges on the concept of *Shannon entropy.* Given a random variable X, it quantifies how much information we gain on average when we learn its value. It is worth mentioning that there is a substantial difference between the meanings given to the same term, as dictated by the common sense and as set in the present discussion.
Here, the term ”information” refers to our prior ”ignorance” or ”uncertainty” about X before learning it. Taken for granted these two complementary views, one can thus employ the words ”uncertainty” and ”information” interchangeably. The Shannon entropy associated with the probability distribution

$$\{p_1, p_2, \ldots, p_k\}$$

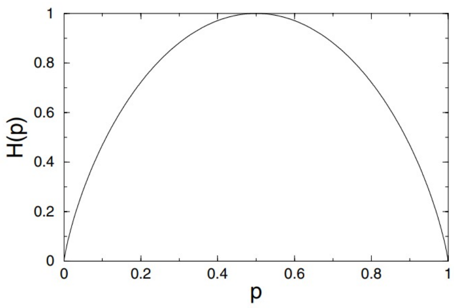

Figure 1.1: The Shannon binary entropy H(p)

is defined by

$$H(p_1, p_2, \ldots, p_k) \equiv -\sum_{i=1}^{k} p_i \log p_i$$

where the logarithm is written in base 2. A clarification of the insightful discussion concerning the meaning of information in this framework is bestowed by the following special case in which $k = 2$. Upon defining $p_1 = p \quad and \quad p_2 = 1 - p \quad (0 \leq p \leq 1)$, the *Shannon binary entropy* can be readily written as

$$H(p) \equiv H(p_1, p_2) = -p \log p - (1 - p) \log (1 - p)$$

A simple plot of H versus p shows that the entropy vanishes for both $p = 0$ and $p = 1$, whereas it attains its maximum value H = 1 when $p = 1/2$ (look at the figure 1.1): this confirms what we stated above: the Shannon entropy is a measure of our uncertainty. If we know we shall receive letter a_1 for sure (and therefore $p = 1$), then there would not be any information gain when receiving that letter; the same argument holds when letter a_2 will be given with certainty ($p = 0$). However, when both letters are equiprobable ($p = 1/2$) we would gain the maximum information, namely a bit of information (H(1/2) = 1) because our a priori ignorance is maximum.

Suppose X is a random variable that takes the value x from the alphabet $\{a_1, a_2, \ldots, a_k\}$with probability $p(x) \in \{p_1, p_2, \ldots, p_k\}$.

Then the Shannon entropy $H(p_1, p_2, p_k)$ is also denoted as H(X) and we can write

$$H(X) = -\sum_{x} p(x)\ \log(p(x)) = -\sum_{i=1}^{k} p_i\ \log p_i$$

Now we can proceed in further definitions of remarkable information quantities.

1.2 Joint Entropy

The *Joint Entropy* endows a generalization of the idea of uncertainty concerning two discrete random variables X and Y taking values x and y with probabilities $p(x)$ and $p(y)$, respectively. We can define the entropy of the joint random variable (X,Y) :

Definition (Joint Entropy)**.** Let X and Y be discrete random variables with joint probability distribution $p(x, y)$. The joint entropy H(X,Y) is defined as

$$H(X, Y) \equiv -\sum_{x,y} p(x, y) \, \log p(x, y)$$

where $p(x, y)$ is the probability that $X = x$ and $Y = y$.

1.3 Conditional Entropy

Suppose X and Y are two discrete random variables sharing correlation because they are not statistically independent. Then the *Conditional Entropy* $H(X \mid Y)$ describes the amount of uncertainty one has about X given the value of Y. It can be shown that its expression is the following :

$$H(X \mid Y) = -\sum_{x,y} p(x, y) \, log \, p(x \mid y)$$

where $p(x \mid y) = p(x, y)/p(y)$ is the probability that $Y = y$ provided $X = x$. Indeed,

$$\begin{aligned} H(X \mid Y) &= H(X, Y) - H(Y) \\ &= -\sum_{x,y} p(x, y) \, \log p(x, y) + \sum_{y} p(y) \, \log p(y) \\ &= -\sum_{x,y} p(x, y) \, \log(p(y) \, p(x \mid y)) + \sum_{x,y} p(x, y) \, \log p(y) \\ &= -\sum_{x,y} p(x, y) \, \log p(x \mid y) \end{aligned} \tag{1.1}$$

where we exploited the marginalization over variable Y,

$$\sum_{x} p(x, y) = p(y),$$

and we gave the first line of equation 1.1 this simple meaning: given the value of Y, H(Y) counts the number of bits of information about the pair (X,Y); the residual ignorance about (X,Y) is associated with the residual lack of knowledge about X, when the knowledge of Y is provided. Rather conceivably, the conditional entropy $H(X \mid Y)$ should be less than or equal to the entropy H(X): we expect this to happen, since having access to a side

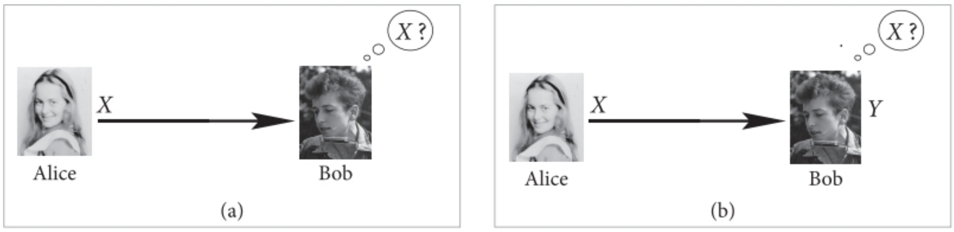

Figure 1.2: (a) The entropy H(X) is the uncertainty that Bob has about random variable X before learning it. (b) The conditional entropy $H(X \mid Y)$ is the uncertainty that Bob has about X when he already possesses Y.

variable Y should only decrease the uncertainty about another variable X it is related with. This compelling intuition is expressed in the following statement :

Theorem (Conditioning does not increase entropy)**.** *The entropy H(X) is greater than or equal to the conditional entropy $H(X \mid Y)$, and the inequality is saturated if and only if X and Y are independent random variables:*

$$H(X) \geq H(X \mid Y)$$

1.4 Mutual Information

A third relevant quantity is the *Mutual Information* content of random variables X and Y:

Definition (Mutual Information)**.** Let X and Y be discrete random variables with joint probability distribution $p(x, y)$. The mutual information $I(X : Y)$ is given by:

$$\begin{aligned} I(X : Y) &\equiv H(X) + H(Y) - H(X, Y) \\ &= H(X) - H(X \mid Y) \end{aligned} \tag{1.2}$$

It quantifies the amount of common information shared by X and Y. This can be devised by the the upper line of the above definition given to $I(X : Y)$: when we add H(X), that is the information content of X to the information content of Y, H(Y), overlapping information will be counted twice, while information which is not common will be merely counted once. Therefore, subtracting off the joint information of (X,Y), $H(X, Y)$, one has the mutual (or common) information of X and Y. Although it is just a consequence of the first line, the second expression has another nice interpretation: knowledge of Y implies an uncertainty $H(X|Y)$ about variable

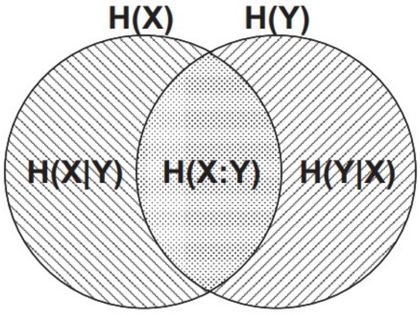

Figure 1.3: Relationships between different entropies

X or, equivalently, accessing to Y supplies an information gain of $H(X|Y)$ bits about X, hence it reduces the entire uncertainty H(X) about X, the one we do not have any side information on. An alternative expression of the mutual information $I(X:Y)$ is provided below in terms of marginal and joint probability density functions $p(x)$, $p(y)$ and $p(x,y)$:

$$I(X,Y) = \sum_{x,y} p(x,y)log(\frac{p(x,y)}{p(x)p(y)})$$

Here the logarithm is vanishing whenever $p(x,y) = p(x)p(y)$ and since this condition occurs for statistically independent random variables, in this circumstance the random variables X and Y possess zero bits of mutual information. In other words, the knowledge of X does not provide any information about Y when the two random variables are statistically independent.

1.5 Conditional Mutual Information

The last classical entropic quantity we present in this section is the *Conditional Mutual Information*, which quantifies the common information between two random variables X and Y when we already have some side information provided by another random variable Z.

Definition (Conditional Mutual Information)**.** Let X,Y and Z be discrete random variables. The conditional mutual information is defined as follows:

$$\begin{aligned} I(X;Y \mid Z) &= H(Y \mid Z) - H(Y \mid X,Z) \\ &= H(X \mid Z) - H(X \mid Y,Z) \\ &= H(X \mid Z) + H(Y \mid Z) - H(X,Y \mid Z) \end{aligned} \tag{1.3}$$

Let us now consider the following statement. We shall later refer to it in order to shed some light on the departure of quantum information theory from the classical one:

Theorem (Strong Subadditivity)**.** *The conditional mutual information $I(X;Y \mid Z)$ is non-negative:*

$$I(X;Y \mid Z) \geq 0,$$

and the inequality is saturated if and only if X - Z - Y is a Markov Chain (i.e.,if $p(x,y \mid z) = p(x \mid z)p(y \mid z)$).

Proof. The proof follows straightforwardly from the non-negativity of the mutual information. Consider the following equality:

$$I(X;Y \mid Z) = \sum_z p(z)I(X;Y \mid Z = z)$$

where $I(X;Y \mid Z)$ is a mutual information with respect to the joint density $(x, y \mid z)$ and the marginal densities $p(x \mid z)$ and $p(y \mid z)$. Combining non-negativity of both $p(z)$ and $I(X;Y \mid Z = z)$ we easily obtain the non-negativity of $I(X;Y \mid Z)$, with the saturation condition fulfilled in the presence of saturation of $I(X;Y \mid Z = z)$(considering that the conditional mutual information is a convex combination of mutual information). □

Chapter 2

Classical vs Quantum Information Theory

You have nothing to do but mention the quantum theory, and people will take your voice for the voice of science, and believe anything
George Bernard Shaw

Shannon's contribution to classical information theory is heralded as one of the single greatest achievements in modern science, albeit it is not sufficient for the purpose we are chasing, because it is now necessary to delve into the convergence of information theory and quantum theory. In principle, this task is seemingly overwhelming, since it is very difficult to understand the quantum theory intuitively and the phenomena it predicts are not amenable to daily experience. Nevertheless, we shall limit ourselves to rephrase many previous concepts, enlightening their differences and common traits within the quantum scenario with respect to the classical one. Driven by this motivation, we take into account some useful information measures for quantifying the amount of information within composite quantum systems and their correlations.

