Observable Lower Bounds to Quantum Information

Politecnico di Torino

1859

Master degree in

Physics of Complex Systems

CANDIDATE:
Francois Payn

SUPERVISOR:
Prof. Davide Girolami

Academic Year 2024/2025

Abstract

Entanglement is a fundamental property of quantum mechanics and plays a central role in quantum information science, quantum computation, and many-body physics. It represents non-classical correlations between subsystems that cannot be explained by local hidden variables. Despite its conceptual importance, entanglement remains difficult to quantify in practice because it is not directly observable. Various tools have been developed to estimate entanglement, among these, a particularly insightful approach is based on the concept of squashed entanglement, a rigorous entanglement measure defined through the quantum conditional mutual information. Specifically, the squashed entanglement of a bipartite state is given by the infimum of the conditional mutual information over all possible extensions. This measure is of strong theoretical interest because it connects entanglement with information-theoretic quantities, and thus with entropy; however, it poses significant challenges for practical computation, as it typically requires complete knowledge of the global quantum state, which is experimentally inaccessible in most settings. This work addresses the challenge by deriving experimentally accessible lower bounds on the conditional mutual information in many-body quantum systems. Through numerical analysis, we validate the effectiveness of these bounds in several nontrivial scenarios, highlighting their utility not only for entanglement quantification, especially in approximating squashed entanglement, but also for a wide range of tasks crucial to the effective characterization of components in quantum computing.

Contents

Chapter 1

Introduction

1.1 Linear Algebra

A *spanning set* for a vector space is a collection of vectors $|v_1\rangle, \ldots, |v_n\rangle$ such that any vector $|v\rangle$ in the space can be written as a linear combination of the vectors in this collection:

$$|v\rangle = \sum_i a_i |v_i\rangle . \tag{1.1}$$

A set of non-zero vectors $|v_1\rangle, \ldots, |v_n\rangle$ is considered *linearly dependent* if there exists a set of scalars $a_1, \ldots, a_n$, where at least one $a_i \neq 0$, satisfying:

$$a_1 |v_1\rangle + a_2 |v_2\rangle + \cdots + a_n |v_n\rangle = 0. \tag{1.2}$$

Conversely, a set of vectors is called *linearly independent* if it does not satisfy the above condition for any non-zero coefficients. A *basis* of a vector space V is a set of vectors that are both linearly independent and spanning. The number of vectors in the basis is referred to as the *dimension* of the vector space.

[0] This chapter is based on Part 1 Chapter 2 of Nielsen and Chuang [3].

1.1.1 Linear Operators and Matrices

Given two vector spaces V and W, a *linear operator* is a function $A : V \to W$ that satisfies linearity:

$$A\left|v\right\rangle = A\left(\sum_i a_i \left|v_i\right\rangle\right) = \sum_i a_i A(\left|v_i\right\rangle). \tag{1.3}$$

Two significant linear operators worth noting are the *Identity operator* I, which satisfies $I\left|v\right\rangle = \left|v\right\rangle$, and the *zero operator* 0, which satisfies $0\left|v\right\rangle = 0$. To express operators in a matrix form, consider a linear operator $A : V \to W$ and suppose $\{\left|v_1\right\rangle, \ldots, \left|v_n\right\rangle\}$ and $\{\left|w_1\right\rangle, \ldots, \left|w_m\right\rangle\}$ are bases of V and W, respectively. For each j, there exist scalars A_{ij} such that:

$$A\left|v_i\right\rangle = \sum_j A_{ij}\left|w_j\right\rangle. \tag{1.4}$$

The matrix whose elements are A_{ij} is called the *matrix representation* of the operator A.

1.1.2 The Pauli Matrices

Four important matrices that are frequently used in quantum mechanics are the *Pauli matrices.* These are defined as:

$$\begin{aligned} \sigma_0 &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \sigma_y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned} \tag{1.5}$$

1.1.3 Inner Products

An *inner product* is a function $(\cdot,\cdot) : V \times V \to \mathbf{C}$ that takes a pair of vectors $\left|v\right\rangle$ and $\left|w\right\rangle$ from a vector space V and returns a complex number. This operation is denoted as $(v, w) = \langle v|w\rangle$. The notation $\langle v|$ represents the dual vector of $\left|v\right\rangle$, which is a linear functional mapping vectors in V to the complex numbers $\mathbf{C}$. To qualify as an inner product, the function must satisfy the following properties:

1. Linearity in the second argument:
$$\left(|v\rangle, \sum_i \lambda_i |w_i\rangle\right) = \sum_i \lambda_i(|v\rangle, |w_i\rangle).$$

2. Conjugate symmetry:
$$(|v\rangle, |w\rangle) = (|w\rangle, |v\rangle)^*.$$

3. Positive definiteness:
$$(|v\rangle, |v\rangle) \geq 0,$$
with equality if and only if $|v\rangle = 0$.

A vector space equipped with an inner product is called an *inner product space*. Inner products enable the introduction of several key concepts. Two vectors are said to be *orthogonal* if their inner product is zero. The *norm* of a vector $|v\rangle$ is defined as:

$$\| |v\rangle \| = \sqrt{\langle v|v\rangle}. \tag{1.6}$$

A vector is called a *unit vector* or *normalized* if $\| |v\rangle \| = 1$. If this is not the case, the normalized form of $|v\rangle$ can be obtained as $|v\rangle / \| |v\rangle \|$. A set of vectors $\{|i\rangle\}$ is said to be *orthonormal* if each vector is normalized and orthogonal to every other vector in the set, i.e., $\langle v_i|v_j\rangle = \delta_{ij}$.

1.1.4 Outer Product Representation

Linear operators can be conveniently expressed using the *outer product* notation. For vectors $|v\rangle$ and $|w\rangle$ from the inner product spaces V and W, respectively, the operator $|w\rangle\langle v|$ is defined as the linear map from V to W such that:

$$(|w\rangle\langle v|)(|v'\rangle) = |w\rangle\langle v| \, |v'\rangle = \langle v|v'\rangle \, |w\rangle \, . \tag{1.7}$$

Here, the result is the vector $|w\rangle$ scaled by the scalar $\langle v|v'\rangle$. Linear combinations of such outer product operators can also be constructed. For example, $\sum_i a_i |w_i\rangle\langle v_i|$ represents a linear operator that, when applied to a vector $|v'\rangle$, produces:

$$\sum_i a_i |w_i\rangle \langle v_i|v'\rangle \, . \tag{1.8}$$

1.1.5 Completeness Relation and Operator Representation

A key application of the outer product formalism is the *completeness relation* for orthonormal bases. Let $\{|i\rangle\}$ be an orthonormal basis for a vector space V. Any vector $|v\rangle \in V$ can be expressed as $|v\rangle = \sum_i v_i |i\rangle$, where $v_i = \langle i|v\rangle$. Using this, we have:

$$\left(\sum_i |i\rangle\langle i|\right) |v\rangle = \sum_i |i\rangle\langle i| \, |v\rangle = \sum_i v_i \, |i\rangle = |v\rangle \, . \tag{1.9}$$

Since this equation holds for all $|v\rangle$, it means that:

$$\sum_i |i\rangle\langle i| = I, \tag{1.10}$$

where I is the identity operator. Using the completeness relation, any linear operator $A : V \to W$ can be represented in the outer product form. Let $\{|v_i\rangle\}$ be an orthonormal basis for V and $\{|w_j\rangle\}$ be an orthonormal basis for W. Then, A can be written as:

$$A = I_W A I_V = \sum_{ij} |w_j\rangle\langle w_j| \, A \, |v_i\rangle\langle v_i| = \sum_{ij} \langle w_j| \, A \, |v_i\rangle \, |w_j\rangle\langle v_i| \, . \tag{1.11}$$

This is the outer product representation of the operator A.

1.1.6 Eigenvectors and Eigenvalues

Using this notation, the *eigenvectors* $|v\rangle$ and *eigenvalues* v of an operator A are defined by the equation:

$$A \, |v\rangle = v \, |v\rangle \, .$$

The *eigenspace* associated with an eigenvalue v is the set of all vectors that share this eigenvalue, forming a subspace of the vector space on which A acts. A *diagonal representation* of the operator A is given by:

$$A = \sum_i \lambda_i \, |i\rangle\langle i| \, ,$$

where $\{|i\rangle\}$ is an orthonormal set of eigenvectors of A with corresponding eigenvalues $\{\lambda_i\}$. An operator is called *diagonalizable* if such a representation exists.

1.1.7 Adjoints and Hermitian Operators

For a linear operator A on a Hilbert space V, its *adjoint* (or Hermitian conjugate), denoted $A^\dagger$, is the unique operator satisfying:

$$(|v\rangle, A|w\rangle) = (A^\dagger |v\rangle, |w\rangle),$$

for all $|v\rangle, |w\rangle \in V$. It can be shown that $(AB)^\dagger = B^\dagger A^\dagger$. Moreover, for a vector $|v\rangle$, we adopt the convention $|v\rangle^\dagger = \langle v|$, from which it follows that:

$$(A|v\rangle)^\dagger = \langle v| A^\dagger.$$

An operator A is called *Hermitian* if $A^\dagger = A$. A *projector* is a specific type of operator used to project vectors onto a subspace. Suppose W is a k-dimensional subspace of a d-dimensional vector space V. We can construct an orthonormal basis $\{|1\rangle, \ldots, |d\rangle\}$ for V such that $\{|1\rangle, \ldots, |k\rangle\}$ forms an orthonormal basis for W. The *projection operator* onto W is then defined as:

$$P := \sum_{i=1}^{k} |i\rangle\langle i|.$$

An operator A is called *normal* if $AA^\dagger = A^\dagger A$ and *unitary* if $A^\dagger A = I$. Unitary operators are important because they preserve inner products between vectors.

Theorem 1 (Spectral decomposition). *Any operator is normal if and only if it is diagonalizable*

1.1.8 Tensor Products

Let V and W be vector spaces of dimensions m and n, respectively. Their *tensor product* $V \otimes W$ is a vector space of dimension mn. The elements of $V \otimes W$ are linear combinations of *tensor products* $|v\rangle \otimes |w\rangle$, where $|v\rangle \in V$ and $|w\rangle \in W$. If $\{|i\rangle\}$ and $\{|j\rangle\}$ are orthonormal bases for V and W, then $\{|i\rangle \otimes |j\rangle\}$ forms an orthonormal basis for $V \otimes W$.

1.1.9 Operator Functions

Functions can also be defined for operators and matrices. Given a function $f : \mathbf{C} \to \mathbf{C}$ and the spectral decomposition of an operator $A = \sum_a a\, |a\rangle\langle a|$, the operator function is defined as:

$$f(A) = \sum_a f(a) \left|a\right\rangle\!\left\langle a\right| .$$

One important operator function is the *trace* of a matrix, defined as:

$$\mathrm{tr}(A) := \sum_i A_{ii}.$$

The trace has two key properties: it is *cyclic*, meaning $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, and *linear*, meaning $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$ and $\mathrm{tr}(zA) = z\mathrm{tr}(A)$ for any matrices A, B and scalar z. Additionally, the trace is invariant under *unitary similarity transformations*, such that for $A \to UAU^\dagger$, we have:

$$\mathrm{tr}(UAU^\dagger) = \mathrm{tr}(A).$$

The trace of an operator A is defined as the trace of any matrix representation of A.

1.1.10 The Commutator and Anti-Commutator

The *commutator* of two operators A and B is defined as:

$$[A, B] := AB - BA.$$

If $[A, B] = 0$, then A and B are said to *commute*. The *anti-commutator* of A and B is defined as:

$$\{A, B\} := AB + BA,$$

and A and B are said to *anticommute* if $\{A, B\} = 0$.

Theorem 2 (Simultaneous Diagonalization Theorem)**.** *Let A and B be Hermitian operators. Then $[A, B] = 0$ if and only if there exists an orthonormal basis in which both A and B are diagonal. In this case, A and B are said to be* simultaneously diagonalizable*.*

1.1.11 The Polar and Singular Value Decompositions

The *polar decomposition* and the *singular value decomposition* (SVD) are powerful techniques for expressing linear operators in terms of simpler components.

Theorem 3 (Polar Decomposition)**.** *Let A be a linear operator on a vector space V.*

There exist a unitary operator U and positive operators J and K such that:

$$A = UJ = KU,$$

The positive operators J and K are uniquely defined as $J := \sqrt{A^\dagger A}$ and $K := \sqrt{AA^\dagger}$. Also, if A is invertible, the unitary operator U is also unique.

