Improving the Security Analysis of Device-Independent Quantum Key Distribution via Bell Nonlocality and Semidefinite Programming

A Dissertation Submitted

in Partial Fulfilment of the Requirements

for the Degree of

BACHELOR OF TECHNOLOGY

in

Computer Science and Engineering

by

Sawan Bhattacharyya

Semester : 8

Roll Number : T91/CSE/216017

Registration Number : D01-1111-0236-22

Under the Joint Supervision of

Prof. Pankaj Agrawal and Prof. Sunirmal Khatua

to

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
UNIVERSITY OF CALCUTTA
KOLKATA, INDIA

June, 2025

Dedicated to Ma, Baba, Didi, Gullu, Raja da
and Prasenjit Sir

DECLARATION

I, **Sawan Bhattacharyya (Registration No: D01-1111-0236-22, Roll No: T91/CSE/216017)**, hereby declare that this report, submitted to the Department of Computer Science and Engineering, University of Calcutta, Kolkata, in partial fulfillment of the requirements for my **Bachelor of Technology** in **Computer Science and Engineering**, is my own original work. It was conducted under the joint guidance of Prof. Pankaj Agrawal, Centre for Quantum Engineering, Research and Education, TCG Centre for Research and Education in Science and Technology, Kolkata, and Prof. Sunirmal Khatua, Department of Computer Science and Engineering, University of Calcutta, Kolkata. I confirm that this work has not previously been submitted for any degree or diploma at this or any other institution. I have diligently upheld academic ethics and honesty throughout this report, ensuring that all external sources of information, statements, or results are properly acknowledged and cited.

Place: Kolkata, West Bengal, India
Date: June 2025

Sawan Bhattacharyya

CERTIFICATE

This is to certify that the thesis entitled **Improving the Security Analysis of Device-Independent Quantum Key Distribution via Bell Nonlocality and Semidefinite Programming** submitted by **Sawan Bhattacharyya** (**Registration No: D01-1111-0236-22, Roll No: T91/CSE/216017**) to the Department of Computer Science and Engineering, University of Calcutta, Kolkata, in partial fulfillment of the requirements for the degree of **Bachelor of Technology** in **Computer Science and Engineering**, is an original work carried out under my supervision. This thesis has not been submitted elsewhere for the award of any degree or diploma.

Date: June 12, 2025 **Kolkata, West Bengal, India**

....
Prof. Pankaj Agrawal
Supervisor
Centre for Quantum Engineering Research and Education
TCG Centre for Research and Education in Science and Technology

....
Prof. Sunirmal Khatua
Co-Supervisor
Department of Computer Science and Engineering
University of Calcutta

....
External Examiner(S)

....
Head of the Department
Department of Computer Science and Engineering
University of Calcutta

ACKNOWLEDGEMENT

It is with profound gratitude that I acknowledge the invaluable contributions of numerous individuals, without whom the completion of this thesis would not have been possible.

I extend my deepest appreciation to my supervisors, Prof. Pankaj Agrawal of TCG CREST and Prof. Sunirmal Khatua of the University of Calcutta. Their exceptional guidance, insightful feedback, and unwavering encouragement were instrumental throughout this work. Their valuable suggestions were of immense help, and their boundless enthusiasm to explore my ideas, no matter how unconventional, truly compelled me to approach the problem from entirely new perspectives.

My sincere thanks also go to Dr. Prasenjit Deb of TCG CREST. This work simply could not have been realized without his guidance. His motivation and mentorship, both within and beyond the project, ensured that I never experienced a moment of distress. His philosophy on research and life allowed me to genuinely enjoy every step of this journey.

I am equally grateful to Mr. Turbasu Chatterjee of Virginia Tech, whose unwavering dedication and meticulous care were crucial in achieving conceptual clarity and fostering out-of-the-box thinking. Our tireless efforts, often burning the midnight oil to cross-validate proofs and computational results, solidified my conceptual understanding and confidence like never before.

A special acknowledgement goes to Mr. Debdeep Mitra of the University of Calcutta. His energetic nature and constant words of motivation sustained me throughout this endeavor. Debdeep was a steadfast presence during my entire tenure at the University of Calcutta, and I am privileged to have him as an elder brother who consistently supported me in all technical and non-technical matters.

I will be forever thankful to all the members of TCG CREST for their support. Mr. Sudhindu Mandal for introducing me to Prof. Agrawal; Prof. Snehal Mandal for meticulously reviewing my non-technical content, including the thesis structure; Mr. Biplab Biswas for his guidance upon tailoring the thesis and Mr. Souvik Chatterjee for his invaluable assistance with mathematical code. I could not have reached this point without their contributions, present from the very genesis of this work.

Beyond academic collaborations, I am incredibly fortunate to my two best friends Anamika and Anushree, the only two true friends I have ever had. They were the pillar of support during my toughest times—when frustrations, anxiety, and fear peaked, when life seemed broken, and when I regretted choosing an academic career. I always found Anushree and Anamika to be the nearest and most reliable persons with whom I could share everything.

Finally, I dedicate this thesis to my father, mother, and elder sister, whose unconditional love and care have always been my greatest inspiration. I also express my profound gratitude and respect to my ideal, Rajdeep Mukherjee, who planted the seed of being a researcher and, more importantly, a curious student. As a true mentor, he inspired me to choose this path; I might never have considered a career in Computer Science had Rajdeep not been there.

Sawan Bhattacharyya
Univeristy Of Calcutta

ABSTRACT

Security analysis is the bedrock of any cryptographic protocol, classical or quantum, ensuring the confidentiality of distributed keys. Often, this analysis involves framing the problem as a challenging optimization task. In this article, we tackle this challenge within the context of Device-Independent Quantum Key Distribution (DIQKD) with a random key basis protocol.

Our work demonstrates a significant reduction in the computational cost of the existing security analysis without compromising the key rate. We achieve this by reframing the entire security analysis as a strongly convex optimization problem. This novel approach allows for a more efficient optimization of Bob's measurement angles, specifically for determining a lower bound on Eve's uncertainty about Alice's key generation basis. Unlike the original security proof, our method incurs a considerably lower optimization cost for this crucial step.

Furthermore, we provide an explicit form of the pessimistic error that emerges during the optimization of both parties' measurement angles. This detailed error analysis adds a layer of precision and rigor to the security proof. We also clarify several aspects of the original security proof, making the overall analysis more robust and complete. This refinement is vital for establishing the protocol's practical viability and trustworthiness.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Quantum key distribution (QKD) is one of the fundamental tasks in quantum communication and cryptography, which allows two remote users to set up a shared cryptographic key via an *insecure* quantum channel and an *authenticated* classical channel. An adversary can perform arbitrary quantum operations on the quantum states transmitted through the quantum channel. However, he cannot modify the messages sent through the classical channel – those messages can only be read by the adversary. Under a certain well-defined set of assumptions, the secret keys exchanged using a QKD protocol can be proven to be *information-theoretically secure.*

Traditional QKD protocols, like BB84, assume that the devices used by Alice and Bob are perfectly controlled and behave as expected. However, this assumption is not always realistic. Device-Independent QKD removes these assumptions and guarantees security based solely on the laws of quantum mechanics and the observed correlations in the measurement outcomes.

Essential requirements for any QKD system are Independence and privacy of the measurement settings, and Privacy of the measurement outcomes. In addition to these two essential requirements, a device-dependent QKD system also holds a third requirement, which is being eliminated by DIQKD, *Perfect control over the state preparation and measurement devices.* If the devices that implement the protocol do not behave as advertised, one may consider that the security of QKD has been compromised.

Quantum hacking by an adversary consists of attacks that cause the behaviour of the devices to deviate from the model that is used in the security proof. To rule out these scenarios of device-dependent QKD protocol (DD-QKD), a security proof was devised that is agnostic to the device modelling. The key distribution protocols that aim to prove their security without specifying the states and measurement devices used are referred to as device- independent QKD (DI-QKD)protocols. With device-independent security, all side-channels in the quantum communications layer can be eliminated conclusively and elegantly. This is achieved by using Bell non-locality to certify that the uncharacterised devices are producing genuinely random outputs to the adversary. Hence, Bell's non-locality is a necessary condition for DI-QKD's security – As long as some minimal assumptions are satisfied, the security of this protocol can be guaranteed based solely on the violation of a Bell inequality. The security of device-independent QKD has been rigorously studied, thereby emphasising that the DI-QKD protocol is perhaps the ultimate key exchange protocol in terms of security.