2.1 The Von Neumann Entropy

We pointed out that Shannon entropy refers to the uncertainty associated with a classical probability distribution. In a similar fashion, we can endow the Von Neumann entropy of the same meaning, with the caveat that quantum states are described by density operators rather than probability distributions. Since density operators settle the probabilities for the measurement outcomes of any system and capture the notion of uncertainty arising from the uncertainty principle, it is reasonable to express the quantum measure of uncertainty as a function of density matrices. Therefore, it is customary to represent the *Von Neumann entropy* associated to a quantum system described by its own density matrix as follows:

$$S(\rho) = -Tr(\rho \log \rho)$$

As classical entropy gives a precise meaning to the notion of `information bit`, the quantum entropy assigns a meaning to the `information qubit`. The latter is a rather peculiar concept and despite it seems similar to that of physical qubit, it is instead quite different: while the physical qubit describes a two-level quantum state belonging to any quantum object (such as a photon or an electron), the information qubit reveals the amount of information which is present in a quantum system. However, classical and quantum measures of uncertainty are not always far one away from the other. It turns out that the former is a special case of the latter. Consider the following example.

Source of Orthogonal pure states

Let us suppose to deal with two orthogonal pure states for a qubit, namely $|0\rangle$ and $|1\rangle$, providing a basis for the single qubit Hilbert space. Assuming that the states $|0\rangle$ or $|1\rangle$ occur with probabilities $p_0 = p$ and $p_1 = 1 - p$, and given their relevant density matrices $\rho_0 = |0\rangle\langle 0|$ and $\rho_1 = |1\rangle\langle 1|$, we can write

$$\rho = p_0\ |0\rangle\langle 0| + p_1\ |1\rangle\langle 1| = \begin{bmatrix} p_0 & 0 \\ 0 & p_1 \end{bmatrix}$$

Hence, the Von Neumann entropy is readily computed :

$$\begin{aligned} S(\rho) &= -Tr(\rho \, log \, \rho) \\ &= -Tr\left(\begin{bmatrix} p_0 & 0 \\ 0 & p_1 \end{bmatrix} \begin{bmatrix} \log\ p_0 & 0 \\ 0 & \log\ p_1 \end{bmatrix}\right) \\ &= -p_0 \log p_0 - p_1 \log p_1 = H(p_0, p_1) \end{aligned} \tag{2.1}$$

This example showed that in the simplest case of orthogonal pure quantum states, the situation is classical, from the perspective of information theory (it is reasonable since orthogonal pure states are maximally distinguishable); thus, in this case the Von Neumann entropy is the Shannon entropy in disguise. As final remark of this section, we mention these two important properties for quantum entropy, some of them being useful for the rest of the discussion:

- The entropy is non-negative. It is zero if and only if the state is pure;
- Suppose a composite system AB is in a pure state. Then S(A) = S(B);
- The Von Neumann entropy is invariant under unitary temporal evolution. Indeed, $S(\rho)$ depends only on the eigenvalues of ρ, which are basis-independent.Therefore, $S(U\rho U^{\dagger}) = S(\rho)$;

By analogy with the classical case, it is possible to define quantum joint and conditional entropies and quantum mutual information for many-body quantum systems.

2.2 Joint Quantum Entropy

Given the density operator associated to a bipartite system AB, i.e. $\rho_{AB} \in \mathcal{D}\{\mathcal{H_A} \otimes \mathcal{H_B}\}$, the *Joint quantum entropy* is defined as

$$S(AB) = -Tr\{\rho_{AB} \log \rho_{AB}\}.$$

If ρ_{ABC} is a tripartite system, i.e. $\rho_{ABC} \in \mathcal{D}\{\mathcal{H_A} \otimes \mathcal{H_B} \otimes \mathcal{H_C}\}$, the entropy S(AB) is defined in the same manner, with $\rho_{AB} = Tr_C\{\rho_{ABC}\}$. The Joint quantum entropy allows us to point out the first critical departure of quantum theory from the classical world. Recall that the following inequalities for classical entropy

$$H(X,Y) \geq H(X), \qquad H(X,Y) \geq H(Y)$$

hold rather intuitively because they state that we have more uncertainty about the joint state of X and Y than that of X alone. A major stumbling block which prevents the quantum theory from being understood via common sense is represented by the fact that this intuition fails for quantum states. Let us consider a purely entangled bipartite system AB, whose state is described by an *EPR pair*:

$$|\beta_{00}\rangle = \left(|00\rangle + |11\rangle\right) / \sqrt{2}$$

This is a pure state, hence S(A,B) = 0 by a previous property. However, the marginal state on subsystem A is the maximally mixed state, because its density operator is $\frac{I}{2}$, thus its entropy in unitary. Indeed, the above *Bell state* has density operator

$$\begin{aligned} \rho &= \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)\left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) \\ &= \frac{|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|}{2} \end{aligned} \tag{2.2}$$

After tracing out the second qubit we easily find the reduced density operator of the first qubit:

$$\begin{aligned}
\rho^1 &= tr_2\,(\rho) \\
&= \frac{tr_2(|00\rangle\,\langle 00|) + tr_2(|11\rangle\,\langle 00|) + tr_2(|00\rangle\,\langle 11|) + tr_2(|11\rangle\,\langle 11|)}{2} \\
&= \frac{|0\rangle\,\langle 0|\,\langle 0|0\rangle + |1\rangle\,\langle 0|\,\langle 0|1\rangle + |0\rangle\,\langle 1|\,\langle 1|0\rangle + |1\rangle\,\langle 1|\,\langle 1|1\rangle}{2} \\
&= \frac{|0\rangle\,\langle 0| + |1\rangle\,\langle 1|}{2} \\
&= \frac{I}{2}.
\end{aligned} \tag{2.3}$$

2.3 Conditional Quantum Entropy

The most natural definition of *Conditional quantum entropy* and also the most useful one in quantum information theory is the following simple one.

Definition (Conditional Quantum Entropy)**.** Let $\rho_{AB} \in \mathcal{D}\{\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}\}$. The conditional quantum entropy $S(A \mid B)$ of ρ_{AB} is equal to the difference of the joint quantum entropy S(A,B) and the marginal entropy S(B) :

$$S(A \mid B) \equiv S(A, B) - S(B)$$

By virtue of this definition, we can rephrase alternatively the result of the previous subsection: for an *entangled state*, such as a Bell state, the conditional quantum entropy can be negative. In particular, recall that for the marginal state we found

$$\rho^1 = tr_2(\rho) = tr_2\Big(\Big(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\Big)\Big(\frac{\langle 11| + \langle 00|}{\sqrt{2}}\Big)\Big) = \frac{I}{2}$$

and since $tr\Big(\big(\frac{I}{2}\big)^2\Big) = \frac{1}{2} < 1$ and the joint state of the entangled system is a pure state $\big(S(A, B) = 0\big)$, the quantity

$$S(B \mid A) = S(A, B) - S(A)$$

is negative. Stated otherwise, the joint state of the EPR pair is known exactly because its associated entropy is vanishing, but being the first qubit in a mixed state, i.e. a state about which we do not have maximal knowledge, the marginal entropy is non-vanishing. This paradoxical property that the joint state of a system can be completely known, yet we do not have a complete knowledge of its subsystems, is a hallmark of *quantum entanglement* and it is in stark contrast with our common sense, because the only reason that conditional quantum entropy can be negative occurs when we are actually more certain about the joint state of a quantum system than we are about any of its individual parts. It is perhaps the very essence of the radical departure between the classical and quantum world from an informational point of view, definitely defying our intuition.

2.4 Coherent Information

Negativity of Conditional quantum entropy is overarching in the framework of quantum information theory. Its importance is such that it is customary to employ another information quantity in its own right, namely the *Coherent Information*, which we shall refer later in this discussion:

Definition (Coherent Information)**.** The coherent information $I(A > B)$ of a bipartite state $\rho_{AB} \in \mathcal{D}(\mathcal{H_A} \otimes \mathcal{H_B})$ is defined as

$$I(A > B) \equiv S(B) - S(A, B)$$

where the Dirac symbol '>' is present to indicate that this is a quantum information measure having no classical counterpart. Its expression, given as the negative of the conditional quantum entropy, suggests that it is a proper measure of the extent to which we know less about a part of the system than we do about its whole.

2.5 Quantum Mutual Information

Just as the Mutual information measures classical correlations, the *Quantum Mutual Information* is the information quantity that measures both classical and quantum correlations.

Definition (Quantum Mutual Information)**.** The quantum mutual information of a bipartite state $\rho_{AB} \in \mathcal{D}(\mathcal{H_A} \otimes \mathcal{H_B})$ is defined as

$$\begin{aligned} S(A : B) &\equiv S(A) + S(B) - S(A, B) \\ &= S(A) - S(A \mid B) \\ &= S(B) - S(B \mid A) \end{aligned} \tag{2.4}$$

The second line of the above definition leads to the following relations between coherent information and quantum mutual information :

$$\begin{aligned} S(A : B) &= S(A) + I(A > B) \\ &= S(B) + I(B > A). \end{aligned} \tag{2.5}$$

Chapter 3

Approximations to QCMI

It is the mark of an educated mind to rest satisfied with the degree of precision which the nature of the subject admits and not to seek exactness where only an approximation is possible
Aristotle

This section can be conceivably considered the core of the discussion and it copes with the problem outlined at the end of the introduction. Once we stated the formal definition of QCMI, we will deal with its expression made of an algebraic sum of entropic terms in order to get some insights on the bounds we are looking for. Our aim is motivated by the fact that the exact knowledge of the Von Neumann entropy requires the reconstruction of the full spectrum of quantum states, which is clearly a computationally-demanding task for which there is no known algorithmic speed-up, and since much of information theory rests upon quantum entropy's shoulders (we could safely say that it is one of the bedrocks of Quantum Information Theory), as suggested by the fact that it stems from the definition of many important information measures, it urges a clever method to measure entropy efficiently, as well as all the other physical quantities related to it and we are interested in.