Theorem 4 (Singular Value Decomposition)**.** *Given a square matrix A, there exist unitary matrices U and V and a diagonal matrix D with non-negative entries such that:*

$$A = UDV,$$

The diagonal elements of D are the singular values of A.

1.2 The postulates of Quantum Mechanics

In this chapter, we provide a description of the fundamental postulates of quantum mechanics. These postulates establish the connection between the physical reality of quantum systems and the mathematical framework.

1.2.1 State Space

> **Postulate 1**: Every isolated physical system is associated with a complex vector space equipped with an inner product (a Hilbert space), referred to as the state space of the system. The state of the system is fully described by a state vector, which is a unit vector in this Hilbert space.

The simplest quantum mechanical system has a two-dimensional state space and is known as a *qubit.* Its state is a vector of a Hilbert space spanned by the orthonormal basis $\{|0\rangle, |1\rangle\}$. Any state $|\psi\rangle$ in this space can be expressed as:

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle$$

where a and b are complex numbers. These coefficients must satisfy the normalization condition:

$$\langle\psi|\psi\rangle = |a|^2 + |b|^2 = 1.$$

A linear combination of the form $\sum_i \alpha_i\,|\psi_i\rangle$ is called a *superposition* of the states $|\psi_i\rangle$, with α_i denoting the *amplitude* for the state $|\psi_i\rangle$.

1.2.2 Evolution

The second postulate of quantum mechanics prescribes how a physical system evolves over time:

> **Postulate 2**: A closed quantum system evolve according to a unitary transformation. Specifically, the state $|\psi\rangle$ of the system at time t is related to the state $|\psi\rangle$ at time t' via a unitary operator U, which depends only on the times t and t':
>
> $$|\psi(t')\rangle = U\,|\psi(t)\rangle\,. \tag{1.12}$$

This postulate does not specify the particular unitary operator U that governs the dynamics of a given system. Examples of unitary operators include the Pauli matrices σ_x, σ_y, σ_z, and the *Hadamard gate* H, which acts on the qubit basis as follows:

$$H\left|0\right\rangle = \frac{\left|0\right\rangle + \left|1\right\rangle}{\sqrt{2}}, \quad H\left|1\right\rangle = \frac{\left|0\right\rangle - \left|1\right\rangle}{\sqrt{2}}. \tag{1.13}$$

There exist also an alternative formulation of Postulate 2, which describes the evolution of a quantum system in continuous time:

> **Postulate 2':** The time evolution of the state of a closed quantum system is governed by the *Schrödinger equation*:
>
> $$i\hbar\frac{\mathrm{d}\left|\psi\right\rangle}{\mathrm{d}t} = H\left|\psi\right\rangle. \tag{1.14}$$
>
> In this equation, $\hbar$ is the reduced Planck constant. In many practical contexts, it is common to set $\hbar = 1$. The operator H, known as the *Hamiltonian*, is an Hermitian operator that characterizes the energy and dynamics of the closed system.

Given the Hamiltonian of a system and the value of $\hbar$, the dynamics of the system can, in principle, be completely known. However, identifying the appropriate Hamiltonian that describes a specific physical system is often a challenging problem. Since the Hamiltonian is a Hermitian operator, it admits a spectral decomposition:

$$H = \sum_E E\left|E\right\rangle\left\langle E\right|, \tag{1.15}$$

the states $\left|E\right\rangle$ are called *energy eigenstes* and E is the *energy* of the state $\left|E\right\rangle$. The lowest of these energy is called *ground state*. We can find a connection between the Hamiltonian picture of dynamics and the Unitary operator picture just by finding a solution to the Shrodinger equation:

$$\left|\psi(t')\right\rangle = \exp\left[-\frac{iH(t'-t)}{\hbar}\right]\left|\psi(t)\right\rangle = U(t,t')\left|\psi(t)\right\rangle, \tag{1.16}$$

where

$$U(t,t') \equiv \exp\left[-\frac{iH(t'-t)}{\hbar}\right]. \tag{1.17}$$

It can be shown that this is a unitary operator. Any unitary operator can be written in a form: $U = exp(iK)$ for some Hermitian operator K.

1.2.3 Quantum Measurement

We now explain what happens when the physical system is not closed anymore, Interacting with an environment or being measured.

> **Postulate 3:** Quantum measurements are characterized by a set of *measurement operators* $\{M_m\}$, which act on the state space of the quantum system. The index m corresponds to the possible outcomes of the measurement. If the state of the system before measurement is $|\psi\rangle$, the probability of obtaining the outcome m is given by:
>
> $$p(m) = \langle\psi| M_m^\dagger M_m |\psi\rangle . \tag{1.18}$$
>
> After the measurement, the quantum state collapses to:
>
> $$\frac{M_m |\psi\rangle}{\sqrt{\langle\psi| M_m^\dagger M_m |\psi\rangle}}. \tag{1.19}$$
>
> The measurement operators must satisfy the *completeness relation*, which ensures that probabilities sum to 1:
>
> $$\sum_m M_m^\dagger M_m = I, \tag{1.20}$$
>
> where I is the identity operator.

1.2.4 Distinguishing Quantum States

In this section, we address the statement: orthogonal quantum states can be distinguished, whereas non-orthogonal states cannot. Consider a quantum state $|\psi_i\rangle$ where $(1 \leq i \leq n)$ that needs to be identified. If the states $|\psi_i\rangle$ are orthonormal, we can define a quantum measurement that distinguishes them. For each state $|\psi_i\rangle$, we introduce a measurement operator $M_i = |\psi_i\rangle \langle\psi_i|$, and an additional operator M_0 defined as the positive square root of: $I - \sum_{i\neq 0} |\psi_i\rangle \langle\psi_i|$. Measuring this observable on the state $|\psi_i\rangle$ gives i with probability:

$$p(i) = \langle\psi_i| M_i |\psi_i\rangle = 1,$$

which confirms that orthogonal states can be perfectly distinguished.
In contrast, if the states $|\psi_i\rangle$ are not orthonormal there is no quantum measurement able to distinguish the states with certainty. Lets say we try to do a measurement described by the measurement operator M_j with outcome j and given j we could use some rule to guess the index i of the state like $i = f(j)$. The problem is that if we have non-orthogonal states, for example $|\psi_1\rangle$ and $|\psi_2\rangle$, and we measure a j such that $f(j) = 1$, since $|\psi_2\rangle$ has a non vanishing component parallel to $|\psi_1\rangle$, we will also sometimes get j while the state prepared is $|\psi_2\rangle$, but in that case we would make an error saying that the state is $|\psi_1\rangle$, so in general it is not possible to accurately distinguish the states.

1.2.5 Projective Measurements

A *projective measurement* is defined by an observable M, which is an Hermitian operator acting on the state space of the system under observation. The observable M can be decomposed as:

$$M = \sum_m m P_m, \tag{1.21}$$

where P_m is the projector onto the eigenspace of M corresponding to the eigenvalue m. The eigenvalues m represent the possible outcomes of the measurement. If the system is in the state $|\psi\rangle$, the probability of observing the outcome m is:

$$p(m) = \langle\psi| P_m |\psi\rangle . \tag{1.22}$$

After observing the outcome m, the state of the system collapses to:

$$\frac{P_m |\psi\rangle}{\sqrt{p(m)}}. \tag{1.23}$$

Projective measurements simplify the computation of expectation values for observables. The expectation value of M is:

$$\begin{aligned}\mathbb{E}(M) &= \sum_m m\, p(m) \\ &= \sum_m m \langle\psi|P_m|\psi\rangle \\ &= \langle\psi|\left(\sum_m mP_m\right)|\psi\rangle \\ &= \langle\psi|M|\psi\rangle .\end{aligned}$$

Using the shorthand $\langle M\rangle \equiv \langle\psi| M |\psi\rangle$ for the average value of M, the variance of M is given by:

$$[\Delta M]^2 = \left\langle (M - \langle M\rangle)^2 \right\rangle = \left\langle M^2 \right\rangle - \langle M\rangle^2 . \tag{1.24}$$

Heisenberg Uncertainty Principle

The Heisenberg Uncertainty Principle is a cornerstone of quantum mechanics, governing the intrinsic limits on the precision with which certain pairs of observables can be simultaneously measured. Consider two Hermitian operators, A and B, and a quantum state $|\psi\rangle$. Let $\langle\psi| AB |\psi\rangle = x + iy$, where $x, y \in \mathbb{R}$. From this it is easy to get: $\langle\psi| [A, B] |\psi\rangle = 2iy$ and $\langle\psi| \{A, B\} |\psi\rangle = 2x$, and it follows that:

$$|\langle\psi| [A, B] |\psi\rangle|^2 + |\langle\psi| \{A, B\} |\psi\rangle|^2 = 4|\langle\psi| AB |\psi\rangle|^2. \tag{1.25}$$

Applying the Cauchy-Schwarz inequality to $\langle\psi| AB |\psi\rangle$ gives:

$$|\langle\psi| AB |\psi\rangle|^2 \leq \langle\psi| A^2 |\psi\rangle \langle\psi| B^2 |\psi\rangle . \tag{1.26}$$

Combining this result with Eq. (1.25), we deduce:

$$|\langle\psi| [A, B] |\psi\rangle|^2 \leq 4 \langle\psi| A^2 |\psi\rangle \langle\psi| B^2 |\psi\rangle . \tag{1.27}$$

Now we consider two observables C and D, defined by $A = C - \langle C\rangle$ and $B = D - \langle D\rangle$, substituting back yields the uncertainty relation:

$$\Delta(C)\Delta(D) \geq \frac{|\langle\psi| [C, D] |\psi\rangle|}{2}. \tag{1.28}$$

This means that if an ensemble of systems is prepared in the same state $|\psi\rangle$, and measurements of C and D are performed on different subsets of the ensemble, the product of the standard deviations $\Delta(C)$ and $\Delta(D)$ will obey Eq. (1.28).

1.2.6 POVM Measurements

Consider a quantum measurement characterized by the measurement operators M_m, applied to a quantum system in the state $|\psi\rangle$. The probability of observing the outcome m is again:

$$p(m) = \langle\psi| M_m^\dagger M_m |\psi\rangle\,. \tag{1.29}$$

Defining $E_m = M_m^\dagger M_m$, we see that E_m is a positive operator satisfying:

$$\sum_m E_m = I, \quad p(m) = \langle\psi| E_m |\psi\rangle\,.$$

The set $\{E_m\}$ is known as a *Positive Operator-Valued Measure* (POVM), representing the measurement. Unlike projective measurements, POVMs generalize quantum measurements to include scenarios involving noise or partial information.

1.2.7 Phase

The notion of *phase* plays a critical role in quantum mechanics. Two key concepts are the *global phase* and the *relative phase*. A global phase factor is a complex scalar of the form $e^{i\theta}$, where $\theta \in \mathbb{R}$. If a quantum state $|\psi\rangle$ is multiplied by a global phase $e^{i\theta}$, the resulting state $e^{i\theta}|\psi\rangle$ is physically indistinguishable from $|\psi\rangle$. To see this, consider a measurement operator M_m:

$$\langle\psi| e^{-i\theta} M_m^\dagger M_m e^{i\theta} |\psi\rangle = \langle\psi| M_m^\dagger M_m |\psi\rangle\,. \tag{1.30}$$

Since global phases have no observable effect, they are often disregarded in quantum descriptions. In contrast, a relative phase captures the phase difference between components of a quantum state. For example, consider the states:

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

we see that the component $|1\rangle$ has the same $\frac{1}{\sqrt{2}}$ amplitude in two states, but a phase shift of -1. Relative phase differences may influence interference patterns and measurable outcomes, making them physically significant. More generally, two amplitudes a and b differ by a relative phase ϕ if:

$$\frac{a}{b} = e^{i\phi}.$$

Quantum states differing by relative phase are in general not physically equivalent.

1.2.8 Composite Systems

The following postulate defines the structure of the state space for composite quantum systems:

> **Postulate 4:** The state space of a composite system is the tensor product of the state spaces of its components. Consider n quantum states prepared as $|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_n\rangle$, the joint state of the composite system is:
>
> $$|\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle .$$

Using this postulate, we can define *entanglement*. Consider the state:

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}. \tag{1.31}$$

This state cannot be expressed as a product $|a\rangle |b\rangle$ of single-qubit states. Such states, which cannot be written as separable products, are named *entangled*.