With the advent of this protocol, several loopholes and side-channels that make current QKD vulnerable can be closed. Even though DI-QKD, as an entanglement-based protocol, allows the users to realise a secure QKD with unknown and uncharacterized devices, the practical implementation of this protocol is technically challenging with the current technologies.

The genesis of Device-Independent Quantum Key Distribution can be traced back to the seminal work on using the Bell pair in Quantum Key Distribution by Ekert. In this work, we aim to study the three variants of the DI-QKD protocol, viz. DI-QKD with noisy pre-processing, random key basis, and random post-selection. In the context of DD-QKD, it is known that the robustness of some protocols to experimental imperfections, such as channel loss or noise, can be improved by randomly flipping some of the bits in the raw key before performing error correction and privacy amplification. This step is known as noisy pre-processing.

The reason for this is that the random flips would increase Eve's uncertainty about the raw key, though such random flips also reduce the correlation between Alice and Bob (hence increasing the error correction cost). It was shown that the same effect can be observed in DI-QKD. This noisy pre-processing yields significant improvements for the photonic implementation of the protocol In standard DI-QKD, Alice uses only one measurement setting for generating keys; the other two measurement settings are used to test the violation of the CHSH inequality. In such a case, Eve could focus on minimising her uncertainty about that fixed key-generating measurement at the expense of having higher uncertainty about the other measurement. To overcome this problem, the standard The DI-QKD protocol has been modified to one that uses both of Alice's measurements to generate the raw key. In this modified protocol, Bob needs to perform an additional measurement to obtain outcomes that are better correlated with Alice's second measurement. This process is known as a random key basis. In DD-QKD, post-selection is a common practice as photons are occasionally lost in the quantum channel. However, in DI-QKD, some care is needed when post-selection is employed, as discarding some events might open up the detection loophole. For example, when one naively discards the "no- Click" events, it is possible to violate the CHSH inequality using a classical strategy. To overcome this problem, if the users simply assign a deterministic outcome for rounds in which the detectors do not click, then it would decrease the amount of certified randomness produced by the measurement and consequently pose a challenge to the implementation of photonic DI-QKD. A DI-QKD protocol with a random post-selection step has been proposed to circumvent this problem. Recently, an experimental result has been reported showing the efficiency of DI-QKD with random post-selection. The two most important figures-of-merit that determine the experimental implementations of different DI-QKD are the critical quantum bit error rate and the critical detection efficiency. In this project, our goal will be to design a DI-QKD protocol to maximise the asymptotic key rate and/or to minimise the experimental requirements given in terms of the mentioned figure-of-merit. We have worked on the security analysis of device-independent quantum key distribution using random key basis[48] in Chapter 8, where we have derived a closed form of the pessimistic error needed for optimising Alice and Bob angles.

The next chapter introduces the mathematical prerequisites one would need to understand the upcoming chapters. It typically introduces the concept of *Complex Eucildean Space* and *Linear Operator* are widely used throughout the thesis. Chapter 3 focuses upon the basics of *Quantum Information Theory*, which extensively uses the concepts from the previous chapter. It also discusses various entropic measures as well as entropic inequalities, which are extensively used in the security analysis in Chapter 8. Chapter 4 focuses upon the concept of *Quantum Entanglement* and its various properties. It also discusses upon *Bell nonlocality*. The chapter ends with a note on the difference between *Classical and Quantum Correlation*. The next chapter, Chapter 5, discusses upon various Device Independent Quantum Key Distribution protocols discussed earlier, primarily focusing on the protocol setup, key results and key rate. The chapter concludes with a comparison between these protocols. Chapter 6 focuses upon the security proof of quantum key distribution, particularly highlighting the security definition [45] and various tools used for security analysis. It also shed light upon the role of semi-definite programming in security proof and hence set the motivation for the next chapter. The next chapter, Chapter 7, discusses various concepts related to semidefinite programming, including its standard form, duality and *Slater's condition.* Chapter 8 is all focused upon improving the security proof of DIQKD using Random Key Basis [48]. The next chapter, Chapter 9, presents the computational results of Bell-nonlocality using SDP. The thesis concludes with remarks upon the work and direction of future work.

Chapter 2

Mathematical Preliminaries

2.1 Complex Euclidean Space:

The notion of Complex Euclidean Space is well defined in the literature as the set of all functions from an alphabet set Σ to complex numbers $\mathbb{C}$.The set forms a vector space of dimension $|\Sigma|$ over the field of $\mathbb{C}$ is denoted as $\mathbb{C}^{\Sigma}$. The two operations of addition and scalar multiplication are defined in the following standard way:

1. Addition : Let $x, y \in \mathbb{C}^{\Sigma}$ then the vector $x + y \in \mathbb{C}^{\Sigma}$ is defined in the following standard way $(x + y)(a) = x(a) + y(a) \; \forall \; a \in \Sigma$

2. Scalar multiplication: Let $x \in \mathbb{C}^{\Sigma}$ and a scalar $\alpha \in \mathbb{C}$, then the vector $\alpha x \in \mathbb{C}^{\Sigma}$ is defined in the following standard way $(\alpha x)(a) = \alpha x(a) \; \forall a \in \Sigma$.

The alphabet is defined as a finite non-empty set whose elements are called as symbols. The alphabet is denoted using capital Greek letters, viz. Σ, Γ, Λ and the symbols are denoted using lower case roman letters, viz. a, b, c, d. The most common example of an alphabet is the binary alphabet $\{0, 1\}$ frequently used to define the vector space of a quantum bit or qubit. Other examples include n-fold Cartesian product of the binary alphabet with itself, typically used to define the vector space involving n qubits. Other examples of alphabet include $\{1, 2, 3, 4, ..., n\}$ n being a positive integer, which is typically used to define the vector space involving an n dimensional system. The function $u(a)$ is defined as $u : \Sigma \to \mathbb{C}$, the entry of the vector u indexed by a, for each $u \in \mathbb{C}^{\Sigma}$ and $a \in \Sigma$. The vector whose entries are all zero is simply denoted 0 refer to as zero vector.
Complex Euclidean spaces are typically denoted by scripted capital letters viz. $\mathcal{X}, \mathcal{Y}$, and $\mathcal{Z}$ and the vectors are denoted by lowercase Roman letters viz. u, v, w, x, y, and z. For a vector $u \in \mathbb{C}^n$ its often viewed as n-tuple $u = (\alpha_1, \ldots, \alpha_n)$ or as a column vector $u = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ where $\alpha_i \in \mathbb{C}$ for all $i \in \Sigma$.
One can always fix a bijection $f : \{1, \ldots, n\} \longrightarrow \Sigma$, for any alphabet Σ, the complex euclidean space $\mathbb{C}^{\Sigma}$ may be viewed as being equivalent to $\mathbb{C}^n$ for $n = |\Sigma|$ being the dimension of the space and connecting all vector $u \in \mathbb{C}^{\Sigma}$ with the vector in $\mathbb{C}^n$ whose k-th entry is $u(f(k))$.

2.1.1 Inner Products and Norms of Vectors:

Let us have two vectors $x, y \in \mathbb{C}^{\Sigma}$ then their *Inner product* $\langle x, y \rangle$ is defined as

$$\langle x, y \rangle = \sum_{a \in \Sigma} x(a)\overline{y(a)}. \tag{2.1.1}$$

Inner product of two vectors satisfy certain properties as,

- Linearity in the second argument: $\langle x, \beta y + \alpha w \rangle = \beta \langle x, y \rangle + \alpha \langle x, w \rangle \; \forall x, y, w \in \mathbb{C}^{\Sigma}$ and $\beta, \alpha \in \mathbb{C}$.

- Conjugate symmetry: $\langle x, y \rangle = \overline{\langle y, x \rangle}\ \forall x, y, w \in \mathbb{C}^{\Sigma}$
- Positive semi definiteness: $\langle x, x \rangle \geq 0\ \forall x \in \mathbb{C}^{\Sigma}$ with equality iff $x = 0$.