As much as the Classical Conditional Mutual Information captures the notion of information between two random variables while retaining additional information gathered by a third random variable, we can assign the same task to the *Quantum Conditional Mutual Information*:

Definition (Quantum Conditional Mutual Information)**.** For any tripartite state $\rho_{ABC} \in \mathcal{D}(\mathcal{H_A} \otimes \mathcal{H_B} \otimes \mathcal{H_C})$, the quantum conditional mutual information is defined as

$$I(A : C \mid B) \equiv S(\rho_{AB}) + S(\rho_{BC}) - S(\rho_B) - S(\rho_{ABC})$$

The QCMI quantifies the correlations established between A and C from the perspective of B. Assuming B is quantum (otherwise $I(A : C|B)$ would be simply the average over the values b taken by B of the mutual information for the conditional state on the system AC, QCMI is significantly less known than its classical counterpart. In particular, its non-negativity property is not obvious at all and deserves particular attention because it

can be thought of as the consequence of a non trivial statement for the Von Neumann entropy, namely the strong subadditivity property.
Before embarking into technical details, it is worth emphasizing that the solution given to bound estimation problem for QCMI will be analogous to the bound estimation problem's solution given for the coherent information, whose expression written below explicitly shows its dependence on the Von Neumann entropy:

$$I(A > B) = S(\rho_B) - S(\rho_{AB})$$

where A and B are subsystems and $\rho_B = Tr_A(\rho_{AB})$ is the reduced density operator of subsystem B. As suggested by the definition, the exact knowledge of $I(A > B)$ requires to know the spectrum reconstruction of the system state in order to compute the von Neumann entropy. However, here we address a remarkable strategy based on Lagrange multipliers that allows us to infer conceivable upper and lower bounds on the quantum entropy in terms of global and marginal purity of the state, defined as the trace of the squared density matrix:

$$\mathcal{P}(\rho) \equiv \sum_{i=1}^{d} \lambda_{i,\rho}^2$$

where the sum runs up to d, i.e. the dimension of the quantum state which admits a Schmidt decomposition

$$\rho = \sum_{i=1}^{d} \lambda_{i,\rho} \left|\psi_i\right\rangle \left\langle\psi_i\right|$$

where λ_i are non-negative real numbers satisfying

$$\lambda_{1,\rho} \geq \lambda_{2,\rho} \geq \cdots \geq \lambda_{d,\rho}, \qquad \sum \lambda_{i,\rho} = 1$$

and $\left|\psi_i\right\rangle, \left|\psi_j\right\rangle$ are two vectors drawn from an orthogonal basis fulfilling

$$\langle\psi_i|\psi_j\rangle = \delta_{i,j} \; = \begin{cases} 1 & if \quad i = j \\ 0 & otherwise \end{cases}$$

Disregarding the full derivation, which is faithfully reported in 4.3 it can be shown that the spectrum $\{\lambda_{i,\rho}^M\}$ that maximizes

$$S(\rho) = -\sum_{i=1}^{d} \lambda_{i,\rho} \log \lambda_{i,\rho}$$

is given by

$$\lambda_{1,\rho}^M = \frac{1}{d} + \sqrt{\frac{d-1}{d}\Big(\mathcal{P}(\rho) - \frac{1}{d}\Big)}, \quad \lambda_{2,\rho}^M = \cdots = \lambda_{d,\rho}^M = \frac{1 - \lambda_{1,\rho}^M}{d-1}$$

whereas the spectrum $\{\lambda_{i,\rho}^m\}$ that minimizes the entropy $S(\rho)$ is given by

$$\lambda_{1,\rho}^m = \lambda_{2,\rho}^m = \cdots = \lambda_{k_\rho - 1,\rho} = \frac{1-\alpha_\rho}{k_\rho - 1}, \quad \lambda_{k_\rho,\rho}^m = \alpha_\rho,\, \lambda_{k_\rho+1,\rho}^m = \cdots = \lambda_{d,\rho}^m = 0,$$

where

$$\alpha_\rho = 1/k_\rho - \sqrt{\Big(1 - \frac{1}{k_\rho}\Big)\Big(\mathcal{P}(\rho) - \frac{1}{k_\rho}\Big)}$$

and k_ρ is the integer such that

$$\frac{1}{k_\rho} \leq \mathcal{P}(\rho) < \frac{1}{k_\rho - 1}$$

Provided the vectors solving the minimization and the maximization, that are $\{\lambda_{i,\rho}^m\}, \{\lambda_{i,\rho}^M\}$, respectively, by minimizing (maximizing) the marginal purity on subsystem B and maximizing (minimizing) the global purity, one has lower(upper) bounds to the Coherent Information :

$$\ell_e(\rho_{AB}) \leq I(A > B) \leq u_e(\rho_{AB}),$$

where

$$\begin{aligned} \ell_e(\rho_{AB}) = &\Big(\lambda_{k_{\rho_B},\rho_B} - 1\Big) \log \lambda_{1,\rho_B}^m - \lambda_{k_{\rho_B},\rho_B} \log \lambda_{k_{\rho_B},\rho_B}^m \\ &+ \Big(1 - \lambda_{1,\rho_{AB}}^M\Big) \log \frac{1 - \lambda_{1,\rho_{AB}}^M}{d-1} + \lambda_{1,\rho_{AB}}^M \log \lambda_{1,\rho_{AB}^M} \end{aligned} \tag{3.1}$$

and

$$\begin{aligned} u_e(\rho_{AB}) = &\Big(1 - \lambda_{k_{\rho_{AB}},\rho_{AB}}^m\Big) \log \lambda_{1,\rho_{AB}}^m + \lambda_{k_{\rho_{AB}},\rho_{AB}}^m \log \lambda_{k_{\rho_{AB}},\rho_{AB}}^m \\ &- \Big(1 - \lambda_{1,\rho_{AB}}^M\Big) \log \frac{1 - \lambda_{1,\rho_B}^M}{d_B - 1} - \lambda_{1,\rho_B}^M \log \lambda_{1,\rho_B}^M \end{aligned} \tag{3.2}$$

Quantitative bounds to the conditional mutual information in terms of purity functions can be obtained by a straightforward generalization of this method. Given the spectra $\{\lambda_{i,\rho}^M\}$ and $\{\lambda_{i,\rho}^m\}$ maximizing and minimizing $S(\rho)$, respectively, by minimizing (maximizing) the first two terms and maximizing (minimizing) the third and the fourth term of the expression for $I(A : C \mid B)$, one can obtain lower (upper) bounds to the conditional quantum mutual information. Indeed, it is bounded as follows :

$$\ell_I(\rho_{ABC}) \leq I(A : C \mid B) \leq u_I(\rho_{ABC}),$$

where

$$
\begin{aligned}
\ell_I(\rho_{ABC}) = &\Big(\lambda_{k_{\rho_{AB}},\rho_{AB}}^{m} - 1\Big)\,\log \lambda_{1,\rho_{AB}}^{m} - \lambda_{k_{\rho_{AB}},\rho_{AB}}^{m}\,\log \lambda_{k_{\rho_{AB}},\rho_{AB}}^{m} \\
&+ \Big(\lambda_{k_{\rho_{BC}},\rho_{BC}}^{m} - 1\Big)\,\log \lambda_{1,\rho_{BC}}^{m} - \lambda_{k_{\rho_{BC}},\rho_{BC}}^{m}\,\log \lambda_{k_{\rho_{BC}},\rho_{BC}}^{m} \\
&+ \Big(1 - \lambda_{1,\rho_B}^{M}\Big)\,\log \frac{1-\lambda_{1,\rho_B}^{M}}{d_B - 1} + \lambda_{1,\rho_B}^{M}\,\log \lambda_{1,\rho_B}^{M} + \Big(1 - \lambda_{1,\rho_{ABC}}^{M}\Big)\log\Big(\frac{1-\lambda_{1,\rho_{ABC}}^{M}}{d_{ABC} - 1}\Big) \\
&+ \lambda_{1,\rho_{ABC}}^{M}\,\log \lambda_{1,\rho_{ABC}}^{M}
\end{aligned} \tag{3.3}
$$

and

$$
\begin{aligned}
u_I(\rho_{ABC}) = &\Big(\lambda_{1,\rho_{AB}}^{M} - 1\Big)\,\log \frac{1-\lambda_{1,\rho_{AB}}^{M}}{d_{AB} - 1} - \lambda_{1,\rho_{AB}}^{M}\,\log \lambda_{1,\rho_{AB}}^{M} \\
&+ \Big(\lambda_{1,\rho_{BC}}^{M} - 1\Big)\,\log \frac{1-\lambda_{1,\rho_{BC}}^{M}}{d_{BC} - 1} - \lambda_{1,\rho_{BC}}^{M}\,\log \lambda_{1,\rho_{BC}}^{M} \\
&+ \Big(1 - \lambda_{k_{\rho_B},\rho_B}^{m}\Big)\,\log \lambda_{1,\rho_B}^{m} + \lambda_{k_{\rho_B},\rho_B}^{m}\,\log \lambda_{k_{\rho_B},\rho_B}^{m} \\
&+ \Big(1 - \lambda_{k_{\rho_{ABC}},\rho_{ABC}}^{m}\Big)\,\log \lambda_{1,\rho_{ABC}}^{m} + \lambda_{k_{\rho_{ABC}},\rho_{ABC}}^{m}\,\log \lambda_{k_{\rho_{ABC}},\rho_{ABC}}^{m}.
\end{aligned} \tag{3.4}
$$

In the next section, we delve into the problem of showing that $\ell_I(\rho_{ABC})$ constitutes a suitable refinement for another lower bound to the QCMI.

Chapter 4

Further refinements to QCMI's bounds

A mathematical problem should be difficult in order to entice us, yet not completely inaccessible, lest it mock at our efforts. It should be to us a guidepost on the mazy paths to hidden truths, and ultimately a reminder of our pleasure in the successful solution
David Hilbert

The Strong Subadditivity Inequality for Quantum Entropy, which is the statement that for all tripartite states $\rho_{123} \in \mathcal{D}(H_1 \otimes H_2 \otimes H_3)$, the *Conditional Mutual Information of 1 and 2 given 3* is non-negative,

$$I(1,2 \mid 3) \equiv S_{13} + S_{23} - S_{123} - S_3 \geq 0, \tag{4.1}$$

sets naturally a lower bound to QCMI. However, the clarity of this result should not beguile us into false complacency: indeed, RHS of equation 4.1 does not represent the only lower bound existing, as many tighter lower bounds are possible too. One of them is a nice result by Eric A. Carlen and Elliot H. Lieb (see also [9]), who obtained a lower bound which follows, in a relatively simple manner, from strong subadditivity and it is dubbed **Extended Strong Subadditivity**. The main thrust of this chapter is to demonstrate that our bound 3.3 is also tighter than the Carlen and Lieb's one and we will prove its sharpness on a particular case study.