1.3 Superdense coding

Superdense coding is a remarkable application of elementary quantum mechanics It involves two parties conventionally known as Alice and Bob, Bob's goal here is to transmit two bits of classical information to Alice just sending her only one qubit. So they initially share and entangled pair of qubits prepared in the state:

$$|\psi\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}. \tag{1.32}$$

Bob can encode two bits of classical information into the entangled state by applying specific operations to his qubit. The transformations are as follows:

$$\begin{aligned} 00 : |\psi\rangle &\rightarrow \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \\ 01 : |\psi\rangle &\rightarrow \frac{|00\rangle - |11\rangle}{\sqrt{2}}, \\ 10 : |\psi\rangle &\rightarrow \frac{|10\rangle + |01\rangle}{\sqrt{2}}, \\ 11 : |\psi\rangle &\rightarrow \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \end{aligned}$$

The interpretation of this is:

- To transmit the bit string "00", Bob performs no operation on his qubit.
- To send "01", he applies the Z operator.
- To encode "10", he applies the X operator.
- To transmit "11", he applies the iY operator.

The four resulting states form the *Bell basis*, that is an orthonormal basis of the two-qubit system:

$$\left\{ \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \frac{|00\rangle - |11\rangle}{\sqrt{2}}, \frac{|10\rangle + |01\rangle}{\sqrt{2}}, \frac{|01\rangle - |10\rangle}{\sqrt{2}} \right\}.$$

Once Bob applies the appropriate operation to his qubit, he sends it to Alice. After receiving Bob's qubit, Alice can perform a measurement in the Bell basis, allowing her to determine which of the four states the system is in. This enables Alice to decode the two bits of classical information encoded by Bob.

1.4 The density operator

It is possible to describe quantum mechanics using an alternative framework to the state vector formalism, employing a tool known as the *density matrix.*

1.4.1 Ensemble of Quantum States

A quantum system whose state $|\psi\rangle$ is known exactly is said to be a *pure state*, in this case the density matrix is simply $\rho = |\psi\rangle \langle\psi|$, otherwise ρ is in a mixed state and in this case the system is in a state $|\psi_i\rangle$ with probability p_i and the corresponding density matrix is defined as a weighted sum of pure state density matrices:

$$\rho \equiv \sum_i p_i |\psi_i\rangle \langle\psi_i| \tag{1.33}$$

In this chapter we reformulate the postulates of quantum mechanics using the density operator formalism. We start with some examples.

- **Evolution:** The time evolution of a closed quantum system is governed by a unitary operator U. If the system starts in the state $|\psi_i\rangle$ with probability p_i, then after the evolution, it transitions to $U|\psi_i\rangle$ with the same probability. Accordingly, the evolution of the density operator is:

  $$\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i| \longrightarrow \sum_i p_i U |\psi_i\rangle \langle\psi_i| U^\dagger = U\rho U^\dagger.$$

- **Measurement:** For a measurement described by the operator M_m, if the system initially occupies the state $|\psi_i\rangle$, the probability of obtaining the measurement result m is:

  $$p(m|i) = \langle\psi_i| M_m^\dagger M_m |\psi_i\rangle = \mathrm{tr}(M_m^\dagger M_m |\psi_i\rangle \langle\psi_i|).$$

  The total probability of observing the outcome m is then:

  $$p(m) = \sum_i p(m|i) p_i = \mathrm{tr}(M_m^\dagger M_m \rho).$$

After measuring outcome m, the new density matrix is:

$$\rho_m = \frac{M_m \rho M_m^\dagger}{\operatorname{tr}(M_m M_m^\dagger \rho)}.$$

1.4.2 General Properties

The set of all density operators is defined by the following characterization theorem: **Theorem 5.** Characterization of Density Operators*: An operator ρ represents a density matrix corresponding to some ensemble $\{p_i, |\psi_i\rangle\}$ if and only if it satisfies the following conditions:*

1. ***Trace Condition:*** *ρ has trace one i.e.* $\operatorname{tr}(\rho) = 1$.
2. ***Positivity Condition:*** *ρ is a positive semidefinite operator.*

Using this theorem, we can redefine a density operator as a positive operator that has trace equal to one acting on the state space of the system. Now, using this definition, we can provide an alternative version of the four postulates of quantum mechanics in terms of the density operator.

- **Postulate 1:** Every isolated quantum system has associated a complex vector space equipped with an inner product, known as the *state space*. The system is completely characterized by its density operator ρ, which is positive and has trace equal to one. For an ensemble where the system is in state ρ_i with probability p_i, the density operator is $\rho = \sum_i p_i \rho_i$.
- **Postulate 2:** The evolution of a closed quantum system is described by a unitary transformation. If ρ_1 represents the state of the system at time t_1, then the state ρ_2 at time t_2 is given by:

  $$\rho_2 = U \rho_1 U^\dagger,$$

  where U is a unitary operator that depends only on times t_1 and t_2.
- **Postulate 3:** Quantum measurements are represented by a set of measurement operators $\{M_m\}$, where each M_m corresponds to a possible measurement outcome m. For a system in state ρ, the probability of obtaining the outcome m is given by:

  $$p(m) = \operatorname{tr}(M_m^\dagger M_m \rho),$$

and the post-measurement state of the system is:

$$\rho_m = \frac{M_m \rho M_m^\dagger}{\mathrm{tr}(M_m^\dagger M_m \rho)}.$$

These operators satisfy the completeness relation:

$$\sum_m M_m^\dagger M_m = I.$$

Now we will try to answer an important question that arises: which ensembles of states can produce a given density matrix? To address this, we introduce vectors $|\tilde{\psi}_i\rangle$, which are not necessarily normalized. These vectors are used to construct the density operator as:

$$\rho = \sum_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|.$$

The connection to the standard ensemble representation is given by $|\tilde{\psi}_i\rangle = \sqrt{p_i}\,|\psi_i\rangle$. Then, under what conditions do two sets of vectors $\{|\tilde{\psi}_i\rangle\}$ and $\{|\tilde{\varphi}_i\rangle\}$ generate the same density matrix ρ?

Theorem 6 (Unitary Freedom in the Ensemble Representation of Density Matrices)**.** *Two sets of states $\{|\tilde{\psi}_i\rangle\}$ and $\{|\tilde{\varphi}_i\rangle\}$ generate the same density matrix ρ if and only they satisfy this relation:*

$$|\tilde{\psi}_i\rangle = \sum_j u_{ij} |\tilde{\varphi}_j\rangle,$$

here u_{ij} is a unitary matrix of complex coefficients. If the two ensembles have different cardinalities, we pad the smaller one with zero vectors until both have the same number of elements.

1.4.3 The Reduced Density Operator

In the context of composite quantum systems, the *reduced density operator* provides a description of a subsystem's state. For a system composed of two parts, A and B, described by the density operator ρ^{AB}, the reduced density operator for system A is:

$$\rho^A = \mathrm{tr}_B(\rho^{AB}),$$

where tr_B is known as the *partial trace* over the system B. The partial trace is explicitly defined for product states as:

$$\mathrm{tr}_B\left(|a_1\rangle\langle a_2|\otimes|b_1\rangle\langle b_2|\right)=|a_1\rangle\langle a_2|\ \mathrm{tr}(|b_1\rangle\langle b_2|),$$

where $|a_1\rangle$ and $|a_2\rangle$ belong to the Hilbert space of subsystem A, and $|b_1\rangle$ and $|b_2\rangle$ belong to that of subsystem B.

1.5 The Schmidt decomposition and purifications

We now discuss two foundational concepts in quantum information theory and quantum computation: the Schmidt decomposition and the process of purification.

1.5.1 Schmidt Decomposition

Theorem 7 (Shmidt decomposition)**.** *For any pure quantum state $|\psi\rangle$ of a composite system AB, there exist orthonormal state $|i_A\rangle$ for subsystem A and $|i_B\rangle$ for subsystem B such that:*

$$|\psi\rangle = \sum_i \lambda_i \, |i_A\rangle \, |i_B\rangle \,, \tag{1.34}$$

where λ_i are non-negative real numbers called the Schmidt coefficients*, satisfying the normalization condition $\sum_i \lambda_i^2 = 1$.*

This result is particularly useful in analyzing quantum systems. For a pure state $|\psi\rangle$ of a composite system AB, the reduced density matrices for the subsystems A and B can be written as:

$$\rho^A = \sum_i \lambda_i^2 \, |i_A\rangle\!\langle i_A| \,, \tag{1.35}$$

$$\rho^B = \sum_i \lambda_i^2 \, |i_B\rangle\!\langle i_B| \,. \tag{1.36}$$

This demonstrates that the eigenvalues of ρ^A and ρ^B are identical and equal to λ_i^2. This means that for a pure state of a composite system, properties determined by the eigenvalues of the reduced density matrix are the same for both subsystems. The orthonormal bases $\{|i_A\rangle\}$ and $\{|i_B\rangle\}$ are known as the *Schmidt bases* for subsystems A and B, respectively. The number of non-zero Schmidt coefficients, λ_i, is referred to as the *Schmidt number* of the state $|\psi\rangle$. This quantity serves as a measure of entanglement between subsystems A and B.

An important property of the Schmidt decomposition is that the Schmidt number remains unchanged under unitary transformations applied to either subsystem A or B individually. To see this, consider the Schmidt decomposition of $|\psi\rangle$:

$$|\psi\rangle = \sum_i \lambda_i \, |i_A\rangle \, |i_B\rangle \,.$$

Applying a unitary operator U to subsystem A, the Schmidt decomposition of the state $U\left|\psi\right\rangle$ will be:

$$U\left|\psi\right\rangle = \sum_i \lambda_i (U\left|i_A\right\rangle)\left|i_B\right\rangle.$$

The new state retains the Schmidt decomposition structure, with the same Schmidt coefficients λ_i.

1.5.2 Purification

Another key concept in quantum information theory is **purification**. Any mixed state ρ^A of a quantum system A can be associated with a pure state $\left|AR\right\rangle$ of an extended system AR, where R is an auxiliary system known as the *reference system.* This pure state is constructed such that:

$$\rho^A = \mathrm{tr}_R(|AR\rangle\langle AR|),$$

where tr_R denotes the partial trace over the reference system R. The reference system has no physical interpretation and is introduced purely as a mathematical construct. To demonstrate this, consider a mixed state ρ^A with an orthonormal decomposition:

$$\rho^A = \sum_i p_i \left|i^A\right\rangle \left\langle i^A\right|.$$

Introduce a reference system R with the same state space as A, and let $\{\left|i^R\right\rangle\}$ form an orthonormal basis for R. Define a pure state for the combined system AR as:

$$\left|AR\right\rangle = \sum_i \sqrt{p_i}\left|i^A\right\rangle\left|i^R\right\rangle.$$

To verify that $\left|AR\right\rangle$ purifies ρ^A, calculate the reduced density matrix for system A:

$$\mathrm{tr}_R(|AR\rangle\langle AR|) = \sum_{i,j}\sqrt{p_i p_j}\left|i^A\right\rangle\left\langle j^A\right| \mathrm{tr}(\left|i^R\right\rangle\left\langle j^R\right|) \tag{1.37}$$

$$= \sum_{i,j}\sqrt{p_i p_j}\left|i^A\right\rangle\left\langle j^A\right| \delta_{ij} \tag{1.38}$$

$$= \sum_i p_i \left|i^A\right\rangle\left\langle i^A\right| \tag{1.39}$$

$$= \rho^A. \tag{1.40}$$

Thus, $|AR\rangle$ satisfies the condition for purification. The relationship between Schmidt decomposition and purification is really important. When purifying a mixed state ρ^A, the procedure involves constructing a pure state where the Schmidt basis for system A is the eigenbasis of ρ^A, and the Schmidt coefficients are the square roots of the eigenvalues of ρ^A.