Any function satisfying these three properties is called as an inner product but in the context of quantum information inner product of vectors in complex euclidean space is primarily considered.
In the literature there is a class of p-norms, which are defined for each $x \in \mathbb{C}^{\Sigma}$ in the following standard form

$$\|x\|_p = \left(\sum_{a \in \Sigma} |x(a)|^p \right)^{\frac{1}{p}} \tag{2.1.2}$$

The classes of *p-norms* holds certain properties

- Positive semi definiteness : $\|x\| \geq 0\ \forall x \in \mathbb{C}^{\Sigma}$ with equality iff $x = 0$.
- Positive scalability : $\|\beta x\| = |\beta| \|x\|\ \forall x \in \mathbb{C}^{\Sigma}$ and $\beta \in \mathbb{C}$
- Triangle inequality: $\|x + y\| \leq \|x\| + \|y\|\ \forall x, y \in \mathbb{C}^{\Sigma}$

We will now consider two space cases when $p = 2$ and $p = \infty$.
Let us first consider the case when for for $p < \infty$ then we can defined the *max-norm* of the given vector as

$$\|x\|_\infty = \max\{|x(a)| : a \in \Sigma\}. \tag{2.1.3}$$

It essentially gives the maximum of the absolute values of all entry, in our case it gives the maximum modulo among all complex entries.
Now we will defined the case when $p = 2$ which is widely known an *Euclidean norm* which is defined for a vector $x \in \mathbb{C}^{\Sigma}$ as:

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{\sum_{a \in \Sigma} |x(a)|^2}. \tag{2.1.4}$$

The above three norm properties (positive definiteness, positive scalability, and the triangle inequality) hold for $\| \cdot \|$ replaced by $\| \cdot \|_p$ for any choice of $p \in [1, \infty]$.

2.1.2 The Cauchy-Schwarz inequality:

The Cauchy-Schwarz inequality defines a relation between inner product of two vectors and the product of their norms.

$$|\langle x, y \rangle| \leq \|x\| \|y\|. \tag{2.1.5}$$

$\forall x, y \in \mathbb{C}^{\Sigma}$, the equality condition hold only when u and v are linearly dependent.
The collection of all those unit vector in a complex euclidean space $\mathcal{X}$ forms a *unit sphere* in $\mathcal{X}$ and denoted in the standard form in the literature as

$$\mathcal{S}(\mathcal{X}) = \{u \in \mathcal{X} : \|u\| = 1\} \tag{2.1.6}$$

2.1.3 Orthogonality and Orthonormality:

In order to define the the concept of *Orthogonality and Orthonormality* we need to first define the concept of two orthogonal vectors. Let us have two vectors $x, y \in \mathbb{C}^{\Sigma}$ will be called as orthogonal if $\langle x, y \rangle = 0$. This property signifies that the two vectors are perpendicular in the context of the complex Euclidean space, meaning they do not share any component in a given direction. An alternative notation often used to express this orthogonality relation is $x \perp y$, which explicitly denotes that x and y are mutually orthogonal.
In general for any set $\mathcal{M} \subseteq \mathbb{C}^{\Sigma}$ we will say that a vector x is orthogonal to $\mathcal{M}$ if u is orthogonal with every vector in $\mathcal{M}$. A collection of vector indexed by some alphabet Λ, will be called as orthogonal set if for any two distinct vectors their inner product is 0. Mathematically we can formulate it as

$$\langle x_a, y_b \rangle = 0 \tag{2.1.7}$$

for all choices of $a, b \in \Lambda$ with $a \neq b$. This definition establishes that each vector in the set is mutually perpendicular to the others. Additionally, a crucial property of an orthogonal collection of nonzero vectors is that it inherently possesses linear independence, meaning that no vector in the set can be written as a linear combination of the others. When an orthogonal set consists exclusively of unit vectors, it is classified as an orthonormal set. Moreover, if such a set spans the entire space, it is termed an *orthonormal basis.* A fundamental criterion for an *orthonormal set* $\{x_b : b \in \Lambda\}$ to be considered a basis for the space $\mathbb{C}^\Sigma$ is that its cardinality must match the dimension of the space, i.e., $|\Lambda| = |\Sigma|$.

A well-known example of an orthonormal basis in $\mathbb{C}^\Sigma$ is the standard basis, which consists of the vectors $\{e_b : b \in \Sigma\}$, where each standard basis vector e_b is defined component-wise as

$$e_b(c) = \begin{cases} 1, & \text{if } b = c \\ 0, & \text{otherwise} \end{cases} \tag{2.1.8}$$

for all $b, c \in \Sigma$. This definition ensures that each basis vector is aligned with a single coordinate axis in the space.

2.2 Linear Operator:

Given two complex Euclidean spaces $\mathcal{P}$ and $\mathcal{Q}$, the notation $L(\mathcal{P}, \mathcal{Q})$ represents the collection of all linear mappings that take elements from $\mathcal{P}$ and transform them into elements of $\mathcal{Q}$. More formally, the function satisfies the properties of linearity belongs to this set:

$$L : \mathcal{P} \longrightarrow \mathcal{Q}. \tag{2.2.1}$$

Such mappings are commonly referred to as linear operators, or simply operators, when the context is clear. When representing the action of a linear operator on a vector $x \in \mathcal{P}$, parentheses are typically omitted in notation whenever there is no ambiguity. That is, instead of writing $L(x)$, it is standard practice to write Lx to denote the vector obtained by applying the operator L to x. This convention simplifies expressions and enhances readability in algebraic manipulations. The collection of all such linear operators $L(\mathcal{P}, \mathcal{Q})$ forms a vector space itself, equipped with operations of vector addition and scalar multiplication, defined as follows:

1. Addition: Given two operators $L, M \in L(\mathcal{P}, \mathcal{Q})$, their sum $L + M$ is also an operator in $L(\mathcal{P}, \mathcal{Q})$ and is defined by applying each operator separately and then summing the results:

$$(L + M)x = Lx + Mx, \quad \forall x \in \mathcal{P}. \tag{2.2.2}$$

   This property ensures that the space of linear operators is closed under addition, making it a well-defined vector space.

2. Scalar Multiplication: Given an operator $L \in L(\mathcal{P}, \mathcal{Q})$ and a scalar $\beta \in \mathbb{C}$, the scaled operator βL is defined as:

$$(\beta L)x = \beta(Lx), \quad \forall x \in \mathcal{P}. \tag{2.2.3}$$

   This operation preserves the structure of scalar multiplication and ensures that the set of linear operators remains closed under scaling by elements of $\mathbb{C}$.

The vector space $L(\mathcal{P}, \mathcal{Q})$ has a dimension given by:

$$\dim(L(\mathcal{P}, \mathcal{Q})) = \dim(\mathcal{P}) \cdot \dim(\mathcal{Q}). \tag{2.2.4}$$

This result reflects the fact that a linear operator is completely determined by how it acts on a basis of $\mathcal{P}$, with each basis element mapping to a vector in $\mathcal{Q}$. A crucial aspect of any linear operator L is its kernel and image, which describe its null space and range, respectively.

1. Kernel: The kernel (or null space) of L, denoted $\ker(L)$, consists of all vectors in $\mathcal{P}$ that are mapped to the zero vector in $\mathcal{Q}$:
$$\ker(L) = \{x \in \mathcal{P} \mid Lx = 0\}. \tag{2.2.5}$$
This subspace indicates the set of elements that are annihilated by the transformation, which is important in determining whether L is injective (one-to-one).
2. Image: The image (or range) of L, denoted $\mathrm{im}(L)$, consists of all vectors in $\mathcal{Q}$ that can be written as Lx for some $x \in \mathcal{P}$:
$$\mathrm{im}(L) = \{Lx \mid x \in \mathcal{P}\}. \tag{2.2.6}$$
This subspace represents the portion of $\mathcal{Q}$ that L maps onto, determining whether the operator is surjective (onto).

The rank of the operator L, denoted $\mathrm{rank}(L)$, is defined as the dimension of its image:
$$\mathrm{rank}(L) = \dim(\mathrm{im}(L)). \tag{2.2.7}$$

A fundamental result in linear algebra, known as the rank-nullity theorem, states that the sum of the dimensions of the kernel and the image of L equals the dimension of $\mathcal{P}$:
$$\dim(\ker(L)) + \mathrm{rank}(L) = \dim(\mathcal{P}). \tag{2.2.8}$$

This theorem provides crucial insight into how linear transformations distribute dimensions between the space of solutions and the range of transformation, reinforcing the fundamental structure of vector spaces and mappings between them.