4.1 A fundamental bound to QCMI

Recall that the irrefutable positivity of the conditional entropies in classical probability theory, namely

$$S_{12} - S_1 \geq 0 \quad \textit{and} \quad S_{12} - S_2 \geq 0$$

has no analog in quantum mechanics, where entanglement ensures that the failure of either one of these inequalities can occur. This sentence turns into the sharp inequality

$$E \geq max\{S_1 - S_{12}, S_2 - S_{12}, 0\},$$

where E denotes either E_f or E_{sq} (take a look at [7]), the latter (SQUASHED ENTANGLEMENT) being a non-negative minorant of the former (FAITHFUL ENTANGLEMENT), defined by

$$E_{sq}(\rho_{12}) = \frac{1}{2}\inf\{I(1,2 \mid 3) : \rho_{123}\,\text{is } \textit{any} \text{ tripartite extension of } \rho_{12}\} \quad (4.2)$$

Equation 4.2 immediately implies an extension of 4.1, as dictated by the following result:

Theorem (Extended strong subadditivity). *For all tripartite states ρ_{123},*

$$I(1,2 \mid 3) \geq 2\,max\{S_1 - S_{12}, S_2 - S_{12}, 0\}.$$

The inequality $I(1,2 \mid 3) \geq \lambda max\{S_1 - S_{12},\, S_2 - S_{12},\, 0\}$ can be violated for all $\lambda > 2$.

The connection between this section and its sibling, the next one, occurs with the proof of the above statement, as the final conclusion is centered on it.

Proof. The theorem resorts to a widely used tool in Quantum Information, i.e. PURIFICATION (see section on pag.35) and to the following useful result (already mentioned in section 2.1 on page 15): the marginal entropies S(A) and S(B) of a pure bipartite state $|\phi\rangle$ are equal:

$$S(A)_\phi = S(B)_\phi, \quad (4.3)$$

while the joint entropy $S(AB)$ vanishes:

$$S(AB)_\phi = 0.$$

This statement also applies to several systems, provided a bipartite cut of them. Therefore, the following equations (as many others obtained by permuting subscripts) hold for the state $|\psi\rangle_{ABCDE}$:

$$S(A)_\psi = S(BCDE)_\psi \quad (4.4a)$$
$$S(AB)_\psi = S(CDE)_\psi \quad (4.4b)$$
$$S(ABC)_\psi = S(DE)_\psi \quad (4.4c)$$
$$S(ABCD)_\psi = S(E)_\psi \quad (4.4d)$$

This fact brings us to the right stance for building the proof. Consider any purification ρ_{1234} of ρ_{123}. Being ρ_{1234} pure, then $S_{23} = S_{14}$ and $S_{123} = S_4$, by virtue of equation 4.3. As a consequence, we can set the equality

$$S_{12} + S_{23} - S_1 - S_3 = S_{12} + S_{14} - S_{124} - S_1, \quad (4.5)$$

proving that $S_{12}+S_{23} \geq S_1+S_3$, because the right hand side of equation 4.5 is non-negative by 4.1. By adding $S_{12}+S_{23} \geq S_1+S_3$ and $S_{13}+S_{23} \geq S_1+S_2$ we obtain

$$S_{12} + S_{13} + 2\,S_{23} \geq 2\,S_1 + S_2 + S_3 \tag{4.6}$$

At the end, by using $S_{12} = S_{34}$, $S_{23} = S_{14}$ and $S_2 = S_{134}$ and rearranging terms in equation 4.6, we can write

$$S_{13} + S_{34} - S_{134} - S_3 \geq 2\,(S_1 - S_{14}),$$

which is nothing but the same equation in the theorem with different indices.

□

4.2 A case study

The importance of this theorem lies in its dual effectiveness: first of all, it allows to limit the Quantum Conditional Mutual Information from below, which notably involves at least three subsystems, by means of a function that ostensibly depends only on two subsystems, disregarding the third. Secondly, it undoubtedly constitutes a step forward in improving the lower bound established by 4.1, as it admits a quantity that can be potentially greater than zero. This premise was necessary to tackle the actual problem, that is to plot the upper and lower bounds of $I(A:B\mid C)$ as the probability $p \in [0,1]$ varies for the following three-qubit mixed state:

$$\begin{aligned}\rho &= p\,\rho_W + (1-p)(\rho_{EPR}\otimes\rho_1)\\ &= p\,|W\rangle\langle W| + (1-p)\Big(|\psi_{EPR}\rangle\langle\psi_{EPR}|\otimes|1\rangle\langle 1|\Big)\end{aligned} \tag{4.7}$$

In this case,

$$|\mathbf{W}\rangle = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ 0 \\ \frac{1}{\sqrt{3}} \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

shares entanglement among the three subsystems A,B and C, whereas the second term encloses a state where C is not correlated with AB. It consists of

$$|\psi_{\mathbf{EPR}}\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

also dubbed *Bell state* or *EPR pair* and of the *one-state*:

$$|\mathbf{1}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Since the dimension of any possible n-qubit subsystem is given by 2^n, we have:

$$d_\rho = 8$$
$$d_{\rho_{AB}} = d_{\rho_{BC}} = d_{\rho_{AC}} = 4$$
$$d_{\rho_A} = d_{\rho_B} = d_{\rho_C} = 2$$

Now the idea is to raise the stakes, showing that lower bound 3.3 is tighter than the Carlen and Lieb's one, thus improving their *Extended Strong Subadditivity*. That said, it was relatively simple to draw the figure showing the behaviors of both the lower bound obtained by the two physicists and the QCMI itself, as provided in the output of the first code at pag.36:

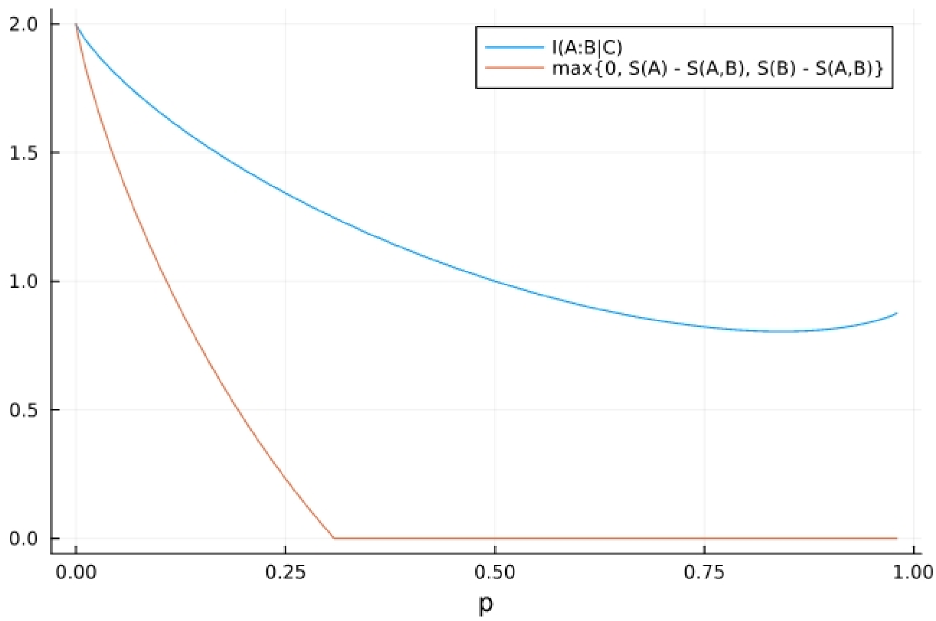

Figure 4.1: QCMI and its lower bound provided by the Extended strong subadditivity.

Furthermore, for the sake of completeness, we cannot exempt ourselves from representing all the quantities involved, in order to exhibit their mutual relationships, as the following Julia figure attempts to supply:

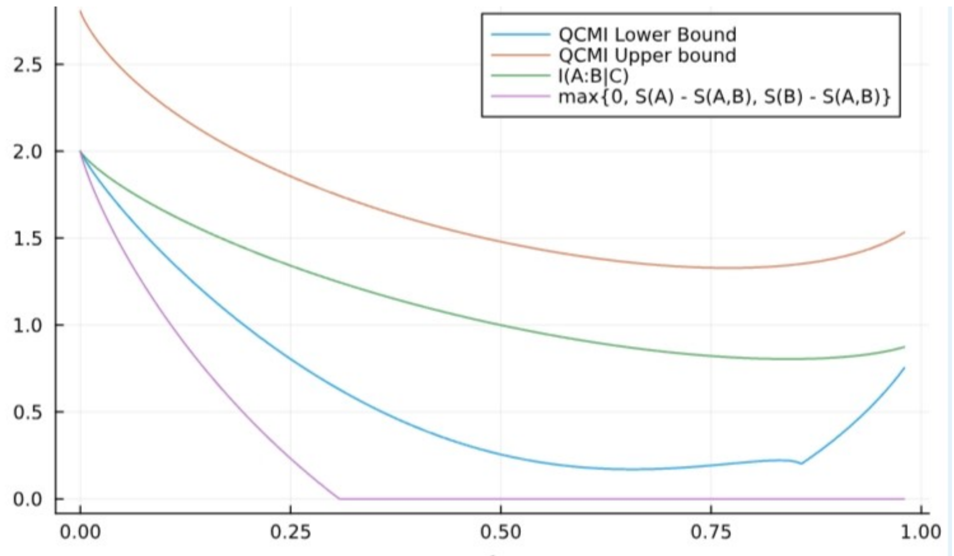

Figure 4.2: Joint representation of $\ell_{I(A:B|C)}$, $u_{I(A:B|C)}$, $I(A : B \mid C)$ and the Carlen-Lieb lower bound for the density operator $\rho = p\,\rho_W + (1-p)(\rho_{EPR} \otimes \rho_1)$ as p varies in $[0, 1]$.