1.6 EPR and the Bell inequality

In classical physics, an object's properties are well-defined and independent of observation. However, quantum mechanics introduces a different perspective: physical properties often lack definite values until measurement. This concept is central to quantum mechanics and challenges classical intuition. Einstein, Podolsky, and Rosen (EPR) famously critiqued this interpretation, arguing that quantum mechanics was incomplete. They proposed the concept of *elements of reality*: if the value of a physical property can be predicted with certainty without direct measurement, it must correspond to an element of reality. To illustrate this concept we consider an entangled pair of qubits shared between two parties Alice and Bob, represented by:

$$\frac{|01\rangle - |10\rangle}{\sqrt{2}}. \tag{1.41}$$

If Alice measures the spin along the $\vec{v}$ direction, she determines Bob's corresponding spin with certainty. For instance, if Alice measures $+1$, Bob is guaranteed to measure -1 when measuring along the same direction. This suggests that Alice is always able to know the value of Bob's measurement before he makes it. This implies that the result is an element of reality and should be represented in any complete physical theory according to the three scientists. Einstein, Podolsky, and Rosen wanted to prove that quantum mechanics is incomplete. They argued that it was missing certain elements of reality, meaning things that exist whether we measure them or not. Their goal was to bring back a more classical view of the world, where systems have definite properties that don't depend on measurements. However, experiments show that this classical view is wrong and that quantum mechanics is correct. The main idea behind this experimental proof is called the **Bell Inequality**. This result doesn't come from quantum mechanics itself. Instead, it is based on a thought experiment that uses our everyday understanding of how the world works—something Einstein believed nature followed. By looking at this idea, we will see that quantum mechanics actually disagrees with it.

Thought experiment: Suppose we prepare two particles in such a way that we are able to reproduce at wish this preparation, and we send one of them to Alice and one to Bob. Alice possesses two different measurement apparatuses: one for the property P_Q and one for P_R. For Bob the same thing happens, but he measures different properties P_S and P_T. None of them know in advance which property is going to be measured over their particle, but they have to decide it based on a random event, like a coin toss (both independently). We suppose also that all of these 4 properties can only produce as output ±1. The timing of the two experiments should be such that one's result cannot influence the other's so they must act in a causally disconnected manner. We begin by analyzing the expression $QS+RS+RT-QT$, observing that:

$$QS + RS + RT - QT = (Q + R)S + (R - Q)T. \tag{1.42}$$

Since $R, Q \in \{\pm 1\}$, it follows that either $(Q + R)S = 0$ or $(R - Q)T = 0$. Consequently, from (1.42), we deduce that $QS + RS + RT - QT = \pm 2$.

Now, let $p(q, r, s, t)$ represents the probability that the system is initially in the state characterized by $Q = q$, $R = r$, $S = s$, and $T = t$ prior to any measurement. These probabilities can depend on experimental conditions. Denoting the expectation value of a quantity by $E(\cdot)$, we have:

$$\begin{aligned} E(QS + RS + RT - QT) &= \sum_{q,r,s,t} p(q, r, s, t)(qs + rs + rt - qt) && (1.43) \\ &\leq \sum_{q,r,s,t} p(q, r, s, t) \times 2 && (1.44) \\ &= 2. && (1.45) \end{aligned}$$

Additionally, we can write:

$$\begin{aligned} E(QS + RS + RT - QT) &= \sum_{q,r,s,t} p(q, r, s, t)qs + \sum_{q,r,s,t} p(q, r, s, t)rs && (1.46) \\ &\quad + \sum_{q,r,s,t} p(q, r, s, t)rt - \sum_{q,r,s,t} p(q, r, s, t)qt && (1.47) \\ &= E(QS) + E(RS) + E(RT) - E(QT). && (1.48) \end{aligned}$$

Comparing (1.45) and (1.48), we derive the Bell inequality:

$$E(QS) + E(RS) + E(RT) - E(QT) \leq 2. \tag{1.49}$$

By repeating the experiment multiple times, Alice and Bob can determine the values on the left-hand side of the inequality with arbitrary precision.

Quantum Mechanical Experiment: Now we consider the preparation of the specific quantum state:

$$|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \tag{1.50}$$

The first qubit is given to Alice, and the second qubit to Bob, they will perform measurements of the following observables:

$$Q = Z_1, \quad S = \frac{-Z_2 - X_2}{\sqrt{2}}, \tag{1.51}$$

$$R = X_1, \quad T = \frac{Z_2 - X_2}{\sqrt{2}}. \tag{1.52}$$

Simple calculations give the following expectation values for these observables:

$$\langle QS\rangle = \frac{1}{\sqrt{2}}, \quad \langle RS\rangle = \frac{1}{\sqrt{2}}, \tag{1.53}$$

$$\langle RT\rangle = \frac{1}{\sqrt{2}}, \quad \langle QT\rangle = -\frac{1}{\sqrt{2}}. \tag{1.54}$$

Thus, we find:

$$\langle QS\rangle + \langle RS\rangle + \langle RT\rangle - \langle QT\rangle = 2\sqrt{2}. \tag{1.55}$$

This result violates the Bell inequality, and it can be experimentally confirmed.

The violation implies that at least one of the assumptions that we have made to get the Bell inequality must be incorrect. The two best candidates are:

1. The assumption that physical properties P_Q, P_R, P_S, P_T have definite values Q, R, S, T independent of observation. This is often referred to as the assumption of realism.

2. The assumption that Alice's measurement does not influence Bob's measurement outcome, and vice versa. This is known as locality.

Together, these two principles form the concept of *local realism.* The violation of Bell's inequality demonstrates that at least one of these assumptions must be incorrect.

Chapter 2

Mutual information bounds for quantum systems

2.1 Introduction

In this section, we explore fundamental quantities from information theory, such as mutual information and conditional mutual information, within the framework of quantum systems. Our primary goal is to investigate how these quantities can be estimated using only the observable characteristics of the system under study. This approach aims to make these quantities more physically meaningful and accessible, especially when direct computation is impractical. We focus on identifying rigorous lower bounds for these information-theoretic quantities that are closely related to quantum observables, such as correlation functions and operator average values. By grounding these bounds in physical quantities, we aim to provide tools that can be applied across a wide range of quantum systems, from many-body lattice models to qubit systems. This perspective not only offers a concrete handle on total correlations within a system but also helps reveal the minimal informational content that must exist given certain observable features. Through this approach, we seek to build a more operational understanding of mutual and conditional mutual information in quantum settings, one that connects theoretical quantities with the realities of experimental and numerical analysis.

2.2 Information theory

Information theory deals with the quantification, storage, and transmission of information inside and between systems of various nature. Since Shannon's foundational work, the field has expanded significantly, shaping disciplines such as physics, computer science, biology, and statistics. In our work, we use information theory to explore correlation properties from a purely theoretical standpoint, as well as in the context of specific physical systems. Through this lens, we will show how information-theoretic quantities, introduced and discussed throughout, can be related to experimentally observable quantities.

2.2.1 Self-Information and Entropy

The foundational concept in information theory is the **self-information** (also called information content) of an event x [4]. Given that the probability of the event is $P(x)$, the self-information quantifies how surprising or informative that event is:

$$I(x) = -\log_b P(x) \tag{2.1}$$

Here, the base b is typically taken to be 2, in which case the information is measured in bits. Intuitively, rare events (low $P(x)$) carry more information, while more probable events convey less. Building on this, we define the **Shannon entropy**, which captures the expected (average) self-information of a random variable X with probability distribution $P(x)$ and domain $\mathcal{X}$:

$$H(X) = -\sum_{x\in\mathcal{X}} P(x)\log_b P(x) \tag{2.2}$$

Entropy serves as a quantitative measure of uncertainty or unpredictability in a system. The greater the entropy, the more uncertain we are about the outcome; conversely, reducing entropy corresponds to gaining information. We now introduce some really important quantity that derives from these concepts. For classical systems, given two random variables X and Y we define the **Joint Entropy** as:

$$H(X,Y) = -\sum_{x,y} P(x,y)\log P(x,y) \tag{2.3}$$

And the **Conditional Entropy** as:

$$H(X|Y) = -\sum_{x,y} P(x,y) \log P(x|y) \tag{2.4}$$

If we have two different distributions over the same domain $\mathcal{X}$, namely P and Q, we can define the Kullback-Leibler (KL) divergence between them as:

$$D_{\mathrm{KL}}(P\|Q) = \sum_{x\in\mathcal{X}} P(x) \log \frac{P(x)}{Q(x)} \tag{2.5}$$

KL divergence measures how different the distribution Q is from the true distribution P. It quantifies the expected amount of additional information (or "surprise") incurred when using Q to approximate P. Although not a true distance metric, since it is not symmetric and does not satisfy the triangle inequality, KL divergence is widely used as a measure of dissimilarity between probability distributions.

These ideas can be extended naturally to quantum systems. In the **quantum setting**, a system is described by a density matrix ρ, and the analogue of Shannon entropy is the **von Neumann Entropy**:

$$S(\rho) = -\mathrm{Tr}[\rho \log \rho] \tag{2.6}$$

The von Neumann entropy plays a central role in quantum information theory, quantifying the uncertainty, or mixedness, of a quantum state. Like its classical counterpart, it provides insight into the information content and correlations present in the system.

If we consider a quantum system composed of two subsystems, A and B. The full system is described by the density matrix ρ_{AB}, while the reduced states of the subsystems are given by $\rho_A = \mathrm{Tr}_B[\rho_{AB}]$ and $\rho_B = \mathrm{Tr}_A[\rho_{AB}]$. The **Joint Entropy** of the system is then defined as:

$$S(AB) = -\mathrm{Tr}[\rho_{AB} \log \rho_{AB}] \tag{2.7}$$

Which is simply the entropy of the total system A + B. We define also the **Conditional Entropy**:

$$S(A|B) = S(AB) - S(B) \tag{2.8}$$

And finally an object that quantifies how hard it is to distinguish between two states, and this is called **Relative Entropy**:

$$S(\rho||\sigma) = Tr[\rho(log(\rho) - log(\sigma))] \tag{2.9}$$

Note: In the following, we will use the notations $S(A)$ and $S(\rho_A)$ interchangeably to denote the Von Neumann entropy of quantum system A.

2.2.2 Mutual Information

In classical information theory, **mutual information** measures how much information two random variables share. For two classical random variables X and Y with joint probability distribution $P(X, Y)$, the mutual information is defined as:

$$I(X : Y) = \sum_{x,y} P(x, y) \log \left(\frac{P(x, y)}{P(x)P(y)} \right) \tag{2.10}$$

This expression quantifies how much knowing the value of X reduces uncertainty about Y, and vice versa. It can also be written in terms of Shannon entropies as:

$$I(X : Y) = H(X) + H(Y) - H(X, Y) \tag{2.11}$$

Mutual information is always non-negative, and equals zero if and only if X and Y are independent, that is, if their joint distribution factors as $P(x, y) = P(x)P(y)$. In the quantum scenario, mutual information is defined analogously, but using the von Neumann entropy instead of Shannon entropy. For a bipartite quantum system described again by a joint density matrix ρ_{AB}, the **quantum mutual information** is given by:

$$I(A : B) = S(A) + S(B) - S(AB) \tag{2.12}$$

Just like in the classical case, quantum mutual information quantifies the total correlations between subsystems, but in this setting, it captures both *classical correlations* and uniquely *quantum correlations*, such as entanglement. It is a central quantity in quantum information theory, playing key roles in communication, thermodynamics, and entanglement theory.

2.2.3 Conditional Mutual Information

In classical information theory, the **conditional mutual information** between three random variables X, Y, and Z is defined as:

$$I(X : Y \mid Z) = H(X \mid Z) - H(X \mid Y, Z) \tag{2.13}$$

where $H(X \mid Z)$ is the conditional Shannon entropy. This expression quantifies how much knowing Y helps reduce the uncertainty about X beyond what is already known from Z. An equivalent expression in terms of entropy is:

$$I(X : Y \mid Z) = H(X, Z) + H(Y, Z) - H(Z) - H(X, Y, Z) \tag{2.14}$$

this quantity plays a key role in understanding conditional dependencies. For example, in a Markov chain $X \to Y \to Z$, we have:

$$I(X; Z \mid Y) = 0$$

This indicates that once Y is known X provides no additional information about Z, encapsulating the essence of the Markov property where the future is independent of the past given the present. In this case, however, it reflects more of a spatial version of the concept.

In the quantum setting, for a tripartite quantum state described by ρ_{ABC}, the **quantum conditional mutual information** is defined as:

$$I(A : B \mid C) = S(AC) + S(BC) - S(ABC) - S(C) \tag{2.15}$$

This quantifies how much information subsystems A and B share, given knowledge of subsystem C. It captures both classical and quantum correlations in the system and quantifies its deviation from being Markovian.

Conditional mutual information in quantum systems is also fundamental in defining certain entanglement measures. The most prominent example is the **squashed entanglement** [2] , defined for a bipartite quantum state ρ_{AB} as:

$$E_{\text{sq}}(\rho_{AB}) = \frac{1}{2} \inf_{\rho_{ABC}} I(A : B \mid C) \tag{2.16}$$

where the infimum is over all extensions ρ_{ABC} such that $\mathrm{Tr}_C(\rho_{ABC}) = \rho_{AB}$. Thus, CMI serves not just as a diagnostic for Markovianity, but also as a powerful tool in quantifying entanglement.