2.2.1 Matrices and their correspondences with Linear operator:

A matrix over the field of complex numbers is formally defined as a mapping of the form:
$$K : \Xi \times \Omega \longrightarrow \mathbb{C}, \tag{2.2.9}$$

where Ξ and Ω are finite, nonempty index sets. The collection of all such matrices is denoted by $\mathcal{K}_{\Xi,\Omega}(\mathbb{C})$. For specific elements $x \in \Xi$ and $y \in \Omega$, the value $K(x, y)$ is referred to as the (x, y)-entry of the matrix K, where x serves as the row index and y as the column index. The collection $\mathcal{K}_{\Xi,\Omega}(\mathbb{C})$ forms a vector space under the usual operations of matrix addition and scalar multiplication, defined as follows:

1. Addition: Given matrices $A, B \in \mathcal{K}_{\Xi,\Omega}(\mathbb{C})$, their sum $A + B \in \mathcal{K}_{\Xi,\Omega}(\mathbb{C})$ is defined entry-wise as
$$(A + B)(x, y) = A(x, y) + B(x, y), \quad \forall x \in \Xi, \quad \forall y \in \Omega. \tag{2.2.10}$$
2. Scalar multiplication: Given a scalar $\beta \in \mathbb{C}$ and a matrix $C \in \mathcal{K}_{\Xi,\Omega}(\mathbb{C})$, the matrix βC is defined by
$$(\beta C)(x, y) = \beta C(x, y), \quad \forall x \in \Xi, \quad \forall y \in \Omega. \tag{2.2.11}$$

As a vector space, $\mathcal{K}_{\Xi,\Omega}(\mathbb{C})$ is therefore isomorphic to the complex Euclidean space $\mathbb{C}^{\Xi\times\Omega}$. Matrix multiplication is defined in the conventional manner. Given matrices $A \in \mathcal{K}_{\Xi,\Lambda}(\mathbb{C})$ and $B \in \mathcal{K}_{\Lambda,\Omega}(\mathbb{C})$, where Λ is another finite, nonempty index set, the product $AB \in \mathcal{K}_{\Xi,\Omega}(\mathbb{C})$ is given by the summation rule:
$$(AB)(x, y) = \sum_{z\in\Lambda} A(x, z)B(z, y), \quad \forall x \in \Xi, \quad \forall y \in \Omega. \tag{2.2.12}$$

Linear operators between complex Euclidean spaces naturally correspond to matrices. Consider the spaces $\mathcal{P} = \mathbb{C}^{\Omega}$ and $\mathcal{Q} = \mathbb{C}^{\Xi}$. Each linear operator $L \in L(\mathcal{P}, \mathcal{Q})$ can be represented by a matrix $K_L \in \mathcal{K}_{\Xi,\Omega}(\mathbb{C})$, where the (x, y)-entry is given by:
$$K_L(x, y) = \langle e_x, Le_y\rangle, \tag{2.2.13}$$

where $\{e_y\}_{y\in\Omega}$ and $\{e_x\}_{x\in\Xi}$ denote the standard basis vectors of $\mathcal{P}$ and $\mathcal{Q}$, respectively. Conversely, any matrix $K \in \mathcal{K}_{\Xi,\Omega}(\mathbb{C})$ defines a linear operator $L_K \in L(\mathcal{P}, \mathcal{Q})$ via its action on a vector $u \in \mathcal{P}$:

$$(L_K u)(x) = \sum_{y\in\Omega} K(x,y)u(y), \quad \forall x \in \Xi. \tag{2.2.14}$$

These mappings, $L \mapsto K_L$ and $K \mapsto L_K$, are linear and mutually inverse. Moreover, the composition of linear operators corresponds to matrix multiplication:

$$K_{LM} = K_L K_M, \tag{2.2.15}$$

whenever $L \in L(\mathcal{Q}, \mathcal{R})$ and $M \in L(\mathcal{P}, \mathcal{Q})$ for complex Euclidean spaces $\mathcal{P}$, $\mathcal{Q}$, and $\mathcal{R}$. Equivalently, given matrices $A \in \mathcal{K}_{\Xi,\Lambda}(\mathbb{C})$ and $B \in \mathcal{K}_{\Lambda,\Omega}(\mathbb{C})$, we have:

$$L_{AB} = L_A L_B. \tag{2.2.16}$$

From this point onward, we will freely interchange between discussing matrices and operators, depending on which perspective is more convenient. When dealing with an operator $A \in L(\mathcal{X}, \mathcal{Y})$, we will often denote its corresponding matrix representation simply as A, and refer to its (x, y)-entry as $A(x, y)$ when necessary.

2.2.2 Conjugate, Transpose, and Adjoint of Linear Operators:

Consider two complex Euclidean spaces, denoted as $\mathcal{P} = \mathbb{C}^{\Xi}$ and $\mathcal{Q} = \mathbb{C}^{\Lambda}$. For any linear operator $L \in L(\mathcal{P}, \mathcal{Q})$, where $L(\mathcal{P}, \mathcal{Q})$ represents the space of bounded linear operators from $\mathcal{P}$ to $\mathcal{Q}$, we define three additional associated operators:

1. Conjugate Operator: The operator $K \in L(\mathcal{P}, \mathcal{Q})$ has a matrix representation whose entries are the complex conjugates of the matrix representation of A:

   $$\overline{K}(x,y) = \overline{K(x,y)} \tag{2.2.17}$$

   for all $x \in \Lambda$ and $y \in \Xi$. This operator is useful in preserving the structure of complex-valued transformations while considering conjugation.

2. Transpose Operator: The transpose $K^T \in L(\mathcal{Q}, \mathcal{P})$ is the operator whose matrix representation is obtained by interchanging the rows and columns of the matrix representation of K:

   $$K^T(y,x) = K(x,y) \tag{2.2.18}$$

   for all $x \in \Lambda$ and $y \in \Xi$. This operation effectively swaps the domain and codomain while preserving the structure of the transformation.

3. Adjoint Operator: The adjoint $K^* \in L(\mathcal{Q}, \mathcal{P})$ is uniquely defined by the inner product relation:

   $$\langle y, Kx\rangle = \langle K^* y, x\rangle \tag{2.2.19}$$

   for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. The adjoint operator can be computed by applying both conjugation and transposition to the matrix representation of A, yielding the relation:

   $$A^* = \overline{(A^T)} \tag{2.2.20}$$

   which effectively generalizes the Hermitian adjoint concept from finite-dimensional matrices to abstract operators in Hilbert spaces.

The mappings defined above obey the following linearity properties:

$$\begin{aligned}\overline{\eta K+\zeta M} &= \overline{\eta}\overline{K}+\overline{\zeta}\overline{M},\\ (\eta K+\zeta M)^* &= \overline{\eta}K^*+\overline{\zeta}M^*,\\ (\eta K+\zeta M)^T &= \eta K^T+\zeta M^T,\end{aligned} \tag{2.2.21}$$

for all $K, M \in L(\mathcal{P}, \mathcal{Q})$ and $\eta, \zeta \in \mathbb{C}$. These mappings are bijective, and each serves as its own inverse. Each vector $x \in \mathcal{P}$ in the complex Euclidean space $\mathcal{P}$ can be naturally identified with the linear operator in $L(\mathbb{C}, \mathcal{P})$ that maps $\beta \mapsto \beta x$. Under this identification, the operators $\overline{x} \in L(\mathbb{C}, \mathcal{P})$ and their transposes and adjoints, $x^T, x^* \in L(\mathcal{P}, \mathbb{C})$, are defined analogously. The vector $\overline{x} \in \mathcal{P}$, the vector representation satisfying:

$$\overline{x}(y) = \overline{x(y)} \tag{2.2.22}$$

for every $y \in \Delta$ if $\mathcal{P}$ is defined over alphabet set Δ. Also, for each vector $x \in \mathcal{P}$, the mapping $x^* \in L(\mathcal{P}, \mathbb{C})$ satisfies:

$$x^* y = \langle x, y\rangle, \quad \forall y \in \mathcal{P}. \tag{2.2.23}$$

This space of linear operators $L(\mathcal{P}, \mathbb{C})$ is commonly referred to as the *dual space* of $\mathcal{P}$, and is often denoted as $\mathcal{P}^*$ rather than $L(\mathcal{X}, \mathbb{C})$.