4.3 Theoretical justifications

We cannot say to get all things done yet, until an explanation of the relation between $I(A : B \mid C)$ and the lower bound refinement by Carlen and Lieb is given. This topic deserves more attention, so in this section we delve into the problem of finding a justification of what figure 4.1 shows. Let us focus on the lower bound: the function decreases as $p \in [0, 1]$ increases or, equivalently, when the correlations supplied by the kronecker product between ρ_{EPR} and ρ_1 exceed the entanglement provided by the W-state, for the qubits A and B. It seems that the more the system tends to be separable in its components, the closer the QCMI gets to zero; conversely, the more entangled a system is, the further the QCMI moves away from vanishing values. In other words, we can say that the Quantum Conditional Mutual Information corresponds to the amount by which Strong Subadditivity of quantum entropy fails to be saturated.
Our observations can be justified theoretically because they agree with a widely recognized result involving the separability of quantum states and Quantum Conditional Mutual Information. It can be shown (see [6] for further details) that *Quantum Mutual Information*

$$I(A : B) = S(A) + S(B) - S(AB)$$

obeys to the inequality

$$I(A:B)_\rho \geq \frac{1}{2\log 2} \quad \|\rho_{AB} - \rho_A \otimes \rho_B\|^2 \tag{4.8}$$

Thus, a bipartite state has a zero Mutual Information if and only if it has no correlations (i.e. it is a product state). As for Mutual Information, one can ask which states have zero Conditional Mutual Information. A state ρ_{ABC} has $I(A:B\mid C)=0$ if and only if it is a ***Quantum Markov Chain***, i.e. it admits a decomposition of the C system vector space

$$\mathcal{H}_\mathcal{C} = \bigoplus_j \mathcal{H}_{C_j^L} \otimes \mathcal{H}_{C_j^R}$$

into a direct sum of tensor products such that

$$\rho_{ABC} = \sum_j p_j \rho_{AC_j^L} \otimes \rho_{BC_j^R}$$

with states $\rho_{AC_j^L}$ on $\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_{C_j^L}, \rho_{BC_j^R}$ on $\mathcal{H}_\mathcal{B} \otimes \mathcal{H}_{C_j^R}$ and probabilities p_j [5]. There is an equality analogous to 4.8 for the Conditional Mutual Information, whose right hand side is related to the lower bound of *Squashed Entanglement*, in terms of a suitable distance to separable states. This distance to the set of separable states is measured in terms of the *one-way LOCC*: by analogy with the definition of the trace norm as the optimal probability of distinguishing two quantum states, the following result takes into account a norm that quantifies the distinguishability of quantum states under measurements that are restricted by local operations and one-way classical communication. Writing the distance from a state ρ_{AB} to the set $\mathcal{S}_{A:B}$ of separable states on $A:B$ as

$$\|\rho_{AB} - \mathcal{S}_{A:B}\| = min_{\sigma\in\mathcal{S}_{A:B}} \|\rho_{AB} - \sigma\|$$

we can write for the Squashed Entanglement of every state ρ_{AB}

$$E_{sq} \geq \frac{1}{16\log 2} \|\rho_{AB} - \mathcal{S}_{A:B}\|^2_{LOCC\mapsto}, \tag{4.9}$$

where the Squashed Entanglement is defined as follows:

$$E_{sq}(\rho_{A:B}) \equiv inf\,\{\frac{1}{2}\,I(A:B\mid C) : \rho_{ABC}\,\text{is an extension of}\,\rho_{AB}\} \tag{4.10}$$

Since $\|*\|_{LOCC\mapsto}$ is a norm, 4.9 implies the faithfulness of the Squashed Entanglement, i.e. its property of being strictly positive on every entangled state. Moreover, combining 4.9 and 4.10, we can firmly write that for every tripartite finite-dimensional state ρ_{ABC},

$$I(A:B\mid C) \geq \frac{1}{8\log 2} \|\rho_{AB} - \mathcal{S}_{A:B}\|^2_{LOCC\mapsto},$$

thus strengthening 4.1 by relating it to a distance-like entanglement measure and showing that if a tripartite state has a small Conditional Mutual Information, its AB reduction is close to a separable state.

Appendix A

In this section, some detailed derivations concerning useful results belonging to the previous parts are shown.

Derivation of the bounds for Von Neumann entropy

The knowledge of the bounds for the coherent information $I(A > B)$ relies on the following strategy for bounding the quantum entropy of a quantum state ρ in a d-dimensional Hilbert space with a function of the state purity $\mathcal{P}(\rho) \equiv Tr(\rho^2)$ (see also [4]). Given the spectral decomposition of the quantum state, $\rho = \sum_{i=1}^{d} \lambda_i \left|\psi_i\right\rangle \left\langle\psi_i\right|$, where $\{\left|\psi_i\right\rangle\}$ is an orthogonal basis of the d-dimensional Hilbert space, our task is to solve the following variational problem:

$$\max / \min S(\rho) = -\sum_{i=1}^{d} \lambda_i \log(\lambda_i)$$

such that

$$\sum_{i=1}^{d} \lambda_i^2 = \mathcal{P}(\rho)$$
$$\sum_{i=1}^{d} \lambda_i = 1$$
$$0 \leq \lambda_i \leq 1, \forall i$$

Maximization

For $d = 2$, the maximization problem would be trivial. Therefore, let us consider the special case $d = 3$, which can be shown to be straightforwardly generalized to any value of d. Assuming WLOG $\lambda_1 \geq \lambda_2 \geq \lambda_3$, we turn the problem into

$$\max S(\rho) = -\lambda_1 \log \lambda_1 - \lambda_2 \log \lambda_2 - \lambda_3 \log \lambda_3$$

such that

$$\lambda_1^2 + \lambda_2^2 + \lambda_3^2 = \mathcal{P}(\rho)$$
$$\lambda_1 + \lambda_2 + \lambda_3 = 1$$
$$1 \geq \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$$

We prove that the maximum is reached with the following Lemma.

Lemma 1. *The solution to the maximization problem is given by*

$$\lambda_1 = \frac{1}{3} + \sqrt{\frac{2}{3}\Big(\mathcal{P}(\rho) - \frac{1}{3}\Big)} \tag{4.11}$$
$$\lambda_2 = \lambda_3 = \frac{1 - \lambda_1}{2} \tag{4.12}$$

Proof. Differentiating the entropy function $S(\rho)$ and the other constraints one has

$$dS = -(1 + \log\,\lambda_1)d\lambda_1 - (1 + \log\,\lambda_2)d\lambda_2 - (1 + \log\,\lambda_3)d\lambda_3$$

and

$$\lambda_1 d\lambda_1 + \lambda_2 d\lambda_2 + \lambda_3 d\lambda_3 = 0 \tag{4.13}$$
$$d\lambda_1 + d\lambda_2 + d\lambda_3 = 0 \tag{4.14}$$

respectively. A simple algebra immediately leads to

$$d\lambda_1 = -\frac{\lambda_2}{\lambda_1}d\lambda_2 - \frac{\lambda_3}{\lambda_1}d\lambda_3 = -\frac{\lambda_2}{\lambda_1}\Big(-d\lambda_1 - d\lambda_3\Big) - \frac{\lambda_3}{\lambda_1}d\lambda_3 \implies$$
$$d\lambda_1 = \frac{\lambda_2 d\lambda_1 + \lambda_2 d\lambda_3 - \lambda_3 d\lambda_3}{\lambda_1} \implies$$
$$d\lambda_1\Big(\lambda_1 - \lambda_2\Big) = d\lambda_3\Big(\lambda_2 - \lambda_3\Big)$$
$$\implies d\lambda_1 = \frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_2}d\lambda_3$$

and

$$\lambda_2 d\lambda_2 = -\lambda_3 d\lambda_3 + \lambda_1 d\lambda_1 = -\lambda_3 d\lambda_3 - \lambda_1\Big(-d\lambda_2 - d\lambda_3\Big)$$
$$\implies d\lambda_2 = \frac{-\lambda_3 d\lambda_3 + \lambda_1 d\lambda_2 + \lambda_1 d\lambda_3}{\lambda_2} \implies$$
$$d\lambda_2\Big(\lambda_2 - \lambda_1\Big) = d\lambda_3\Big(\lambda_1 - \lambda_3\Big)$$
$$\implies d\lambda_2 = \frac{\lambda_1 - \lambda_3}{\lambda_2 - \lambda_1}d\lambda_3.$$

Plugging these two expressions for $d\lambda_1$ and $d\lambda_2$ into the differential of the entropy we obtain

$$\begin{aligned}
dS(\rho) &= -(1+\log\,\lambda_1)d\lambda_1 - (1+\log\,\lambda_2)d\lambda_2 - (1+\log\,\lambda_3)d\lambda_3 \\
&= -d\lambda_1 - d\lambda_2 - d\lambda_3 + \frac{d\lambda_3}{\lambda_1-\lambda_2}\Big((\lambda_3-\lambda_2)\,\log\lambda_1 + (\lambda_1-\lambda_3)\,\log\lambda_2 + (\lambda_2-\lambda_1)\,\log\lambda_3\Big) \\
&= -\underbrace{(d\lambda_1+d\lambda_2+d\lambda_3)}_{=0} + \frac{d\lambda_3}{\lambda_1-\lambda_2}\Big((\lambda_3-\lambda_2)\,\log\lambda_1 + (\lambda_1-\lambda_3)\,\log\lambda_2 + (\lambda_2-\lambda_1)\,\log\lambda_3\Big) \\
&= \underbrace{(\lambda_2-\lambda_3)}_{\geq 0}\underbrace{\left(\frac{\log\lambda_2-\log\lambda_1}{\lambda_1-\lambda_2} + \frac{\log\lambda_3-\log\lambda_2}{\lambda_3-\lambda_2}\right)}_{\geq 0} d\lambda_3
\end{aligned} \tag{4.15}$$

where the under-braced terms within the last line stem from the concavity of the function $\log\lambda$ for $\lambda \in [0,1]$, which leads to

$$\frac{\log\,\lambda_1 - \log\,\lambda_2}{\lambda_1-\lambda_2} \leq \frac{\log\,\lambda_3 - \log\,\lambda_2}{\lambda_3-\lambda_2}$$

for $\lambda_1 \geq \lambda_2 \geq \lambda_3$. This means that $dS(\rho)/d\lambda_3 \geq 0$ and therefore the maximum of $S(\rho)$ occurs when λ_3 is maximum, i.e. when $\lambda_2 = \lambda_3$. Thus from the constraints we can write

$$\begin{cases}\lambda_1^2+\lambda_2^2+\lambda_3^2 = \mathcal{P}(\rho) \\ \lambda_1+\lambda_2+\lambda_3 = 1\end{cases} \implies$$

$$\begin{cases}\lambda_1^2+2\lambda_2^2 = \mathcal{P}(\rho) \\ \lambda_1+2\lambda_2 = 1\end{cases} \implies$$

$$\begin{cases}\lambda_1^2+2\left(\frac{1-\lambda_1}{2}\right)^2 - \mathcal{P}(\rho) = 0 \\ \lambda_2 = \lambda_3 = \frac{1-\lambda_1}{2}\end{cases} \implies$$

$$\begin{cases}\lambda_1 = \frac{1+\sqrt{2(3\mathcal{P}(\rho)-1)}}{3} = \frac{1}{3} + \sqrt{\frac{2}{3}\Big(\mathcal{P}(\rho)-\frac{1}{3}\Big)} \\ \lambda_2 = \lambda_3 = \frac{1-\lambda_1}{2}\end{cases}$$ □

Rather intuitively, the generalization of these results for the maximization problem to any value of d proceeds in the same manner, as shown below.