We end this section stating two important relations that we will largely exploit later, namely:

$$I(A:B|C) = I(A:BC) - I(A:C) \tag{2.17}$$

$$I(A:B|C) = I(B:AC) - I(B:C) \tag{2.18}$$

2.2.4 Subadditivity and Strong Subadditivity Theorems

In the quantum setting, the *von Neumann entropy*, shares many properties with the *Shannon entropy*, among its most important structural properties are **subadditivity** and **strong subadditivity**, both of which express fundamental constraints on the entropic content of multipartite quantum systems.

Theorem 8 (Subadditivity)**.** *For any bipartite quantum state ρ_{AB} on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, the following inequality holds:*

$$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$$

with equality if and only if $\rho_{AB} = \rho_A \otimes \rho_B$, i.e., the state is a product state.

Interpretation: Subadditivity tells us that the entropy of the joint system is less than or equal to the sum of the entropies of its parts. If the two subsystems are correlated, the joint entropy will be strictly less than the sum. This theorem also implies that:

$$I(A:B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \geq 0$$

Theorem 9 (Strong Subadditivity)**.** *For any tripartite quantum state ρ_{ABC}, the following inequality holds:*

$$S(\rho_{ABC}) + S(\rho_C) \leq S(\rho_{AC}) + S(\rho_{BC})$$

This can be rewritten in terms of conditional mutual information:

$$I(A:B \mid C) = S(AC) + S(BC) - S(ABC) - S(C) \geq 0 \tag{2.19}$$

Interpretation: Strong subadditivity is a profound and nontrivial property of quantum entropy. It tells us that the conditional mutual information $I(A : B \mid C)$ is always nonnegative, meaning that our uncertainty about A when C and B are known is not more than when only C is known. SSA captures this balance quantitatively and reveals how information cannot be freely shared across multiple systems, a feature that is fundamental to quantum mechanics and has deep implications for quantum communication, cryptography, and the structure of multipartite entanglement. Notice that equality holds in the relation (2.19) if and only if the tripartite state forms a quantum Markov chain: $A \to C \to B$.

2.2.5 Carlen-Lieb Extension of Strong Subadditivity Theorem

A significant advancement in understanding quantum entropies is provided by Carlen and Lieb, who introduced an important extension of the strong subadditivity theorem [1]. Their result refines the conventional strong subadditivity inequality by presenting a sharper lower bound involving explicitly entropic differences. Specifically, their theorem is stated as follows.

Theorem 10 (Extended Strong Subadditivity [1])**.** *For any tripartite quantum state ρ_{ABC}, the following strengthened inequality holds:*

$$I(A : B \mid C) \geq 2 \max\{S(\rho_A) - S(\rho_{AB}), S(\rho_B) - S(\rho_{AB}), 0\}. \tag{2.20}$$

Interpretation: The Carlen-Lieb extension offers a deeper insight into the structure of quantum correlations by connecting conditional mutual information directly to entropic differences among subsystems. Classically, the entropy differences $S(A) - S(AB)$ and $S(B) - S(AB)$ are always non-positive, and hence do not yield additional insights. However, in the quantum setting, these conditional entropy differences can become strictly positive, indicating a quantum regime characterized by entanglement.

Example:

Consider the tripartite state

$$|\psi\rangle_{ABC} = \frac{1}{\sqrt{2}}(|000\rangle + |110\rangle)$$

We compute two key quantities: $S(A) - S(AB)$, and the conditional mutual information $I(A : B|C)$, to illustrate the Carlen–Lieb inequality. First, we observe that tracing out system C leaves AB in the pure Bell state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle+|11\rangle)$, so $S(AB) = 0$. The reduced state of system A is maximally mixed, $\rho_A = \frac{1}{2}I$, and thus $S(A) = 1$. Therefore:

$$S(A) - S(AB) = 1.$$

This quantity is positive, which is a distinctly quantum phenomenon: in classical information theory this wouldn't be possible, signaling the presence of purely quantum correlations. Next, we compute the conditional mutual information. Since the global state is pure, $S(ABC) = 0$. The state ρ_C is also pure, thus $S(C) = 0$. The reduced states ρ_{AC} and ρ_{BC} are each mixtures:

$$\rho_{AC} = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|), \quad \rho_{BC} = \frac{1}{2}(|00\rangle\langle 00| + |10\rangle\langle 10|),$$

and both have entropy $S = 1$. Therefore,

$$I(A : B|C) = 1 + 1 - 0 - 0 = 2.$$

We can conclude that in this case the bound provides extra information beyond that of the standard strong subadditivity theorem.

An important implication of this extended inequality is related to entanglement measures. Carlen and Lieb [1] derived the following bound on squashed entanglement E_{sq}, demonstrating its fundamental relation to entropy differences:

$$E_{sq}(\rho_{AB}) \geq \max\{S(\rho_A) - S(\rho_{AB}), S(\rho_B) - S(\rho_{AB}), 0\} \tag{2.21}$$

This inequality asserts that significant deviations in entropy between the total and individual subsystems necessarily indicate a non-trivial quantum entanglement.

2.3 Lower Bound on N-Partite Mutual Information

We now discuss the behavior of a full quantum system, which we consider to be divided into two subsystems, A and B (these may also represent separate regions that do not necessarily sum up to the entire system). The mutual information between two subsystems, denoted by $I(A : B)$, has already been defined. In 2008, M. M. Wolf *et al.* [5] found an analytical lower bound for this quantity that depends only on the average values of observables. Specifically, considering two observables, M_A, acting on subsystem A, and M_B, acting on subsystem B, it can be shown that:

$$I(A : B) \geq \frac{C(M_A, M_B)^2}{2\|M_A\|^2\|M_B\|^2} \tag{2.22}$$

where $C(M_A, M_B) = \langle M_A \otimes M_B \rangle - \langle M_A \rangle \langle M_B \rangle$ is the correlation function between the two observables. This relation highlights the usefulness of mutual information $I(A : B)$ as a powerful measure of correlations, surpassing traditional correlation functions $C(M_A, M_B)$. Specifically, the inequality demonstrates that mutual information captures all correlations, ensuring that even weak correlations are not overlooked, unlike simpler measures, which may fail to detect quantum correlations.

Following this relation, the authors emphasize that if mutual information decays exponentially, traditional correlation functions must also decay exponentially. Therefore, mutual information serves as a reliable tool for defining correlation lengths in quantum systems. This robustness makes mutual information particularly relevant for the study of area laws, as the existence of a finite correlation length (quantified through mutual information) rigorously implies an area-law scaling of the system's entropy. In this way, the relation effectively connects local, observable-based correlations with a global information-theoretic measure, offering deeper insights into complex many-body states.

Starting from this result, we have worked to generalize it to the case of **N-partite mutual information**, defined as:

$$I(A_1 : A_2 : \ \ldots \ : A_N) = S(A_1) + S(A_2) + \ldots + S(A_N) - S(A_1 A_2 \ldots A_N) \tag{2.23}$$

and we found the relation:

$$I(A_1 : A_2 : \ \ldots \ : A_N) \geq \frac{\left[\left\langle \bigotimes_{i=1}^{N} M_{A_i} \right\rangle - \prod_{i=1}^{N} \langle M_{A_i} \rangle \right]^2}{2 \prod_{i=1}^{N} \|M_{A_i}\|^2} \tag{2.24}$$

Proof. It is a well known fact that, in the context of quantum information, mutual information between different subsystems can be expressed as a relative entropy:

$$I(A_1 : A_2 : \ \ldots \ : A_N) = S(\rho_{a_1 a_2 \ldots a_N} \| \rho_{a_1} \otimes \rho_{a_2} \otimes \ldots \otimes \rho_{a_N})$$

Here, ρ_{a_i} with $i \in [1, N]$ is the reduced density matrix describing subsystem A_i, obtained by tracing out all other subsystems. We now exploit Pinsker's inequality, which in this case reads:

$$S(\rho_{a_1 a_2 \ldots a_N} \| \rho_{a_1} \otimes \rho_{a_2} \otimes \ldots \otimes \rho_{a_N}) \geq \frac{1}{2} \| \rho_{a_1 a_2 \ldots a_N} - \rho_{a_1} \otimes \rho_{a_2} \otimes \ldots \otimes \rho_{a_N} \|_1^2$$

Since this relation involves the trace norm $\|\ldots\|_1$, we can apply the inequality $\|X\|_1 \geq \frac{\mathrm{Tr}[XY]}{\|Y\|}$, where we have the freedom to choose Y. In our derivation, using $Y = (M_{A_1} \otimes M_{A_2} \otimes \ldots \otimes M_{A_N})$, we find:

$$\begin{aligned} &\| \rho_{a_1 a_2 \ldots a_N} - \rho_{a_1} \otimes \rho_{a_2} \otimes \ldots \otimes \rho_{a_N} \|_1 \\ &\geq \frac{\mathrm{Tr}[(\rho_{a_1 a_2 \ldots a_N} - \rho_{a_1} \otimes \rho_{a_2} \otimes \ldots \otimes \rho_{a_N})(M_{A_1} \otimes M_{A_2} \otimes \ldots \otimes M_{A_N})]}{\|M_{A_1}\| \|M_{A_2}\| \cdot \ldots \cdot \|M_{A_N}\|} \\ &= \frac{\mathrm{Tr}[\rho_{a_1 a_2 \ldots a_N}(M_{A_1} \otimes M_{A_2} \otimes \ldots \otimes M_{A_N}) - \rho_{a_1} \otimes \rho_{a_2} \otimes \ldots \otimes \rho_{a_N}(M_{A_1} \otimes M_{A_2} \otimes \ldots \otimes M_{A_N})]}{\|M_{A_1}\| \|M_{A_2}\| \cdot \ldots \cdot \|M_{A_N}\|} \\ &= \frac{\langle M_{A_1} \otimes M_{A_2} \otimes \ldots \otimes M_{A_N} \rangle - \langle M_{A_1} \rangle \langle M_{A_2} \rangle \cdot \ldots \cdot \langle M_{A_N} \rangle}{\|M_{A_1}\| \|M_{A_2}\| \cdot \ldots \cdot \|M_{A_N}\|} \end{aligned} \tag{2.25}$$

Combining these two inequalities, we find:

$$I(A_1 : A_2 : \ \ldots \ : A_N) \geq \frac{(\langle M_{A_1} \otimes M_{A_2} \otimes \ldots \otimes M_{A_N} \rangle - \langle M_{A_1} \rangle \langle M_{A_2} \rangle \cdot \ldots \cdot \langle M_{A_N} \rangle)^2}{2(\|M_{A_1}\| \|M_{A_2}\| \cdot \ldots \cdot \|M_{A_N}\|)^2} \tag{2.26}$$

which is exactly expression (2.24). □

Therefore, even when the system is partitioned into more than two parts, an analogous lower bound to that presented in [5] holds for multipartite mutual information. This result establishes a meaningful connection between local observable-based correlations and global information-theoretic quantities, offering valuable insight into the structure of quantum correlations in complex many-body systems.

Chapter 3

Lower bounds on Conditional Mutual Information

We now consider the behavior of the conditional mutual information. Given a quantum system, we define three subsystems, A, B, and C, and aim to compute an analytical lower bound for $I(A : B|C)$. As explained in the *Information Theory* section, this quantity represents the amount of information that B has about A that is not already provided by C. In the quantum context, in particular, this quantity can be used to quantify entanglement between A and B. Our objective is to strengthen the strong subadditivity theorem by providing a lower bound that is positive and strictly greater than zero at least in certain cases. This would allow us to use the bound $I(A : B|C) \geq \max(0, \mathrm{LB})$ rather than relying solely on the standard inequality $I(A : B|C) \geq 0$. Additionally, we aim to derive a lower bound that depends primarily, or ideally, exclusively, on observable quantities. Such a result would provide an effective tool for estimating entanglement in experimental or physical settings.

3.1 Trivial cases

We begin by identifying some trivial cases that do not require further analysis. These are quantum states described by ρ_{ABC} for which the conditional mutual information satisfies $I(A : B|C) = I(A : B)$. This can occur for two reasons:

- ρ_{ABC} **is a pure state**: In this case, a straightforward computation shows that $I(A : B|C) = I(A : B)$. This result follows from the property that, for any pure quantum state bipartitioned as $Z = X + Y$, the two resulting subsystems carry exactly the same amount of entropy, i.e., $S(X) = S(Y)$, regardless of subsystem characteristics such as dimension. In our case, using relation (2.15), we obtain:

$$\begin{aligned} I(A : B|C) &= S(AC) + S(BC) - S(ABC) - S(C) \\ &= S(B) + S(A) - S(AB) \end{aligned}$$

where we have also used the fact that $S(ABC) = 0$ for a pure state.