Standard Basis in the Space of Operators: Assuming $\mathcal{Q} = \mathbb{C}^\Gamma$ and $\mathcal{R} = \mathbb{C}^\Omega$, for each pair $(x \in \Omega, y \in \Gamma)$, we define the elementary operator $E_{x,y} \in L(\mathcal{Q}, \mathcal{R})$ by:

$$E_{x,y} = e_x e_y^*, \tag{2.2.24}$$

which in component form satisfies:

$$E_{x,y}(z, w) = \begin{cases} 1, & \text{if } (x = z) \text{ and } (y = w),\\ 0, & \text{if } (x \neq z) \text{ or } (y \neq w).\end{cases} \tag{2.2.25}$$

The set $\{E_{x,y} : x \in \Omega, y \in \Gamma\}$ forms a basis for $L(\mathcal{Q}, \mathcal{R})$ and is referred to as the *standard basis* of this operator space.

2.2.3 Direct Sums:

The concept of a direct sum extends the structure of complex Euclidean spaces by combining multiple individual spaces into a larger composite space. Given n complex Euclidean spaces $\mathcal{M}_1 = \mathbb{C}^{\Gamma_1}, \ldots, \mathcal{M}_n = \mathbb{C}^{\Gamma_n}$, their direct sum is denoted as

$$\bigoplus_{i=1}^{n} \mathcal{M}_i = \mathbb{C}^{\sqcup_{i=1}^{i=n}\Gamma_i}, \tag{2.2.26}$$

where the symbol $\sqcup$ represents the disjoint union of the index sets $\Gamma_1, \ldots, \Gamma_n$. Explicitly, this union is described as

$$\begin{aligned}\bigsqcup_{i=1}^{n} \Gamma_i &= \{(1, b_1) : b_1 \in \Gamma_1\} \cup \cdots \cup \{(n, b_n) : b_n \in \Gamma_n\} \quad \forall b_i \in \Gamma_i\\ &= \bigcup_{j=1}^{n} \{(j, b_j) : b_j \in \Gamma_j\} \quad \forall b_j \in \Gamma_i\end{aligned} \tag{2.2.27}$$

where each element in the union is uniquely labeled by a pair consisting of the space index k and the corresponding element b in Γ_k. For vectors $x_1 \in \mathcal{M}_1, \ldots, x_n \in \mathcal{M}_n$, the notation $\bigoplus_{i=1}^{n} x_i$ represents the vector in the direct sum space defined by the relation

$$\bigoplus_{i=1}^{n} x_i(k, b) = x_k(b), \tag{2.2.28}$$

for each $j \in \{1, 2, 3, \cdots n\}$ and $b \in \Gamma_j$.The dimension of a vector $m \in \bigoplus_{i=1}^{n} \mathcal{M}_\rangle$ is given by $d = \sum_{i=1}^{n} |\mathcal{M}_i|$ which systematically assigns the values of each component vector to the appropriate indexed position. The direct sum space maintains fundamental vector space properties, and every element in $\bigoplus_{i=1}^{n} \mathcal{M}_i$ can be uniquely decomposed as $\oplus_{i=1}^{n} x_i$ for some choice of constituent vectors. The following algebraic identities hold for all $x_1, y_1 \in \mathcal{M}_1, \ldots, x_n, y_n \in \mathcal{M}_n$ and any scalar $\beta \in \mathbb{C}$:

$$\oplus_{i=1}^{n} x_i + \oplus_{i=1}^{n} y_i = \oplus_{i=1}^{n} (x_i + y_i), \tag{2.2.29}$$

$$\beta(\oplus_{i=1}^{n} x_i) = \oplus_{i=1}^{n} (\beta x_i), \tag{2.2.30}$$

$$\langle \oplus_{i=1}^{n} x_i, \oplus_{i=1}^{n} y_i \rangle = \sum_{i=1}^{n} \langle x_i, y_i \rangle. \tag{2.2.31}$$

Consider a collection of complex Euclidean spaces defined as $\mathcal{M}_i = \mathbb{C}^{\Omega_i}, \quad \mathcal{N}_j = \mathbb{C}^{\Lambda_j}$ where $i \in \{1, 2, 3, ...n\}$,$i \in \{1, 2, 3, ...m\}$ and $n, m \in \mathbb{N}$ The Ω_i and , Λ_i are finite, nonempty sets that index the dimensions of these spaces. The direct sum of these spaces is given by:

$$\mathcal{M} = \bigoplus_{i=1}^{n} \mathcal{M}_i, \quad \mathcal{N} = \bigoplus_{i=1}^{m} \mathcal{N}_i. \tag{2.2.32}$$

Now, let $K \in L(\mathcal{M}, \mathcal{N})$ be a linear operator mapping $\mathcal{M}$ to $\mathcal{N}$. The matrix representation of K can be naturally identified with a block matrix:

$$K = \begin{bmatrix} K_{1,1} & \cdots & K_{1,n} \\ \vdots & \ddots & \vdots \\ K_{m,1} & \cdots & K_{m,n} \end{bmatrix}, \tag{2.2.33}$$

where each block $K_{i,j}$ represents an operator in $L(\mathcal{M}_j, \mathcal{N}_i)$ for every $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$. These operators are uniquely determined and describe how elements in $\mathcal{M}$ are transformed into elements in $\mathcal{N}$ under K. Given a vector $\bigoplus_{i=1}^{n} x_i \in \mathcal{M}$, its image under K is given by:

$$K(\bigoplus_{i=1}^{n} x_i) = \bigoplus_{j=1}^{m} y_j, \tag{2.2.34}$$

where each $y_j \in \mathcal{N}_j$ is explicitly determined as:

$$y_i = \sum_{j=1}^{n} K_{i,j} x_j \tag{2.2.35}$$

This formulation highlights the structure of K as an operator that acts on the direct sum space $\mathcal{M}$ by applying a collection of linear transformations, each affecting a specific subspace component. The block matrix representation provides a convenient way to analyze and compute the action of K, especially in cases where individual subspaces $\mathcal{M}_k$ and $\mathcal{N}_j$ have different dimensions. The structure of K enables decomposition into smaller operators, simplifying computations in applications such as functional analysis, quantum information theory, and numerical linear algebra.

2.2.4 Tensor Product:

The concept of the tensor product plays a fundamental role in linear algebra, particularly in the study of complex Euclidean spaces. Given n complex Euclidean spaces, denoted as $\mathcal{R}_1 = \mathbb{C}^{\Omega_1}, \mathcal{R}_2 = \mathbb{C}^{\Omega_2}, \ldots, \mathcal{R}_n = \mathbb{C}^{\Omega_n}$, their tensor product forms another complex Euclidean space expressed as:

$$\bigotimes_{i=1}^{n} \mathcal{R}_i = \mathbb{C}^{\Omega_1 \times \cdots \times \Omega_n} \tag{2.2.36}$$

This space encapsulates the combined structure of the original spaces, allowing for a more generalized representation of multi-dimensional systems. For vectors $x_1 \in \mathcal{R}_1, x_2 \in \mathcal{R}_2, \ldots, x_n \in \mathcal{R}_n$, the notation $\bigotimes_{i=1}^n x_i \in \bigotimes_{i=1}^n \mathcal{R}_\rangle$ defines an element in this new space. Specifically, the function defining this tensor product is given by:

$$\left(\bigotimes_{i=1}^n x_i\right)(a_1, a_2, \ldots, a_n) = x_1(a_1)x_2(a_2)\ldots x_n(a_n). \tag{2.2.37}$$

Here, the output of the tensor product function is determined by the component-wise multiplication of the respective vector elements at corresponding indices. This formulation is crucial in various applications, including quantum mechanics and multilinear algebra. Vectors that take this specific form, i.e., $\bigotimes_{i=1}^n x_i$, are referred to as elementary tensors. These vectors form a spanning set for the tensor product space $\bigotimes_{i=1}^n \mathcal{R}_\rangle$. However, it is important to note that not every element of $\bigotimes_{i=1}^n \mathcal{R}_\rangle$ can be written as a single elementary tensor; rather, general elements may be linear combinations of such elementary tensors. A number of essential algebraic properties govern tensor products. For all vectors $x_1, y_1 \in \mathcal{R}_1, x_2, y_2 \in \mathcal{R}_2, \ldots, x_n, y_n \in \mathcal{R}_n$, along with scalars $\beta, \gamma \in \mathbb{C}$, and an index $k \in \{1, \ldots, n\}$, the following fundamental identities hold, First, the tensor product operation distributes over vector addition, meaning that for a fixed index k, a linear combination of two vectors within the tensor product behaves as,

$$\bigotimes_{i=1}^{k-1} x_i \otimes (x_k + y_k) \otimes \bigotimes_{i=k+1}^{n} x_i = \bigotimes_{i=1}^{k-1} x_i \otimes (x_k) \otimes \bigotimes_{i=k+1}^{n} x_i + \bigotimes_{i=1}^{k-1} x_i \otimes (y_k) \otimes \bigotimes_{i=k+1}^{n} x_i \tag{2.2.38}$$