Theorem 1. *Suppose that $\lambda_1 \geq \lambda_2 \geq \dots \lambda_d$. The solution to the maximization problem is*

$$\lambda_1 = \frac{1}{d} + \sqrt{\frac{d-1}{d}\Big(\mathcal{P}(\rho) - \frac{1}{d}\Big)} \tag{4.16}$$

$$\lambda_2 = \lambda_3 = \cdots = \lambda_d = \frac{1-\lambda_1}{d-1} \tag{4.17}$$

Proof. The proof is done by contradiction. By absurd, suppose the maximization problem is not provided by $\lambda_2 = \lambda_3 = \cdots = \lambda_d$. We will show that a change in their values would increase $S(\rho)$, thus proving that a configuration for which (3.11) is not true does not provide the maximum entropy. This can be shown by slightly modifying the values of $\lambda_1, \lambda_2, \lambda_d$ while keeping all the other values fixed, i.e $\lambda_3, \lambda_4, \ldots, \lambda_{d-1}$. Setting new constraints for λ_1, λ_2 and λ_d we have

$$\lambda_1^2 + \lambda_2^2 + \lambda_d^2 = a \tag{4.18}$$

$$\lambda_1 + \lambda_2 + \lambda_d = b \tag{4.19}$$

equivalently re-written as follows :

$$\lambda_1'^2 + \lambda_2'^2 + \lambda_d'^2 = a/b^2 \tag{4.20}$$

$$\lambda_1' + \lambda_2' + \lambda_d' = 1 \tag{4.21}$$

upon rescaling $\lambda_1' = \lambda_1/b, \lambda_2' = \lambda_2/b, \lambda_d' = \lambda_d/b$. The entropy function is

$$\begin{aligned} S(\rho) &= -\sum_{i=1}^{d} \lambda_i \log \lambda_i \\ &= S_{1,2,d}(\rho) + S_r(\rho), \end{aligned} \tag{4.22}$$

where $S_{1,2,d}(\rho) = -\lambda_1 \log \lambda_1 - \lambda_2 \log \lambda_2 - \lambda_d \log \lambda_d$ and $S_r(\rho) = -\sum_{i=3}^{d-1} \lambda_i \log \lambda_i$. Since $S_r(\rho)$ is fixed, we need to maximize $S_{1,2,d}(\rho)$, which can also be represented as

$$\begin{aligned} S_{1,2,d}(\rho) &= -b\lambda_1' \log(b\lambda_1') - b\lambda_2' \log(b\lambda_2') - b\lambda_d' \log(b\lambda_d') \\ &= b\Big(-\lambda_1' \log \lambda_1' - \lambda_1' \log b - \lambda_2' \log \lambda_2' - \lambda_2' \log b - \lambda_d' \log \lambda_d' - \lambda_d' \log b\Big) \\ &= b\Big(-\lambda_1' \log \lambda_1' - \lambda_2' \log \lambda_2' - \lambda_d' \log \lambda_d' - \underbrace{(\lambda_1' + \lambda_2' + \lambda_d')}_{=1} \log b\Big) \\ &= b\Big(-\lambda_1' \log \lambda_1' - \lambda_2' \log \lambda_2' - \lambda_d' \log \lambda_d'\Big) - b \log b \end{aligned} \tag{4.23}$$

Denoting $S_{1,2,d}'(\rho) \equiv -\lambda_1' \log \lambda_1' - \lambda_2' \log \lambda_2' - \lambda_d' \log \lambda_d'$, we recover the same optimization problem for $d = 3$, whose solution (the maximum of $S'1, 2, d(\rho)$) is reached when $\lambda_2' = \lambda_d'$, as dictated by Lemma 1. It turns out that the maximum of $S1, 2, d(\rho)$ given the constraints (3.14) and (3.15) is saturated with $\lambda_2 = \lambda_d$, hence contradicting the assumption $\lambda_2 > \lambda_d$. Therefore, the solution to the maximization problem is given by equations (3.10) and (3.11), as we set out to show. □

Minimization

Next we consider the solution to $\min S(\rho)$. As in the maximization problem, consider the special case with $d = 3$ and $\lambda_1 \geq \lambda_2 \geq \lambda_3$,

$$\min S(\rho) = -\lambda_1 \log \lambda_1 - \lambda_2 \log \lambda_2 - \lambda_3 \log \lambda_3$$

such that

$$\begin{gathered} \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = \mathcal{P}(\rho) \\ \lambda_1 + \lambda_2 + \lambda_3 = 1 \\ 1 \geq \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0 \end{gathered}$$

Lemma 2. *The solution to the minimization problem is reached either when* $\lambda_1 = \lambda_2$ *or* $\lambda_3 = 0$

Proof. From the proof of Lemma 1, it is known that $dS(\rho)/d\lambda_3 \geq 0$. Therefore, the lower bound of $S(\rho)$ is reached when λ_3 takes its minimum. A lower bound for λ_3 is readily obtained as follows :

$$\begin{aligned} 2(\lambda_1^2 + \lambda_2^2) &\geq (\lambda_1 + \lambda_2)^2 \\ \implies 2(\mathcal{P}(\rho) - \lambda_3^2) &\geq (1 - \lambda_3)^2 \\ \implies \lambda_3 &\geq \max\{0, \frac{1 - \sqrt{6\mathcal{P}(\rho) - 2}}{3}\} \end{aligned}$$

Thus, when $\mathcal{P}(\rho) \geq 1/2$, the minimal possible value for λ_3 is 0; when $1/3 \leq \mathcal{P}(\rho) < 1/2$, the minimal possible value for λ_3 is $\frac{1-\sqrt{6\mathcal{P}(\rho)-2}}{3}$ and $\lambda_1 = \lambda_2 = (1 - \lambda_3)/2$. □

The general solution, which encompasses this particular case, follows straightforwardly.

Theorem. *Suppose* $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_k$*, the solution to the minimization problem is*

$$\lambda_1 = \lambda_2 = \cdots = \lambda_{k-1} = \frac{1-\alpha}{k-1}, \tag{4.24}$$

$$\lambda_k = \alpha, \tag{4.25}$$

$$\lambda_{k+1} = \cdots = \lambda_d = 0. \tag{4.26}$$

where

$$\alpha = \frac{1}{k} - \sqrt{(1 - 1/k)(\mathcal{P}(\rho) - 1/k)}$$

and k *is the integer such that* $\frac{1}{k} \leq \mathcal{P}(\rho) \leq \frac{1}{k-1}$.

Proof. By absurd, suppose the solution is not provided by

$$\lambda_1 = \lambda_2 = \cdots = \lambda_{k-1}, \lambda_K, \lambda_{k+1} = \cdots = \lambda_d = 0$$

Then, there must exist three $\lambda_i, \lambda_j, \lambda_k$ such that $\lambda_i > \lambda_j \geq \lambda_k$ and $\lambda_k \neq 0$ obeying to the constraints

$$\lambda_i^2 + \lambda_j^2 + \lambda_k^2 = c \tag{4.27}$$

$$\lambda_i + \lambda_j + \lambda_k = \gamma \tag{4.28}$$

equivalently re-written as

$$\lambda_i'^2 + \lambda_j'^2 + \lambda_k'^2 = c/\gamma^2 \tag{4.29}$$

$$\lambda_i' + \lambda_j' + \lambda_k' = 1 \tag{4.30}$$

upon rescaling $\lambda_i' = \lambda_i/\gamma$, $\lambda_j' = \lambda_j/\gamma$ and $\lambda_k' = \lambda_k/\gamma$. The entropy function is

$$\begin{aligned} S(\rho) &= -\sum_{i=1}^{d} \lambda_i \log \lambda_i \\ &= S_{i,j,k}(\rho) + S_R(\rho), \end{aligned} \tag{4.31}$$

where the entropies $S_{i,j,k}(\rho) \equiv -\lambda_i \log \lambda_i - \lambda_j \log \lambda_j - \lambda_k \log \lambda_k$ and $S_R(\rho) \equiv -\sum_{\ell \neq i,j,k}^{d} \lambda_\ell \log \lambda_\ell$ have been defined. Since $S_R(\rho)$ is fixed, our goal is to minimize $S_{i,j,k}(\rho)$, which can also be represented as

$$\begin{aligned} S_{i,j,k}(\rho) &= -\gamma \lambda_i' \log(\gamma \lambda_i') - \gamma \lambda_j' \log\left(\gamma \lambda_j'\right) - \gamma \lambda_k' \log(\gamma \lambda_k') \\ &= \gamma \underbrace{\left(-\lambda_i' \log \lambda_i' - \lambda_j' \log \lambda_j' - \lambda_k' \log \lambda_k'\right)}_{\equiv S^{\text{resc}}{}_{i,j,k}(\rho)} - \gamma \log \gamma \end{aligned} \tag{4.32}$$

Hence, we recover the optimization problem encountered for $d = 3$, whose solution (the minimum of $S^{\text{resc}}{}_{i,j,k}(\rho) \equiv -\lambda_i' \log \lambda_i' - \lambda_j' \log \lambda_j' - \lambda_k' \log \lambda_k'$) is reached either when $\lambda_i = \lambda_j$ or $\lambda_k = 0$, contradicting the assumption $\lambda_i > \lambda_j \geq \lambda_k$.We are now in the right position to write

$$\begin{aligned} (k-1)\lambda_1^2 + \lambda_k^2 &= \mathcal{P}(\rho) \\ (k-1)\lambda_1 + \lambda_k &= 1 \\ k &\leq d \end{aligned}$$

and show that there is only a single integer value allowed for number k. Indeed,

$$\begin{aligned} k[(k-1)\lambda_1^2 + \lambda_k^2] &\geq [(k-1)\lambda_1 + \lambda_k]^2 \\ &\geq (k-1)[(k-1)\lambda_1^2 + \lambda_k^2] \end{aligned} \tag{4.33}$$

that is

$$\begin{aligned} k\mathcal{P}(\rho) \geq 1 \geq (k-1)\mathcal{P}(\rho) \\ \implies \frac{1}{P(\rho)} \leq k \leq \frac{1}{\mathcal{P}(\rho)+1} \\ \implies \frac{1}{k} \leq \mathcal{P}(\rho) \leq \frac{1}{k-1} \end{aligned}$$