- **The state is a product between AB and C:** $\rho_{ABC} = \rho_{AB} \otimes \rho_C$: In this case, we can use the fact that the entropy of a product state is additive:

  $$\begin{aligned} S(\rho_X \otimes \rho_Y) &= -\operatorname{Tr}_{XY}[(\rho_X \otimes \rho_Y) \log(\rho_X \otimes \rho_Y)] \\ &= -\operatorname{Tr}_X[\rho_X \log(\rho_X)] - \operatorname{Tr}_Y[\rho_Y \log(\rho_Y)] \\ &= S(X) + S(Y) \end{aligned}$$

  Applying this property to our case yields $S(ABC) = S(AB) + S(C)$, which, when substituted into relation (2.15), again leads to $I(A : B|C) = I(A : B)$, for every choice of ρ_{AB} mixed or pure.

In both of these cases, we can trivially conclude that the lower bound for the conditional mutual information reduces to the one given in (2.22). Thus, we can write:

$$I(A : B|C) \geq \frac{C(M_A, M_B)^2}{2\|M_A\|^2\|M_B\|^2} \tag{3.1}$$

3.2 Non-Trivial Cases

We now seek to derive a non-trivial lower bound that is useful for highly correlated states, namely, non-product mixed states. To this end, we start by exploiting relations (2.17) and (2.18) to write:

$$I(A : B|C) = \frac{1}{2}\left(I(A : BC) - I(A : C) + I(B : AC) - I(B : C)\right) \tag{3.2}$$

It is well known that tracing out any degrees of freedom can only reduce mutual information. That is, given a purified state ρ_{ABCD} such that $\rho_{ABC} = \operatorname{Tr}_D[\rho_{ABCD}]$, we always have:

$$I(A : C) \leq I(AD : C) \tag{3.3}$$

Now, focusing on the negative terms in equation (3.2), we can write:

$$\begin{aligned}
-[I(A : C) + I(B : C)] &\geq -[I(AD : C) + I(B : C)] \\
&= -[S(AD) + S(C) - S(ADC) + S(B) + S(C) - S(BC)] \\
&= -[S(BC) + S(C) - S(B) + S(B) + S(C) - S(BC)] \\
&= -2S(C)
\end{aligned}$$

Substituting this back into equation (3.2), we obtain:

$$I(A : B|C) \geq \frac{1}{2} \left(I(A : BC) + I(B : AC) - 2S(C)\right) \tag{3.4}$$

Next, by applying the lower bound from equation (2.22) to both $I(A : BC)$ and $I(B : AC)$, we arrive at:

$$I(A : B|C) \geq \frac{1}{2} \left(\frac{C(M_A, M_{BC})^2}{2\|M_A\|^2\|M_{BC}\|^2} + \frac{C(M_{AC}, M_B)^2}{2\|M_{AC}\|^2\|M_B\|^2} - 2S(C) \right) \tag{3.5}$$

Where M_{AC} and M_{BC} are generic operators on systems ρ_{AC} and on ρ_{BC}, that can be defined as:

$$M_{AC} = \sum_i r_i (M_{A_i} \otimes M_{C_i})$$

$$M_{BC} = \sum_i s_i (M_{B_i} \otimes M_{C_i})$$

Defining the right-hand side of the inequality (3.5) as $LB(M_A, M_B, M_{AC}, M_{BC}, S(C))$, we can state the following result:

Theorem 11. *Given a quantum system partitioned into three subsystems A, B, and C, the following inequality always holds:*

$$I(A : B|C) \geq \max(0, LB(M_A, M_B, M_{AC}, M_{BC}, S(C)))$$

3.3 Testing the Effectiveness of the Lower Bound

We now aim to determine whether there exist actual non-trivial quantum states for which LB is positive, indicating that it serves as a useful lower bound. It is easy to observe that this behavior is governed by the competition between the two positive, observable-dependent terms and the negative entropic term in equation (3.5). To better characterize the behavior of different quantum states, we divide them into distinct categories and analyze each separately. In the following we consider states to be simple qubit states, doing so we will work with an 8 dimensional Hilbert space for the full 3-qubits system A, B and C.

- **Almost Product States**: In this case, we define the full quantum state as:

$$\rho_{ABC} = P(\rho_{AB} \otimes |0\rangle\langle 0|) + (1-P)(\rho_{AB} \otimes |1\rangle\langle 1|) \tag{3.6}$$

where P is a probability. To qualify as an almost product state, P must be close to either 0 or 1. We also choose ρ_{AB} to be a randomly generated mixed state, so that the full state ρ_{ABC} is mixed as well, thus possessing non-trivial quantum correlations. C instead is in a simple known state like $|0\rangle\langle 0|$ or $|1\rangle\langle 1|$. These are low entropy states. For a fixed value of P, we generate approximately ten thousands such states and examine how many of them yield a positive value for LB (see Fig. 3.1).

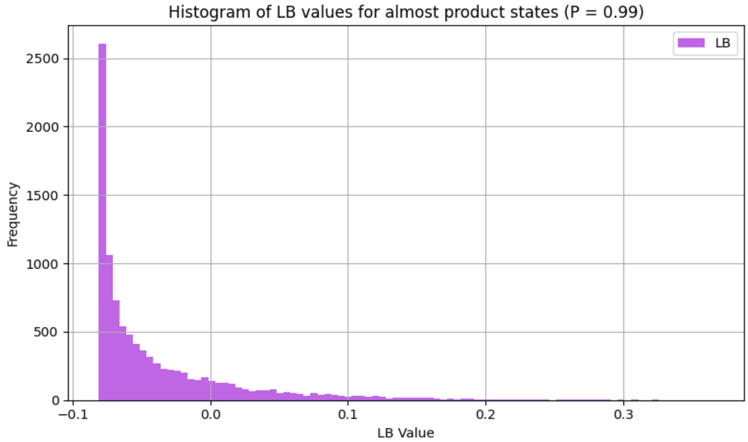

Figure 3.1: Distribution of LB values for almost product states

It is clear that for these simple states, the value of $S(C)$ is entirely determined by P. Indeed, we have:

$$\rho_C = P\,|0\rangle\,\langle 0| + (1-P)\,|1\rangle\,\langle 1|$$

which yields:

$$S(C) = -P\log(P) - (1-P)\log(1-P)$$

This entropy acts as a constant shift applied to the sum of the two positive terms in equation (3.5). If this shift is not large enough to overpower the contributions from the observables, the bound LB remains positive, resulting in a meaningful lower bound. Fixing $P = 0.99$, this occurs for approximately **20%** of the generated states. In Fig. 3.2, we can see the validity of our inequality for some of these states. Each point represents a state, with its $I(A:B|C)$ value plotted against its corresponding (positive) LB value. The diagonal line indicates where $I(A:B|C) = LB$ exactly. Here we can clearly observe that these quantum states deviate from a purely Markovian system, and our lower bound serves as a clear indicator of this behavior.

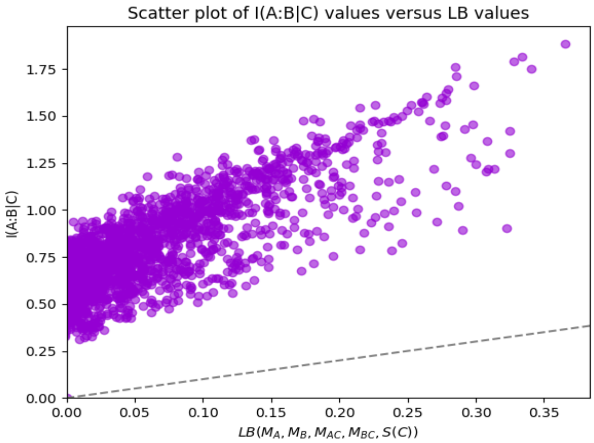

Figure 3.2: $I(A:B|C)$ vs. LB values for almost product states with $P = 0.99$

Based on these results, we conclude that for almost product states of the form considered here, the bound given by equation (3.5) provides a valid refinement of the strong subadditivity theorem in a modest but significant portion of random states.

- **Random-P States**: In this case, we still use a fixed structure for subsystem

C, and the full state retains the same form as before:

$$\rho_{ABC} = P(\rho_{AB} \otimes |0\rangle \langle 0|) + (1 - P)(\rho_{AB} \otimes |1\rangle \langle 1|)$$

However, unlike in the previous analysis, we now allow P to vary across its entire range, $P \in [0, 1]$. This means that the entropy term $S(C)$ will vary from state to state ranging from zero up to its maximum at $P = \frac{1}{2}$, corresponding to a maximally mixed state. To better characterize the behavior of the bound $LB(M_A, M_B, M_{AC}, M_{BC}, S(C))$, we do not sample P uniformly. Instead, we generate P values such that the resulting distribution of $S(C)$ is uniform. Naturally, in this setting we still encounter states for which LB is positive—simply because, with sufficiently high probability, P will occasionally be near 0.99, effectively reproducing the "almost product" scenario.

What is more interesting in this case is to analyze which states yield a positive LB, especially with respect to parameters that were previously fixed, namely, the entropy $S(C)$ and the purity of the full state.

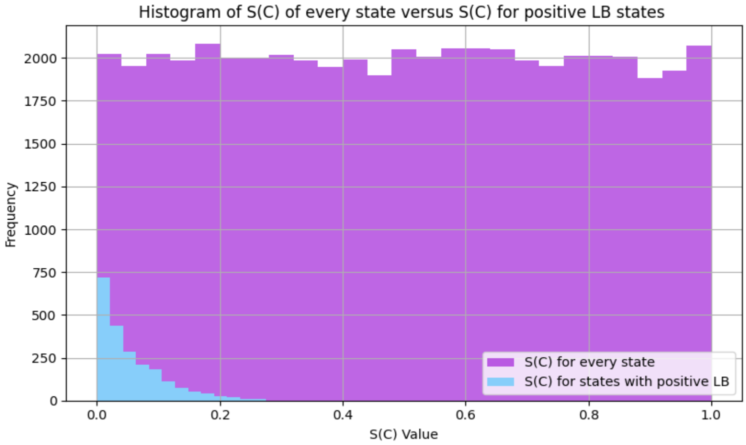

Figure 3.3: S(C) values for all states and for positive LB states

In Fig. 3.3, we show that the distribution of $S(C)$ is indeed uniform, while the subset of states for which LB is positive is mostly localized towards low $S(C)$.

In Fig. 3.4, we present the purity distribution for the same set of states. As expected, those with positive LB tend to have higher purity values.

To produce these figures, we simulated approximately 5×10^3 random quantum

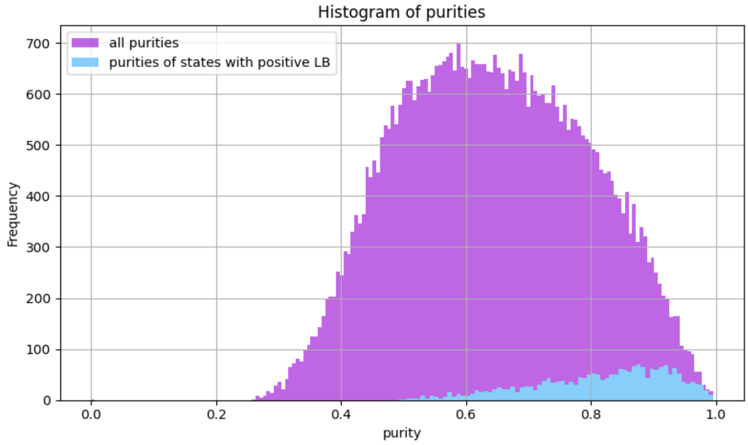

Figure 3.4: Purity values for all states and for positive LB states

states of the form given in equation (3.6), with P randomly generated to ensure uniformity in $S(C)$. We found that approximately **5%** of the generated states exhibited a strictly positive lower bound. Thus, we conclude that Theorem (11) provides a non-trivial refinement of the strong subadditivity inequality in a small but significant number of cases.