This property ensures that the tensor product remains consistent with linearity, which is crucial when extending concepts from single vector spaces to multi-linear settings. Second, the inner product of two tensor product vectors follows a multiplicative property:

$$\langle \bigotimes_{i=1}^{n} x_i, \bigotimes_{i=1}^{n} y_i \rangle = \langle x_1, y_1 \rangle \langle x_2, y_2 \rangle \ldots \langle x_n, y_n \rangle. \tag{2.2.39}$$

This result follows from the fundamental definition of inner products and is widely used in the study of Hilbert spaces, quantum information theory, and multilinear algebra. The third property is crucial to understanding the multiplication of the scaler in the context of the tensor product.

$$\alpha\left(\bigotimes_{i=1}^{n} x_i\right) = (\alpha x_1) \otimes \left(\bigotimes_{i=2}^{n} x_i\right) = x_1 \otimes (\alpha x_2)\left(\bigotimes_{i=3}^{n} x_i\right) = \cdots = \left(\bigotimes_{i=1}^{n-1} x_i\right) \otimes (\alpha x_n) \tag{2.2.40}$$

This result follows from the property of bilinearity of tensor product. The above properties illustrate the structural consistency of tensor products, preserving the foundational operations of vector spaces such as addition, scalar multiplication, and inner products. These characteristics make tensor products an indispensable tool in theoretical and applied mathematics.

2.3 Algebraic Structure of Operator Spaces

For every complex Euclidean space $\mathcal{P}$, the notation $L(\mathcal{P})$ is used as a shorthand for $L(\mathcal{P}, \mathcal{P})$, representing the space of linear operators that map $\mathcal{P}$ to itself. The space $L(\mathcal{P})$ exhibits several significant algebraic properties, making it a fundamental object of study. In particular, it forms an *associative algebra*, meaning it is a vector space equipped with an associative bilinear composition of operators:

$$\begin{aligned}(KM)N &= K(MN),\\ N(\beta K + \alpha M) &= \beta NK + \alpha NM,\\ (\beta K + \alpha M)N &= \beta KN + \alpha MN,\end{aligned} \tag{2.3.1}$$

for all choices of $K, M, N \in L(\mathcal{P})$ and scalars $\beta, \alpha \in \mathbb{C}$. The identity operator $\mathbb{1} \in L(\mathcal{P})$ is defined as the unique operator that satisfies $\mathbb{1}x = x$ for all $x \in \mathcal{P}$. When necessary, it may be denoted explicitly as $\mathbb{1}_{\mathcal{P}}$ to

emphasize its action on $\mathcal{P}$. An operator $K \in L(\mathcal{P})$ is said to be *invertible* if there exists another operator $M \in L(\mathcal{P})$ such that $MK = KM = \mathbb{1}$. If such an operator M exists, it is necessarily unique. This unique inverse is denoted by K^{-1}. The collection of all *invertible operators* within $L(\mathcal{P})$ is denoted by $GL(\mathcal{P})$ and is referred to as the *general linear group* of $\mathcal{P}$.
For any two operators $K, M \in L(\mathcal{P})$, the bracket, also known as the commutator and the braces also known as anti-commutator, is defined as:

$$\begin{aligned} [K, M] &= KM - MK. \\ \{K, M\} &= KM + MK. \end{aligned} \tag{2.3.2}$$

This bracket operation measures the extent to which the two operators fail to commute and plays a fundamental role in various fields such as functional analysis and quantum mechanics.

2.3.1 Trace and Determinant

Operators in the algebra $L(\mathcal{P})$ are represented by *square matrices*, where the rows and columns are indexed by elements of the same finite set. Two fundamental scalar functions that arise in this context are the *trace* and the *determinant*, both of which provide key insights into the properties of linear transformations.

1. The trace of an operator $K \in L(\mathcal{P})$, where $\mathcal{P} = \mathbb{C}^{\Omega}$, is defined as:

   $$\mathrm{Tr}(K) = \sum_{a \in \Omega} K(a, a). \tag{2.3.3}$$

   The trace is a linear function and satisfies the fundamental property:

   $$\mathrm{Tr}(KM) = \mathrm{Tr}(MK), \tag{2.3.4}$$

   for any operators $K \in L(\mathcal{P}, \mathcal{N})$ and $M \in L(\mathcal{N}, \mathcal{P})$, where $\mathcal{P}, \mathcal{N}$ are arbitrary complex Euclidean spaces.

2. The determinant of an operator $K \in L(\mathcal{P})$ for $\mathcal{P} = \mathbb{C}^{\Omega}$ is given by:

   $$\det(K) = \sum_{\pi \in \mathrm{Sym}(\Omega)} \mathrm{sign}(\pi) \prod_{a \in \Omega} K(a, \pi(a)), \tag{2.3.5}$$

   where $Sym(\Omega)$ is the group of all permutations of Ω, and $sign(\pi)$ represents the sign of the permutation π (taking values $+1$ for even permutations and -1 for odd permutations).

The determinant function satisfies the multiplicative property:

$$\det(KM) = \det(K)\det(M), \tag{2.3.6}$$

for all $K, M \in L(\mathcal{P})$, and it is nonzero if and only if K is invertible.

2.3.2 Hilbert–Schmidt Inner Product

The inner product can be defined using the trace function, on the space of operators $L(\mathcal{P}, \mathcal{N})$ as follows:

$$\langle K, M \rangle = \mathrm{Tr}(K^* M), \tag{2.3.7}$$

for all $K, M \in L(\mathcal{P}, \mathcal{N})$, where K^* denotes the adjoint of K. This inner product, known as the Hilbert–Schmidt inner product, satisfies the following properties:

1. Linearity in the second argument:

   $$\langle K, \beta M + \alpha N \rangle = \beta \langle K, M \rangle + \alpha \langle K, N \rangle, \tag{2.3.8}$$

   for all $K, M, N \in L(\mathcal{P}, \mathcal{N})$ and $\beta, \alpha \in \mathbb{C}$.

2. Conjugate symmetry:
$$\langle K, M\rangle = \overline{\langle M, K\rangle}, \tag{2.3.9}$$
ensuring that the inner product respects complex conjugation.

3. Positive semi-definiteness:
$$\langle K, K\rangle \geq 0, \forall A \in L(\mathcal{P}, \mathcal{N}) \quad \text{with equality if and only if} \quad K = 0. \tag{2.3.10}$$

These properties ensure that the Hilbert–Schmidt inner product provides a well-defined inner product structure on the space of linear operators.

2.3.3 Eigenvalues and Eigenvectors

Given an operator $M \in L(\mathcal{P})$, a vector $u \in \mathcal{P}$ and $u \neq 0$ is an eigenvector of M if there exists a scalar $\lambda \in \mathbb{C}$ such that
$$Mx = \lambda x. \tag{2.3.11}$$
The scalar λ is called an eigenvalue of M.

2.4 Important Classes of Operators

The following section is going to introduce some classes of operators frequently used in the theory of quantum information

1. **Normal Operators:** An operator $N \in L(\mathcal{P})$ is normal if it commutes with its adjoint N^*:
$$[N, N^*] = 0 \quad \iff \quad NN^* = N^*N \tag{2.4.1}$$
Normal operators share a common set of eigenvectors with their adjoints and can be simultaneously diagonalized by a unitary transformation. The spectral theorem applies to normal operators. Unitary, Hermitian, and positive semidefinite operators are special cases of normal operators.

Example 2.4.1. Let us consider
$$N = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \text{ then, } N^* = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{2.4.2}$$
$$\begin{aligned} NN^* &= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} (0)(0) + (1)(1) & (0)(-1) + (1)(0) \\ (-1)(0) + (0)(1) & (-1)(-1) + (0)(0) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \\ \text{and} & \\ N^*N &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} (0)(0) + (-1)(-1) & (0)(1) + (-1)(0) \\ (1)(0) + (0)(-1) & (1)(1) + (0)(0) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \end{aligned} \tag{2.4.3}$$
Since $NN^* = N^*N = I$, N is a normal operator.