□

Purification

Purification [1, p.109] is a very useful tool in Quantum Information Theory. Suppose to describe a quantum system A by mean of the density matrix ρ_A. It is always possible to introduce a new fictitious system R, through which we define the pure state $|AR\rangle$ for the joint system AR, such that $\rho_A = Tr_R[|AR\rangle \langle AR|]$. In order to show how to apply this procedure to any system A, let us consider a system in the state ρ_A which admits a spectral decomposition $\rho = \sum_i p_i |i_A\rangle\langle i_A|$ and let R be a system having the same space state as A and basis state $|i\rangle_R$. Let us define a pure state for system AR:

$$|AR\rangle = \sum_i \sqrt{p_i}\ |i_A\rangle |i_R\rangle$$

Computing the reduced density operator for subsystem A we recover the density matrix ρ_A :

$$\begin{aligned} Tr_R[|AR\rangle \langle AR|] &= \sum_{i,j} \sqrt{p_i\, p_j}\ |i_A\rangle \langle j_A|\ Tr[|i_R\rangle \langle j_R|] \\ &= \sum_{i,j} \sqrt{p_i\, p_j}\ |i_A\rangle \langle j_A|\ \delta_{ij} \\ &= \sum_i\ p_i\, |i_A\rangle \langle i_A| = \rho_A \end{aligned} \tag{4.34}$$

Hence, Purification is the procedure which allows to build any mixed state, having no maximal knowledge, by using a pure state, having maximal knowledge, by using the partial trace operation.

Appendix B

You say that I pay too much attention to form. Alas! It is like body and soul: form and content to me are one; I don't know what either is without the other
Gustave Flaubert

Despite its subsidiary importance, this section gathers the main scripts, drawn from the whole code, giving the required plots as output. In order to customize my code, I preferred to switch over Python, by slightly modifying Julia language's syntax, because is expected to be easier for handling syntax highlighting.

Excerpts from Julia codes

Carlen-Lieb's lower bound

The following snippet realizes the Carlen-Lieb's lower bound plot for QCMI, plotted with $I(A : B|C)$ itself.

```
import qutip
qutip.about()
from qutip import *
import numpy as np
import matplotlib.pyplot as plt
import math
ψ_EPR = (tensor(basis(2,0), basis(2,0))
+ tensor(basis(2,1),basis(2,1))).unit()
ψ1 = basis(2,0)
ψW = (tensor(basis(2,0), basis(2,0),basis(2,1))
+ tensor(basis(2,0), basis(2,1), basis(2,0))
+ tensor(basis(2,1), basis(2,0), basis(2,0))).unit()
ρ1 = ket2dm(ψ1)
ρ_EPR = ket2dm(ψ_EPR)
ρW = ket2dm(ψW)
ρ_joint = tensor(ρ_EPR,ρ1)

# Range of probability values
p_values  = np.linspace(0, 1, 1000)
# Function to compute the density matrix
```

```
def mixed_state_density(p):
    return p * ρW + (1 - p) * ρ_joint

# Compute the entropies for different values of p
S_A = []
S_B = []
S_C = []
S_AB = []
S_AC = []
S_BC = []
S_ABC = []

for p in p_values:
    # Compute the density operator  and its
    # reduced density matrices
    ρ_tot = mixed_state_density(p)
    ρ_A = ptrace(ρ_tot,0)
    ρ_B = ptrace(ρ_tot,1)
    ρ_C = ptrace(ρ_tot,2)
    ρ_AB = ptrace(ρ_tot,[0,1])
    ρ_AC = ptrace(ρ_tot,[0,2])
    ρ_BC = ptrace(ρ_tot,[1,2])

    # Compute the entropies associated with the
    # density matrices
    S_A.append(entropy_vn(ρ_A,2))
    S_B.append(entropy_vn(ρ_B,2))
    S_C.append(entropy_vn(ρ_C,2))
    S_AB.append(entropy_vn(ρ_AB,2))
    S_AC.append(entropy_vn(ρ_AC,2))
    S_BC.append(entropy_vn(ρ_BC,2))
    S_ABC.append(entropy_vn(ρ_tot,2))

# Calculate the function
2 * max{0, S(A) - S(A,B), S(B) - S(A,B)}
max_function_values =
2 * np.maximum(0, np.array(S_A) - np.array(S_AB),
               np.array(S_B) - np.array(S_AB))

# Range of probability values
p_values  = np.linspace(0, 1, 1000)

# Array storing the values for Quantum Conditional
# Mutual Information I(A:B|C)
```

```
QCMI = np.zeros_like(p_values)

# von Neumann entropy function
def von_Neumann_entropy(rho):
 evals = rho.eigenenergies()
 return - np.sum(np.nan_to_num(evals * np.log2(evals)))

# Function realizing the plot of QCMI's exact value
for i,p in enumerate(p_values):

    ρ_tot = p * ρW + (1 - p) * ρ_joint
    ρ_AC = ρ_tot.ptrace([0,2])
    ρ_BC = ρ_tot.ptrace([1,2])
    ρ_C  = ρ_tot.ptrace([2])

# This first method accomplishing the task of computing
# Quantum entropy either uses a customized function,
# called "von_Neumann_entropy" or relies on the
# function enabled by QuTip, i.e. "entropy_vn":
# 1:I)

    #S_C = von_Neumann_entropy(ρ_C)
    #S_AC = von_Neumann_entropy(ρ_AC)
    #S_BC = von_Neumann_entropy(ρ_BC)
    #S_ABC = von_Neumann_entropy(ρ_tot)

# 1:II)

    #S_C = entropy_vn(ρ_C,2)
    #S_AC = entropy_vn(ρ_AC,2)
    #S_BC = entropy_vn(ρ_BC,2)
    #S_ABC = entropy_vn(ρ_tot,2)

    #QCMI[i] = (S_AC + S_BC - S_C - S_ABC)

# A second method to compute the QCMI relies on its
# expression in terms of Quantum Mutual Information:
# I(A:B|C) = I(A:BC) - I(A:C). Here the function
# "entropy_mutual"
# receives 4 arguments : I) The density matrix for
# composite quantum systems;
#  II) an intger or a list highlighting the selected
# density matrix components;
# III) An integer or a list denoting the components
```

```
#  we want to condition over and IV) the
# base of the logarithm (base{2,math.e}).

#2)
    I_A_BC = entropy_mutual(ρ_tot, 0, [1,2],2)
    I_A_C = entropy_mutual(ρ_AC, 0,1,2)

    QCMI[i] = I_A_BC - I_A_C

# A third method hinges on the computation of the QCMI in terms
# of conditional entropies:
# I(A:BC) = S(AC) + S(BC) - S(ABC). Here the function
# "entropy_conditional" receives 3 arguments:
# I) The density matrix of a composite object;
# II) The selected components for the density matrix we
#  want to condition over and
# III) The base of the logarithm (base{2,math.e})
   # S_A_conditioned_to_C =
   # entropy_conditional(ρ_AC,1, 2)
   # S_B_conditioned_to_C =
   # entropy_conditional(ρ_BC,1, 2)
   # S_AB_conditioned_to_C =
   # entropy_conditional(ρ_tot,2, 2)

# 3)
    #QCMI[i] = ((S_A_conditioned_to_C +
    # S_B_conditioned_to_C - S_AB_conditioned_to_C))

plt.figure(figsize = (10,5))
plt.subplot(2,2,1)
plt.plot(p_values, QCMI)
plt.xlabel('p')
plt.ylabel('I(A:B|C)')
plt.title('QCMI versus p')
plt.grid()

plt.subplot(2,2,2)
plt.plot(p_values, max_function_values,color = 'yellow')
plt.xlabel('p')
plt.ylabel('max{0,S(A) - S(A,B), S(B) - S(A,B)}')
plt.title('Carlen-Lieb Lower Bound ')
plt.grid()
```