- **Completely Random States**: In this final case, we consider fully random and mixed states, constructed as follows:

$$\rho_{ABC} = P(\rho_{ABC}) + (1 - P)(\sigma_{ABC}) \tag{3.7}$$

where both ρ_{ABC} and σ_{ABC} are pure tripartite states. Using these two pure states, we generate a mixed state as done previously, with P acting as a mixing parameter. In this scenario, the entropy $S(C)$ depends not only on the value of P but also on the randomly sampled reduced state ρ_C. As a result, the influence of P on determining the positivity of LB is significantly diminished.

We conducted simulations similar to those in the previous sections, now using a Dirichlet distribution for P:

$$f(P) = \frac{1}{\mathrm{B}(\alpha_1, \alpha_2)} \cdot P^{\alpha_1 - 1} \cdot (1 - P)^{\alpha_2 - 1} \tag{3.8}$$

where:

$$\mathrm{B}(\alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1) \cdot \Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)}$$

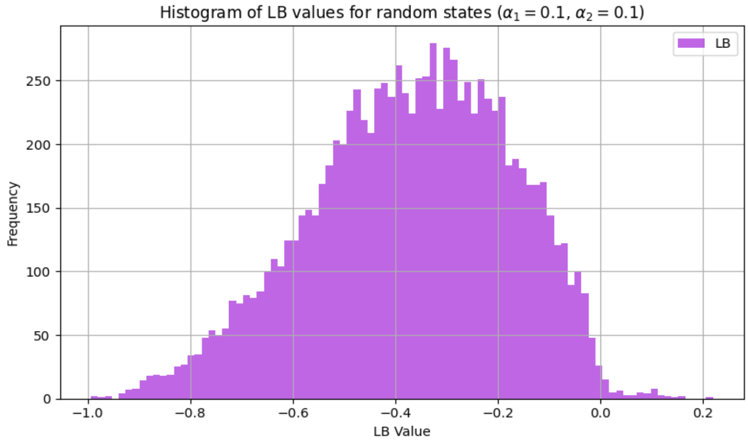

Figure 3.5: LB values for $\alpha_1 = \alpha_2 = 0.1$

In Fig. 3.5, we show the LB values for the case $\alpha_1 = \alpha_2 = 0.1$. Here, the distribution of P is strongly skewed toward the extremes, favoring high-purity (though still mixed) states. Under these conditions, we found that approximately **0.4%** of the generated states exhibit a positive value for LB.

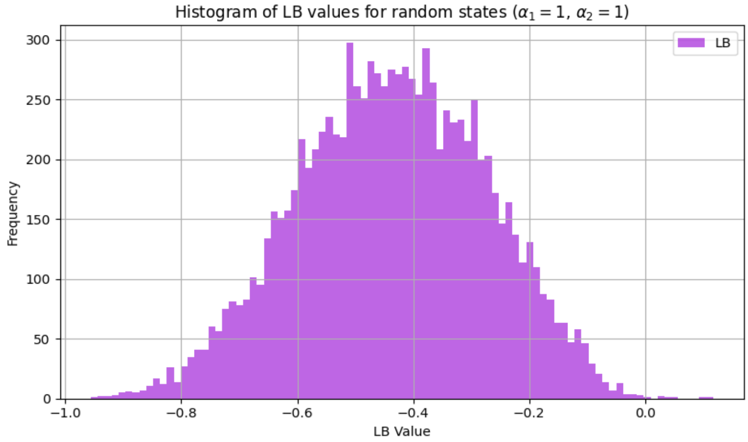

Figure 3.6: LB values for $\alpha_1 = \alpha_2 = 1$

In Fig. 3.6, we set both α_1 and α_2 to 1, resulting in a uniform distribution for P. This case represents the most general scenario. Here, we observed that approximately **0.02%** of the sampled states yield a positive value of LB.

These numerical investigations demonstrate that the proposed lower bound:

$$LB(M_A, M_B, M_{AC}, M_{BC}, S(C)) = \frac{1}{2}\left(\frac{C(M_A, M_{BC})^2}{2\|M_A\|^2\|M_{BC}\|^2} + \frac{C(M_{AC}, M_B)^2}{2\|M_{AC}\|^2\|M_B\|^2} - 2S(C)\right)$$

offers a meaningful refinement of the strong subadditivity inequality in several non-trivial scenarios. For *almost product states*, where the entropy $S(C)$ is minimal and well-controlled, the bound is positive in approximately **20%** of randomly generated cases, clearly validating its effectiveness in detecting correlations beyond the trivial regime. When generalizing to *Random-P* states with uniformly distributed entropies, the positivity rate drops to around **5%**, but still reveals a clear preference for states with lower $S(C)$ and higher purity, confirming that the bound remains informative in broader settings. Finally, for *completely random mixed states*, constructed via convex combinations of pure tripartite states, the bound remains positive in a small fraction of cases, about **0.4%** under skewed Dirichlet-distributed P, and around **0.02%** with uniform mixing, demonstrating that even in highly generic situations, the inequality captures non-trivial structure.

Chapter 4

Conditioning-System Independent Lower Bounds

In this section, we aim to exploit the Carlen–Lieb extension of the Strong Subadditivity theorem to find a lower bound on the conditional mutual information $I(A:B|C)$ that depends only on observable quantities over systems A and B. We arrive at the relation:

$$I(A:B \mid C) \geq \frac{C(M_A, M_B)^2}{2\|M_A\|^2\|M_B\|^2} + f(\langle V\rangle_{\rho_{AB}\otimes\rho_{AB}}, d) \tag{4.1}$$

where M_A and M_B are generic observables over systems A and B, V is the **Swap Operator** evaluated on the system AB and d is the dimension of that system. In this case the function $f(\langle V\rangle_{\rho_{AB}\otimes\rho_{AB}}, d)$ is defined as:

$$\begin{aligned} f(\langle V\rangle_{\rho_{AB}\otimes\rho_{AB}}) = \Big[&\Big(1-\frac{1}{d}-K(d,\langle V\rangle_{\rho_{AB}\otimes\rho_{AB}})\Big)\log\Big(1-\frac{1}{d}-K(d,\langle V\rangle_{\rho_{AB}\otimes\rho_{AB}}))\Big) \\ &+\Big(\frac{1}{d}+K(d,\langle V\rangle_{\rho_{AB}\otimes\rho_{AB}})\Big)\log\Big(\frac{1}{d}+K(d,\langle V\rangle_{\rho_{AB}\otimes\rho_{AB}})\Big)\Big] \end{aligned}$$

and

$$K(d,\langle V\rangle_{\rho_{AB}\otimes\rho_{AB}}) = \sqrt{\left(\frac{d-1}{d}\right)\left(\langle V\rangle_{\rho_{AB}\otimes\rho_{AB}}-\frac{1}{d}\right)}$$

Proof. We begin with the Carlen–Lieb inequality [1]:

$$I(A:B \mid C) \geq 2\max\{0, S(A)-S(AB), S(B)-S(AB)\}$$

from which we deduce:

$$\begin{aligned}\frac{I(A:B\mid C)}{2}&\geq 2\frac{S(A)-S(AB)}{2}\\ \frac{I(A:B\mid C)}{2}&\geq 2\frac{S(B)-S(AB)}{2}\end{aligned}$$

Adding the two, we obtain:

$$\begin{aligned}I(A:B\mid C)&\geq S(A)+S(B)-2S(AB)\\ &=I(A:B)-S(AB)\\ &=S(\rho_{AB}\,\|\,\rho_A\otimes\rho_B)-S(AB)\end{aligned}$$

We now apply the same inequality used in proving equation (2.24):

$$S(\rho_{AB}\,\|\,\rho_A\otimes\rho_B)\geq\frac{C(M_A,M_B)^2}{2\|M_A\|^2\|M_B\|^2},$$

As we can see, this term of the relation depends only on the observables M_A and M_B, so our remaining task is to express the entropic term $S(AB)$ in terms of observables as well. It has been shown [6] that this can be done finding a lower bound to $-S(AB)$, that depends only on the dimension of the system AB, which we call d, and the **Purity** γ of that system. In fact, one can write:

$$\begin{aligned}-S(AB)\geq\Bigg[&\left(1-\frac{1}{d}-\sqrt{\left(\frac{d-1}{d}\right)\left(\gamma-\frac{1}{d}\right)}\right)\log\left(1-\frac{1}{d}-\sqrt{\left(\frac{d-1}{d}\right)\left(\gamma-\frac{1}{d}\right)}\right)\\ &+\left(\frac{1}{d}+\sqrt{\left(\frac{d-1}{d}\right)\left(\gamma-\frac{1}{d}\right)}\right)\log\left(\frac{1}{d}+\sqrt{\left(\frac{d-1}{d}\right)\left(\gamma-\frac{1}{d}\right)}\right)\Bigg]\end{aligned}$$

For the final step we use the fact that the purity of a quantum state can be written as the expectation value of the Swap operator V. For a generic state ρ, we have:

$$\gamma=\mathrm{Tr}[\rho^2]=\mathrm{Tr}[V(\rho\otimes\rho)]=\langle V\rangle_{\rho\otimes\rho}. \tag{4.2}$$

(See Section 4.2.) Substituting this into the previous expression yields the bound in equation (4.1). □

Note: In all these derivations, we have used the Swap operator V, defined either in equation (4.8) or (4.7), depending on the context. In any case, it can be expressed as a sum of simple spin observables

4.1 Numerical Analysis

In this section, we aim to test whether the newly derived lower bounds for the conditional mutual information $I(A:B|C)$ remain non-negative for a relevant number of randomly generated mixed states. As before, we construct states by considering fully random and mixed combinations:

$$\rho_{ABC} = P(\eta_{ABC}) + (1-P)(\sigma_{ABC}), \tag{4.3}$$

where both η_{ABC} and σ_{ABC} are pure tripartite states. The values of P are drawn from a **Dirichlet distribution** (3.8). We vary the parameters α_1 and α_2 to sample states with different purity characteristics. For $\alpha_1 = \alpha_2 = \alpha = 1$, the P-distribution is uniform, while for $\alpha_1 = \alpha_2 = \alpha = 0.1$, the distribution is skewed toward the extremes, favoring higher purity. (In image 4.1, we show the purity distributions obtained through this method.)

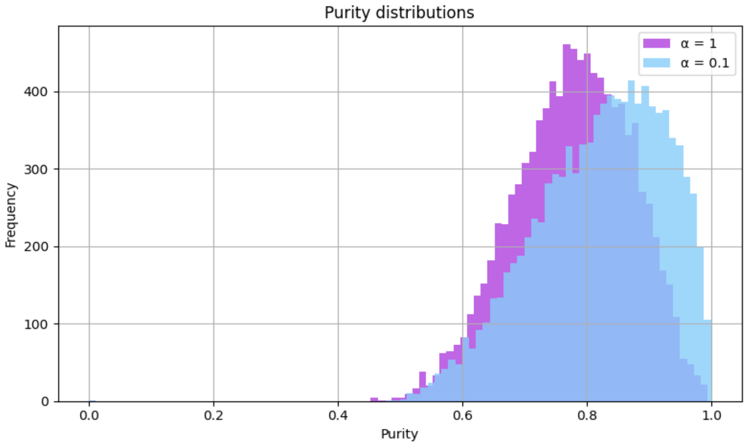

Figure 4.1: Purity distribution

We are now ready to test the bound derived in the previous section namely:

$$I(A:B \mid C) \geq \frac{C(M_A, M_B)^2}{2\|M_A\|^2\|M_B\|^2} + f(\langle V\rangle_{\rho_{AB}\otimes\rho_{AB}}, d) = LB(M_A, M_B, V), \tag{4.4}$$

For this analysis, we have used 3-qubit systems for ρ_{ABC}, consequently, the dimension d in these relations referring to ρ_{AB} is set to 2. In this specific case the function $f(\langle V\rangle_{\rho_{AB}\otimes\rho_{AB}}, d)$ takes the simpler form:

$$f(\gamma) = \left[\left(\frac{1}{2} - \sqrt{\frac{1}{2}\left(\gamma - \frac{1}{2}\right)}\right) \log\left(\frac{1}{2} - \sqrt{\frac{1}{2}\left(\gamma - \frac{1}{2}\right)}\right)\right. \tag{4.5}$$

$$\left. + \left(\frac{1}{2} + \sqrt{\frac{1}{2}\left(\gamma - \frac{1}{2}\right)}\right) \log\left(\frac{1}{2} + \sqrt{\frac{1}{2}\left(\gamma - \frac{1}{2}\right)}\right)\right] \tag{4.6}$$

Where, of course, we have again $\gamma = \langle V\rangle_{\rho_{AB}\otimes\rho_{AB}}$. Moreover, the set of quantum states over which the study was carried out comprised approximately 5000 states.