2. **Hermitian Operators:** A Hermitian (or self-adjoint) operator $H \in L(\mathcal{P})$ is equal to its own adjoint:

$$\text{Herm}(\mathcal{P}) = \{H \in L(\mathcal{P}) : H = H^*\} \tag{2.4.4}$$

where $\text{Herm}(\mathcal{P})$ denotes the set of all Hermitian operators on $\mathcal{P}$. Hermitian operators have real eigenvalues, are diagonalizable with real eigenvalues, and possess an orthonormal basis of eigenvectors. In quantum mechanics, observables are represented by Hermitian operators.

Example 2.4.2. $H = \begin{bmatrix} 2 & i \\ -i & 3 \end{bmatrix}$ is Hermitian because it is equal to its own conjugate transpose H^*. To show this, we first find the conjugate transpose H^*:

$$H^* = \left(\begin{bmatrix} 2 & i \\ -i & 3 \end{bmatrix} \right)^* = \begin{bmatrix} \overline{2} & \overline{-i} \\ \overline{i} & \overline{3} \end{bmatrix} = \begin{bmatrix} 2 & i \\ -i & 3 \end{bmatrix} = H \tag{2.4.5}$$

Since $H^* = H$, the matrix H is indeed Hermitian.

3. **Positive Semidefinite Operators:** A positive semidefinite operator $P \in L(\mathcal{P})$ can be written as the product of an operator and its adjoint:

$$\text{Pos}(\mathcal{P}) = \{P \in L(\mathcal{P}) : P = R^* R\} \tag{2.4.6}$$

where Pos denotes the set of all positive semidefinite operators over $\mathcal{P}$. Positive semidefinite operators include the following properties,

- P is positive semidefinite.
- $P = B^* B$ for some choice of a complex Euclidean space $\mathcal{Y}$ and an operator $B \in L(\mathcal{X}, \mathcal{Y})$.
- $u^* P u$ is a nonnegative real number for every choice of $u \in \mathcal{X}$.
- $\langle Q, P \rangle$ is a nonnegative real number for every $Q \in \text{Pos}(\mathcal{X})$.
- P is Hermitian and every eigenvalue of P is nonnegative.
- There exists a complex Euclidean space $\mathcal{Y}$ and a collection of vectors $\{u_a : a \in \Sigma\} \subset \mathcal{Y}$, such that $P(a, b) = \langle u_a, u_b \rangle$.

Example 2.4.3. Let $R = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$. First, we find the conjugate transpose of R, denoted as R^*. Since all elements in R are real, R^* is simply the transpose of R, R^T.

$$R^* = R^T = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \tag{2.4.7}$$

Now, we compute the product $P = R^* R$:

$$\begin{aligned} P &= R^* R \\ &= \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} (1)(1) + (0)(0) & (1)(2) + (0)(1) \\ (2)(1) + (1)(0) & (2)(2) + (1)(1) \end{bmatrix} \\ &= \begin{bmatrix} 1+0 & 2+0 \\ 2+0 & 4+1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \end{aligned} \tag{2.4.8}$$

This matrix is Hermitian, and its eigenvalues are non-negative, thus it is positive semidefinite.

4. **Positive Definite Operators:** A positive definite operator is a positive semidefinite operator that is also invertible ($\det(P) \neq 0$).

$$\mathrm{Pd}(\mathcal{P}) = \{P \in \mathrm{Pos}(\mathcal{P}) : \det(P) \neq 0\} \tag{2.4.9}$$

where $\mathrm{Pd}(\mathcal{P})$ is the set of all Positive definite operators over $\mathcal{P}$ that satisfy:

- P is positive definite.
- $\langle u, Pu \rangle$ is a positive real number for every nonzero vector $u \in \mathcal{X}$.
- P is Hermitian, and every eigenvalue of P is positive.
- P is Hermitian, and there exists a positive real number $\varepsilon > 0$ such that $P \geq \varepsilon \mathbb{1}$.

Example 2.4.4. $P = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ is positive definite because its eigenvalues are 2 and 3 (both > 0) and $\det(P) = 2 \cdot 3 - 0 \cdot 0 = 6 \neq 0$.

5. **Density Operators:** A density operator ρ on a complex Euclidean space $\mathcal{X}$ describes the state of a quantum system. It satisfies two key properties:

- *Positive semidefinite:* All eigenvalues are non-negative ($\rho \geq 0$).
- *Trace equals 1:* The sum of its diagonal elements is one ($\mathrm{Tr}(\rho) = 1$).

Formally, a matrix $\rho \in \mathrm{Pos}(\mathcal{X})$ is a density operator if $\mathrm{Tr}(\rho) = 1$. The set of all density operators on $\mathcal{X}$ is denoted by $D(\mathcal{X})$.

$$D(\mathcal{P}) = \{\rho \in \mathrm{Pos}(\mathcal{P}) : \mathrm{Tr}(\rho) = 1\} \tag{2.4.10}$$

Example 2.4.5. The pure state $|\psi\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ in $\mathbb{C}^2$ has a density operator

$$\rho = |\psi\rangle\langle\psi| = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \tag{2.4.11}$$

which satisfies $\mathrm{Tr}(\rho) = 1 + 0 = 1$ and $\rho \geq 0$ (eigenvalues are 1 and 0).

6. **Projection Operators:** A projection operator $\Pi \in \mathrm{Pos}(\mathcal{X})$ is a positive semidefinite operator that is idempotent:

$$\Pi^2 = \Pi \tag{2.4.12}$$

Projection operators are also Hermitian ($\Pi = \Pi^*$) and have eigenvalues of only 0 or 1. The set of all projection operators on $\mathcal{X}$ is denoted by $\mathrm{Proj}(\mathcal{X})$. For any subspace $V \subseteq \mathcal{X}$, there exists a unique projection operator $\Pi_V \in \mathrm{Proj}(\mathcal{X})$ whose image is exactly V. This operator projects any vector in $\mathcal{X}$ onto the subspace V.

Example 2.4.6. For the subspace V spanned by $|0\rangle$ in $\mathbb{C}^2$, the projection operator is

$$\Pi = |0\rangle\langle 0| = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \tag{2.4.13}$$

This operator maps any vector $|\psi\rangle = a|0\rangle + b|1\rangle$ to $a|0\rangle$:

$$\Pi|\psi\rangle = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ 0 \end{bmatrix} = a|0\rangle. \tag{2.4.14}$$

Equivalent notations for projectors include P.

7. **Isometries:** An isometry is a linear map $A : \mathcal{X} \to \mathcal{Y}$ that preserves the length (norm) of vectors:

$$\|Au\| = \|u\| \quad \text{for all } u \in \mathcal{X} \tag{2.4.15}$$

This is equivalent to $A^*A = I_{\mathcal{X}}$, where A^* is the adjoint of A, and $I_{\mathcal{X}}$ is the identity operator on $\mathcal{X}$. The set of all isometries from $\mathcal{X}$ to $\mathcal{Y}$ is denoted by

$$\mathcal{U}(\mathcal{X}, \mathcal{Y}) = \{A \in L(\mathcal{X}, \mathcal{Y}) : A^*A = I_{\mathcal{X}}\} \tag{2.4.16}$$

A necessary condition for an isometry to exist from $\mathcal{X}$ to $\mathcal{Y}$ is $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$. Isometries also preserve inner products:

$$\langle Au, Av \rangle = \langle u, v \rangle \quad \text{for all } u, v \in \mathcal{X} \tag{2.4.17}$$

Example 2.4.7. The map $A : \mathbb{C}^2 \to \mathbb{C}^3$ given by the matrix

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{2.4.18}$$

is an isometry, as

$$A^*A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 + 0 \cdot 0 + 0 \cdot 0 & 1 \cdot 0 + 0 \cdot 1 + 0 \cdot 0 \\ 0 \cdot 1 + 1 \cdot 0 + 0 \cdot 0 & 0 \cdot 0 + 1 \cdot 1 + 0 \cdot 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2. \tag{2.4.19}$$

8. **Unitary Operators:** A unitary operator is an isometry where the domain and codomain are the same, i.e., $U \in \mathcal{U}(\mathcal{X}, \mathcal{X})$. It satisfies:

$$U^*U = UU^* = I_{\mathcal{X}} \tag{2.4.20}$$

This implies that a unitary operator is invertible, with its inverse being its adjoint: $U^{-1} = U^*$. Unitary operators are also normal ($UU^* = U^*U$). The set of all unitary operators on $\mathcal{X}$ is denoted by $\mathcal{U}(\mathcal{X})$. Unitary operators preserve both norms and inner products and are crucial for representing reversible evolutions in quantum computing and quantum information.