Upper-Lower bounds refinements

A tighter lower bound and also an upper bound for QCMI also exist. In order to highlight their relation with Carlen-Lieb's lower bound and the exact value of $I(A:B|C)$, these four quantities are plotted together: the following code accomplishes this task.

```
import qutip
qutip.about()
from qutip import *
import numpy as np
import matplotlib.pyplot as plt
import math

ψ_EPR = (tensor(basis(2,0), basis(2,0))
 + tensor(basis(2,1),basis(2,1))).unit()
# ψ_Bell = ((tensor(basis(2,0), basis(2,1))
#  + tensor(basis(2,1),basis(2,0))).unit())
# ψ0 = basis(2,1)
ψ1 = basis(2,0)
ψW = (tensor(basis(2,0),basis(2,0),basis(2,1))
 + tensor(basis(2,0), basis(2,1), basis(2,0))
 + tensor(basis(2,1), basis(2,0), basis(2,0))).unit()
# ρ0 = ket2dm(ψ0)
ρ1 = ket2dm(ψ1)
# ρ_Bell = ket2dm(ψ_Bell)
ρ_EPR = ket2dm(ψ_EPR)
ρW = ket2dm(ψW)
ρ_joint = tensor(ρ_EPR,ρ1)

# Range of probability values
p_values  = np.linspace(0, 1, 1000)

# Function to compute the density matrix
def mixed_state_density(p):
    return p * ρW + (1 - p) * ρ_joint

# Compute the purity of a given density matrix
def purity(rho):
    return (rho * rho).tr()

# Compute the entropies for different values of p
S_A = []
S_B = []
S_C = []
```

```
S_AB = []
S_AC = []
S_BC = []
S_ABC = []

for p in p_values:
    # Compute the density operator  and its
    # density opeators
    ρ_tot = mixed_state_density(p)
    ρ_A = ptrace(ρ_tot,0)
    ρ_B = ptrace(ρ_tot,1)
    ρ_C = ptrace(ρ_tot,2)
    ρ_AB = ptrace(ρ_tot,[0,1])
    ρ_AC = ptrace(ρ_tot,[0,2])
    ρ_BC = ptrace(ρ_tot,[1,2])

    # Compute the entropies associated with the
    # density matrices
    S_A.append(entropy_vn(ρ_A,2))
    S_B.append(entropy_vn(ρ_B,2))
    S_C.append(entropy_vn(ρ_C,2))
    S_AB.append(entropy_vn(ρ_AB,2))
    S_AC.append(entropy_vn(ρ_AC,2))
    S_BC.append(entropy_vn(ρ_BC,2))
    S_ABC.append(entropy_vn(ρ_tot,2))

# Compute the QCMI
QCMI = (np.array(S_AC) + np.array(S_BC)
     - np.array(S_C) - np.array(S_ABC))

# Calculate the function max{0, S(A) - S(A,B),
#                               S(B) - S(A,B)}
max_function_values =
   2 * np.maximum(0, np.array(S_A) - np.array(S_AB),
       np.array(S_B) - np.array(S_AB))


# Function computing ℓ s.t. 1/ℓ<= P(ρ) < 1/ℓ - 1)
def find_ell(P):
   ℓ = 2
   while 1/ℓ > P:
        ℓ += 1
   return ℓ
```

```
# Function to compute the global state purity for a given p
def global_state_purity(p):
    ρ = p*ρW + (1-p)*ρ_joint
    return purity(ρ)

# Function to compute ρ_C's state purity
# for a given p
def ρ_C_purity(p):
    ρ = p*ρW + (1-p)*ρ_joint
    ρ_C = ptrace(ρ,2)
    return purity(ρ_C)

# Function to compute ρ_AC's state purity for given p
def ρ_AC_purity(p):
    ρ = p*ρW + (1-p) * ρ_joint
    ρ_AC = ptrace(ρ,[0,2])
    return purity(ρ_AC)

# Function to compute ρ_BC's state purity for a given p
def ρ_BC_purity(p):
    ρ = p * ρW + (1 - p) * ρ_joint
    ρ_BC = ptrace(ρ,[1,2])
    return purity(ρ_BC)

# Function to compute λ_k_ABC_m
def lambda_k_ABC(p,ℓ):
    α_ABC = 1/ℓ- np.sqrt
    ((1 - 1/ℓ) * (global_state_purity(p) - 1/ℓ))
    return α_ABC

# Function to compute λ_k_AC_m
def lambda_k_AC(p,ℓ):
    α_AC = 1/ℓ - np.sqrt((1 - 1/ℓ)
        * (ρ_AC_purity(p) - 1/ℓ))
    return α_AC

# Function to compute λ_k_BC_m
def lambda_k_BC(p,ℓ):
    α_BC = 1/ℓ - np.sqrt((1 - 1/ℓ)
        * (ρ_BC_purity(p) - 1/ℓ))
    return α_BC

# Function to compute λ_k_C_m
def lambda_k_C(p,ℓ):
```

```
    α_C = 1/ℓ - np.sqrt((1 - 1/ℓ)
      * (ρ_C_purity(p)- 1/ℓ))
    return α_C

# Function to compute λ_1_ABC_m
def lambda_1_ABC(p,ℓ):
    α = lambda_k_ABC(p,)
    return (1 - α)/(ℓ - 1)

# Function to compute _1_AC_m
def lambda_1_AC(p,ℓ):
    β = lambda_k_AC(p,ℓ)
    return (1 - β)/(ℓ - 1)

# Function to compute λ_1_BC_m
def lambda_1_BC(p,ℓ):
    γ = lambda_k_BC(p,ℓ)
    return (1 - γ)/(ℓ - 1)

# Function to compute λ_1_C_m
def lambda_1_C(p,ℓ):
    δ = lambda_k_C(p,ℓ)
    return (1 - δ)/(ℓ - 1)

# Function to compute Λ_ABC_M
def Lambda_ABC(p,d):
    return 1/d + np.sqrt((d - 1)/d
       * (global_state_purity(p) - 1/d))

# Function to compute Λ_AC_M
def Lambda_AC(p,d):
    return 1/d + np.sqrt((d - 1)/d
       * (ρ_AC_purity(p) - 1/d))

# Function to compute Λ_BC_M
def Lambda_BC(p,d):
    return 1/d + np.sqrt((d - 1)/d
       * (ρ_BC_purity(p) - 1/d))

# Function to compute Λ_C_M
def Lambda_C(p,d):
    return 1/d + np.sqrt((d-1)/d
       * (ρ_C_purity(p) - 1/d))
```

```python
# Overall Lower Bound
def lower_bound(p):
    ρ = p * ρW + (1 - p) * ρ_joint
    ℓ_AC = find_ell(ρ_AC_purity(p))
    ℓ_BC = find_ell(ρ_BC_purity(p))
    rho_C = ptrace(ρ,2)
    d_C = ρ_C.shape[0]
    d_ABC = ρ.shape[0]
    λ_k_AC_m = lambda_k_AC(p,ℓ_AC)
    λ_k_BC_m = lambda_k_BC(p,ℓ_BC)
    λ_1_AC_m = lambda_1_AC(p,ℓ_AC)
    λ_1_BC_m = lambda_1_BC(p,ℓ_BC)
    Λ_C_M = Lambda_C(p,d_C)
    Λ_ABC_M = Lambda_ABC(p,d_ABC)

    ℓ_I = ((λ_k_AC_m - 1) * np.log2(λ_1_AC_m)
                 - λ_k_AC_m * np.log2(λ_k_AC_m)
                 + (λ_k_BC_m - 1) * np.log2(λ_1_BC_m)
                 - λ_k_BC_m * np.log2(λ_k_BC_m)
                 + (1 - Λ_ABC_M)
                   * np.log2((1 - Λ_ABC_M)/(d_ABC - 1))
                 + Λ_ABC_M * np.log2(Λ_ABC_M)
                 + (1-Λ_C_M)*
                 np.log2((1 - Λ_C_M)/(d_C - 1))
                 + Λ_C_M * np.log2(Λ_C_M))

    return ℓ_I

# Overall Upper bound
def upper_bound(p):
    ρ = p * ρW + (1 - p) * ρ_joint
    ℓ_ABC = find_ell(global_state_purity(p))
    ℓ_C = find_ell(ρ_C_purity(p))
    ρ_BC = ptrace(ρ,[1,2])
    ρ_AC = ptrace(ρ,[0,2])
    d_AC = ρ_AC.shape[0]
    d_BC = ρ_BC.shape[0]
    λ_k_C_m = lambda_k_C(p,ℓ_C)
    λ_k_ABC_m = lambda_k_ABC(p,ℓ_ABC)
    λ_1_C_m = lambda_1_C(p,ℓ_C)
    λ_1_ABC_m = lambda_1_ABC(p,ℓ_ABC)
    Λ_AC_M = Lambda_AC(p,d_AC)
    Λ_BC_M = Lambda_BC(p,d_BC)
```

```
    u_I = ((Λ_AC_M - 1)
       * np.log2((1 - Λ_AC_M)/(d_AC - 1))
         - Λ_AC_M * np.log2(Λ_AC_M)
         + (Λ_BC_M-1) * np.log2((1 - Λ_BC_M)/(d_BC-1))
         - Λ_BC_M * np.log2(Λ_BC_M)
         + (1 - λ_k_C_m) *np.log2(λ_1_C_m)
         + λ_k_C_m * np.log2(λ_k_C_m)
         + (1 - λ_k_ABC_m) * np.log2(λ_1_ABC_m)
         + λ_k_ABC_m * np.log2(λ_k_ABC_m))

    return u_I

# Range of probability values
p_values  = np.linspace(0, 1, 1000)

lowerbound_values = [lower_bound(p) for p in p_values]

upperbound_values = [upper_bound(p) for p in p_values]

plt.figure(figsize = (10,5))
plt.subplot(2,2,1)
plt.plot(p_values, QCMI)
plt.xlabel('p')
plt.ylabel('I(A:B|C)')
plt.title('QCMI versus p')
plt.grid()

plt.subplot(2,2,2)
plt.plot(p_values,max_function_values,color='yellow')
plt.xlabel('p')
plt.ylabel(' 2 * max{0,S(A) - S(A,B), S(B) - S(A,B)}')
plt.title('Carlen-Lieb Lower Bound ')
plt.grid()

plt.subplot(2,2,3)
plt.plot(p_values, lowerbound_values)
plt.xlabel("p")
plt.ylabel("|$\ell$|_I(A:B|C)")
plt.title('QCMI Lower Bound')
plt.grid()
plt.show()

plt.subplot(2,2,4)
plt.plot(p_values, upperbound_values)
```

```
plt.xlabel("p")
plt.ylabel("u_I(A:B|C)")
plt.title('QCMI Upper Bound')
plt.grid()
plt.show()
```

Bibliography

[1] Michael A. Nielsen and Isaac L. Chuang, *Quantum Computation and Quantum Information*,Cambridge University Press (2010).

[2] Giuliano Benenti, Giulio Casati and Giuliano Strini, *Principles of Quantum Computation and Information, Vol II* (2007).

[3] Mark M. Wilde, *Quantum Information Theory*,Cambridge University Press (2017).

[4] Tin Zhang, Graeme Smith, Davide Girolami et al., *Quantification of Entanglement and Coherence with Purity Detection.* In : `NPJ QUANTUM INFORMATION 10, 60` (2024) arXiv:2308.07068.

[5] Omar Fawzi and Renato Renner, *Quantum conditional mutual information and approximate Markov chains*, (2015) arXiv:1410.0664.

[6] Ke Li and Andreas Winter, *Relative entropy and squashed entanglement.* In : `Commun.Math.Phys.326(1):63-80`, (2014) arXiv:1210.3181.

[7] Fernando G.S.L. Brandão, Matthias Christandl and Jon Yard, *Faithful Squashed Entanglement.* In : `Commun. Math. Phys. 306, 805`, (2011) arXiv:1010.1750.

[8] Fernando G.S.L. Brandão, Aram W. Harrow, Jonathan Oppenheim and Sergii Strelchuk, *Quantum Conditional Mutual Information, Reconstructed States and State Redistribution.* In : `PHYSICAL REVIEW LETTERS`, (2014) arXiv:1411.4921.

[9] Eric A. Carlen and Elliot H. Lieb, *Bounds for Entanglement via an Extension of Strong Subadditivity of Entropy*, (2012) arXiv:1203.4719.