Results:

For $LB(M_A, M_B, V)$, using $M_A = M_B = \sigma_z$, we find that for $\alpha = 1$, **1.22%** of the states yield a positive bound, while for $\alpha = 0.1$, this increases to **5.1%**. The corresponding values are illustrated in image 4.2.

4.2 Swap Operators

We use this section to prove that we can find a swap operator V that satisfies the relation (4.2) for both 2 and 4 dimensional Hilbert space states.

2D Hilbert space: in this case $\rho \in \mathbb{C}^{2\times 2}$, then, V can be written in terms of Pauli operators as:

$$V = \frac{1}{2}(\mathbb{I}\otimes\mathbb{I} + \sigma_z\otimes\sigma_z + \sigma_x\otimes\sigma_x + \sigma_y\otimes\sigma_y) \tag{4.7}$$

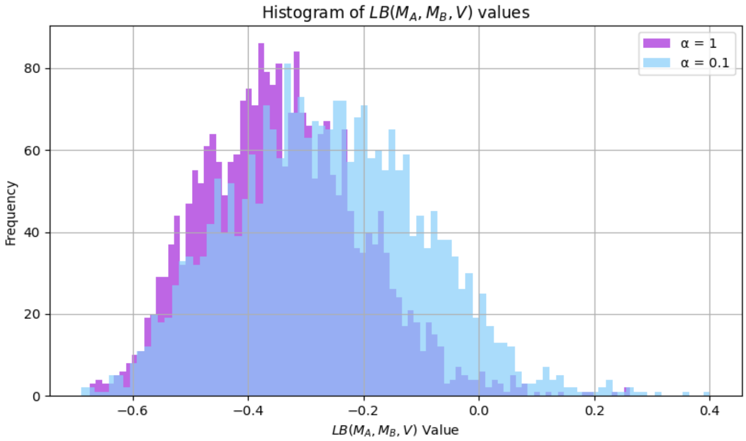

Figure 4.2: Mixed Observables Lower Bound Values

To show (4.2) we use the Bloch vector representation of the density matrix: $\rho = \frac{1}{2}(\mathbb{I} + \vec{r} \cdot \vec{\sigma})$:

$$\mathrm{Tr}[\rho^2] = \mathrm{Tr}\left[\frac{1}{4}(\mathbb{I} + 2\vec{r} \cdot \vec{\sigma} + (\vec{r} \cdot \vec{\sigma})^2)\right]$$

Now using the fact that:

- $(\vec{r} \cdot \vec{\sigma})^2 = \|\vec{r}\|^2 \mathbb{I}$
- $Tr[\vec{r} \cdot \vec{\sigma}] = Tr[r_x \sigma_x] + Tr[r_y \sigma_y] + Tr[r_z \sigma_z] = 0$

We find the relation:

$$\mathrm{Tr}[\rho^2] = \frac{1}{2}(1 + \|\vec{r}\|^2)$$

Going back to the expression for the Swap operator V we have:

$$\begin{aligned}\mathrm{Tr}[V(\rho \otimes \rho)] &= \frac{1}{2}\left(\mathrm{Tr}[(\mathbb{I} \otimes \mathbb{I})(\rho \otimes \rho)] + \mathrm{Tr}[(\sigma_x \otimes \sigma_x)(\rho \otimes \rho)]\right. \\ &\quad \left. + \mathrm{Tr}[(\sigma_y \otimes \sigma_y)(\rho \otimes \rho)] + \mathrm{Tr}[(\sigma_z \otimes \sigma_z)(\rho \otimes \rho)]\right) \\ &= \frac{1}{2}\left(\mathrm{Tr}[\rho] \cdot \mathrm{Tr}[\rho] + \mathrm{Tr}[\sigma_x \rho]^2 + \mathrm{Tr}[\sigma_y \rho]^2 + \mathrm{Tr}[\sigma_z \rho]^2\right) \\ &= \frac{1}{2}\left(1 + \mathrm{Tr}[\sigma_x \rho]^2 + \mathrm{Tr}[\sigma_y \rho]^2 + \mathrm{Tr}[\sigma_z \rho]^2\right)\end{aligned}$$

And since $Tr[\rho \sigma_i] = r_i \quad \forall i$:

$$\mathrm{Tr}[V(\rho \otimes \rho)] = \frac{1}{2}(1 + \|\vec{r}\|^2) = \mathrm{Tr}[\rho^2]$$

4D Hilbert Space: Let $\rho \in \mathbb{C}^{4\times 4}$ be the density matrix of a two-qubit system. Now V will be:

$$V = \frac{1}{4} \sum_{i,j\in\{0,x,y,z\}} \sigma_i \otimes \sigma_j \otimes \sigma_i \otimes \sigma_j \tag{4.8}$$

We aim to prove that also in this case:

$$\mathrm{Tr}[V(\rho \otimes \rho)] = \mathrm{Tr}[\rho^2].$$

To do this we start by noticing that any 2-qubit operator $\rho \in \mathbb{C}^{4\times 4}$ can be expanded in the Pauli basis as:

$$\rho = \sum_{i,j=0}^{3} r_{ij}\, \sigma_i \otimes \sigma_j,$$

where the coefficients are given by:

$$r_{ij} = \frac{1}{4}\mathrm{Tr}[\rho(\sigma_i \otimes \sigma_j)]$$

This is possible because the 16 operators $\sigma_i \otimes \sigma_j$ form an orthonormal basis for 2-qubit operators under the Hilbert-Schmidt inner product. (from now on $\sigma_0 = I$, $\sigma_1 = \sigma_x$, $\sigma_2 = \sigma_y$, $\sigma_3 = \sigma_z$)

Using the expansion, the tensor product $\rho \otimes \rho$ becomes:

$$\rho \otimes \rho = \sum_{i,j,k,l=0}^{3} r_{ij} r_{kl}\, \sigma_i \otimes \sigma_j \otimes \sigma_k \otimes \sigma_l$$

Now we compute:

$$\begin{aligned}\mathrm{Tr}[V(\rho \otimes \rho)] &= \frac{1}{4} \sum_{m,n=0}^{3} \mathrm{Tr}\left[(\sigma_m \otimes \sigma_n \otimes \sigma_m \otimes \sigma_n) \cdot \sum_{i,j,k,l=0}^{3} r_{ij} r_{kl}\, \sigma_i \otimes \sigma_j \otimes \sigma_k \otimes \sigma_l\right] \\ &= \frac{1}{4} \sum_{i,j,k,l=0}^{3} \sum_{m,n=0}^{3} r_{ij} r_{kl}\, \mathrm{Tr}\left[\sigma_m\sigma_i \otimes \sigma_n\sigma_j \otimes \sigma_m\sigma_k \otimes \sigma_n\sigma_l\right]\end{aligned}$$

Using trace factorization:

$$\mathrm{Tr}[A \otimes B \otimes C \otimes D] = \mathrm{Tr}[A]\,\mathrm{Tr}[B]\,\mathrm{Tr}[C]\,\mathrm{Tr}[D]$$

and the property of Pauli matrices:

$$\mathrm{Tr}[\sigma_a \sigma_b] = 2\delta_{ab}$$

we get:

$$\mathrm{Tr}\left[\sigma_m \sigma_i \otimes \sigma_n \sigma_j \otimes \sigma_m \sigma_k \otimes \sigma_n \sigma_l\right] = 2\delta_{mi} \cdot 2\delta_{nj} \cdot 2\delta_{mk} \cdot 2\delta_{nl} = 16\delta_{mi}\delta_{mk}\delta_{nj}\delta_{nl}$$

This enforces $m = i = k$ and $n = j = l$, so only terms with $i = k$, $j = l$ survive. Thus:

$$\mathrm{Tr}[V(\rho \otimes \rho)] = \frac{1}{4} \sum_{i,j=0}^{3} r_{ij}^2 \cdot 16 = 4 \sum_{i,j=0}^{3} r_{ij}^2 \tag{4.9}$$

Now we compute the purity using again the Pauli expansion:

$$\mathrm{Tr}[\rho^2] = \mathrm{Tr}\left[\left(\sum_{i,j=0}^{3} r_{ij} \sigma_i \otimes \sigma_j\right)^2\right] = \sum_{i,j,k,l=0}^{3} r_{ij} r_{kl}\,\mathrm{Tr}[(\sigma_i \sigma_k) \otimes (\sigma_j \sigma_l)]$$

We notice that only terms with $i = k$ and $j = l$ survive, giving:

$$\mathrm{Tr}[\rho^2] = \sum_{i,j=0}^{3} r_{ij}^2 \cdot \mathrm{Tr}[\sigma_i^2] \cdot \mathrm{Tr}[\sigma_j^2] = \sum_{i,j=0}^{3} r_{ij}^2 \cdot 2 \cdot 2 = 4 \sum_{i,j=0}^{3} r_{ij}^2$$

Which is exactly equal to (4.9), therefore,

$$\mathrm{Tr}[V(\rho \otimes \rho)] = \mathrm{Tr}[\rho^2].$$

4.3 Summary of results

In this final section, we summarize all the lower bounds to the conditional mutual information $I(A : B|C)$ derived in this chapter, along with the most significant results from our numerical tests. The bounds presented are:

- $$\begin{aligned} I(A : B|C) &\geq \frac{1}{2}\left(\frac{C(M_A, M_{BC})^2}{2\|M_A\|^2\|M_B\|^2\|M_C\|^2} + \frac{C(M_{AC}, M_B)^2}{2\|M_A\|^2\|M_B\|^2\|M_C\|^2} - 2S(C)\right) \\ &= LB(M_A, M_B, M_{AC}, M_{BC}, S(C)) \end{aligned}$$

- $$\begin{aligned} I(A : B \mid C) &\geq \frac{C(M_A, M_B)^2}{2\|M_A\|^2\|M_B\|^2} + f(\langle V\rangle_{\rho_{AB}\otimes\rho_{AB}}, d) \\ &= LB(M_A, M_B, V) \end{aligned}$$

The table below presents the percentage of randomly generated states (for different Dirichlet parameters α) for which each bound yields a positive value:

Bound	$\alpha = 1$	$\alpha = 0.1$
$LB(M_A, M_B, M_{AC}, M_{BC}, S(C))$	**0.02%**	**0.4%**
$LB(M_A, M_B, V)$	**1.2%**	**5.1%**

Table 4.1: Positivity test results for different α values and bounds

The results in Table 4.1 highlight how the effectiveness of each lower bound varies with the purity of the sampled states, controlled by the Dirichlet parameter α. An important feature of the latest bound, namely $LB(M_A, M_B, V)$, is that it is independent of the conditioning system C, and rely only on observable quantities over subsystems A and B. This allows for experimentally feasible estimation of quantum correlations.

A key implication of these results lies in their connection to entanglement measures. Since each of the proposed lower bounds provides a guaranteed floor for the conditional mutual information $I(A:B|C)$, they also translate into lower bounds for known entanglement quantifiers. Specifically, they imply:

$$E_{sq}(\rho_{AB}) \geq \max\{LB(M_A, M_B, M_{AC}, M_{BC}, S(C)),\ LB(M_A, M_B, V),\ 0\} \tag{4.10}$$

In summary, the lower bounds derived and tested in this chapter offer both theoretical insight and potential practical relevance. Their independence from the conditioning system in some cases, along with their formulation in terms of experimentally accessible quantities, suggests that they could serve as useful tools in quantum information settings, especially in situations where the system size is large, making the direct study of a density matrix inconvenient.

Bibliography

[1] Eric A Carlen and Elliott H Lieb. Bounds for entanglement via an extension of strong subadditivity of entropy. *Letters in mathematical physics*, 101:1–11, 2012.

[2] Matthias Christandl and Andreas Winter. “squashed entanglement”: an additive entanglement measure. *Journal of mathematical physics*, 45(3):829–840, 2004.

[3] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

[4] James V Stone. Information theory: a tutorial introduction. 2015.

[5] Michael M Wolf, Frank Verstraete, Matthew B Hastings, and J Ignacio Cirac. Area laws in quantum systems: mutual information and correlations. *Physical review letters*, 100(7):070502, 2008.

[6] Ting Zhang, Graeme Smith, John A Smolin, Lu Liu, Xu-Jie Peng, Qi Zhao, Davide Girolami, Xiongfeng Ma, Xiao Yuan, and He Lu. Quantification of entanglement and coherence with purity detection. *npj Quantum Information*, 10(1):60, 2024.