Example 2.4.8. The Hadamard operator is given by $H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ To show that H is a unitary operator, we need to verify if $H^*H = I$, where I is the identity matrix and H^* is the conjugate transpose of H.

Since all elements of H are real, its conjugate transpose H^* is simply its transpose H^T:

$$H^* = H^T = \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right)^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{2.4.21}$$

Now, let's compute the product H^*H:

$$\begin{aligned} H^*H &= \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right) \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right) \\ &= \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} (1)(1) + (1)(1) & (1)(1) + (1)(-1) \\ (1)(1) + (-1)(1) & (1)(1) + (-1)(-1) \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 1+1 & 1-1 \\ 1-1 & 1+1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= I_2 \end{aligned} \tag{2.4.22}$$

Since $H^*H = I_2$, the Hadamard operator H is indeed a unitary operator.

2.5 The Spectral Theorem

The *spectral theorem* is a cornerstone in linear algebra and functional analysis. It tells us that *every normal operator*—that is, an operator that commutes with its adjoint—can be "broken down" or represented as a *weighted sum of projections* onto mutually orthogonal subspaces. These projections effectively isolate the action of the operator on different components of the space. The term *spectral* refers to the fact that the coefficients in this decomposition are the *eigenvalues* of the operator, which collectively form its *spectrum*.

Theorem 2.5.1. Let $\mathcal{X}$ be a complex inner product space, and let $A \in L(\mathcal{X})$ be a normal operator. Suppose the distinct eigenvalues of A are $\lambda_1, \ldots, \lambda_k$. Then there exist unique orthogonal projections $P_1, \ldots, P_k \in \mathrm{Pos}(\mathcal{X})$ such that:

- The projections are *mutually orthogonal*, i.e., $P_i P_j = 0$ for $i \neq j$,
- They form a *resolution of the identity*, meaning $P_1 + \cdots + P_k = \mathbb{1}_X$, and
- The operator A can be expressed as:

$$A = \sum_{i=1}^{k} \lambda_i P_i. \tag{2.5.1}$$

This expression is known as a spectral decomposition of A. Each projection P_i projects onto the eigenspace associated with the eigenvalue λ_i, and the rank of P_i equals the algebraic multiplicity of λ_i.

Example 2.5.2. Consider a diagonal matrix $A = \mathrm{diag}(2, 2, 5)$. Then the spectral decomposition is $A = 2P_1 + 5P_2$, where P_1 is the projection onto the 2-eigenspace (a 2D subspace), and P_2 projects onto the 5-eigenspace.

Theorem 2.5.3. Spectral Theorem (Alternate statement – Vector Form) A slightly different, but equivalent, formulation is given below: Let $A \in L(\mathcal{X})$ be normal, and suppose its spectrum consists of eigenvalues $\lambda_1, \ldots, \lambda_n$. Then, there exists an orthonormal basis $\{x_1, \ldots, x_n\}$ of $\mathcal{X}$ such that:

$$A = \sum_{i=1}^{n} \lambda_i x_i x_i^*. \tag{2.5.2}$$

Here, each x_i is an eigenvector of A with eigenvalue λ_i, and the term $x_i x_i^*$ acts as a projection onto the one-dimensional subspace spanned by x_i. This representation is commonly referred to as a spectral decomposition in vector form.
Uniqueness: The decomposition using projections P_i (first version) is unique. However, the vector form (second version) is generally not unique if eigenvalues have multiplicities greater than one, since the choice of orthonormal eigenvectors within an eigenspace is not unique.

2.5.1 Commuting Normal Operators

An important extension of the spectral theorem involves *simultaneous diagonalization*. If two normal operators A and B commute, i.e., $[A, B] = 0$, then there exists a single orthonormal basis $\{x_1, \ldots, x_n\}$ such that both operators are diagonal in that basis:

$$A = \sum_{i=1}^{n} \lambda_i x_i x_i^*, \quad B = \sum_{i=1}^{n} \mu_i x_i x_i^*. \tag{2.5.3}$$

This result is particularly useful in quantum mechanics and multivariable operator theory, where observables represented by commuting normal operators can be simultaneously measured or analyzed.

2.5.2 Function of a Normal Operator

The spectral theorem provides a powerful tool that allows us to define functions of normal operators in a way that generalizes how we apply functions to scalars. In particular, given a scalar function $f : \mathbb{C} \longrightarrow \mathbb{C}$ and a normal operator A on a complex Euclidean space with spectral decomposition:

$$A = \sum_{i=1}^{k} \lambda_i P_i, \tag{2.5.4}$$

we can define a new operator $f(A)$ by applying the function to the eigenvalues:

$$f(A) = \sum_{i=1}^{k} f(\lambda_i) P_i. \tag{2.5.5}$$

This definition is valid because the spectral decomposition isolates each eigenvalue λ_i via the orthogonal projection P_i onto the corresponding eigenspace. Therefore, $f(A)$ acts on each eigenspace by scaling it with $f(\lambda_i)$, extending the function from scalars to operators in a consistent and meaningful way.

Exponential Function

One important function to extend to operators is the exponential function. For complex scalars $\alpha \in \mathbb{C}$, the exponential function $\exp(\alpha)$ is defined via its power series. For a normal operator A with spectral decomposition as above, we define:

$$\exp(A) = \sum_{i=1}^{k} \exp(\lambda_i) P_i. \tag{2.5.6}$$

This expression coincides with the classical matrix exponential defined by the Taylor series:

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}. \tag{2.5.7}$$

For general bounded operators, the Taylor series converges in norm. However, for normal operators, both the spectral definition and the series definition yield the same result. The operator exponential is used in solving systems of linear differential equations and in quantum mechanics to describe time evolution.

Powers of Operators

Another important class of functions is the power functions. Let Q be a positive semidefinite operator, meaning that all of its eigenvalues λ_i are real and non-negative ($\lambda_i \geq 0$). For any real number $r > 0$, we define:

$$Q^r = \sum_{i=1}^{k} \lambda_i^r P_i. \tag{2.5.8}$$

This definition allows us to meaningfully talk about non-integer powers of operators. A particularly common case is when $r = \frac{1}{2}$, giving the square root of an operator:

$$\sqrt{Q} = Q^{1/2}, \tag{2.5.9}$$

which satisfies the identity:

$$\sqrt{Q} \cdot \sqrt{Q} = Q. \tag{2.5.10}$$

This operator square root is especially useful in statistics (for instance, when computing the square root of a covariance matrix) and in quantum information theory when dealing with density operators or quantum fidelity.

Logarithm of an Operator

For a positive definite operator Q (i.e., an operator with strictly positive eigenvalues), the logarithm function can also be extended from scalars to operators. Assuming all eigenvalues $\lambda_i > 0$, we define:

$$\log(Q) = \sum_{i=1}^{k} \log(\lambda_i) P_i. \tag{2.5.11}$$

This operator logarithm is extremely important in quantum information theory. For instance, it appears in the formula for the *von Neumann entropy*, which measures the uncertainty (or mixedness) of a quantum state represented by a density operator ρ:

$$S(\rho) = -\mathrm{Tr}(\rho \log \rho). \tag{2.5.12}$$

Here, the logarithm is computed using the spectral decomposition as above. Overall, the spectral theorem provides a robust framework for extending scalar functions to operators, enabling rich mathematical and physical applications involving functions like exponentials, roots, and logarithms.
The spectral theorem provides a powerful toolset for understanding and manipulating normal operators by reducing them to their spectral components. Whether we're computing exponentials, roots, or logarithms of operators, the spectral decomposition provides a clean and insightful pathway. Moreover, the ability to simultaneously diagonalise commuting normal operators underpins many physical and computational applications, from quantum mechanics to numerical linear algebra.

2.6 The Singular Value Theorem and Moore-Penrose Pseudo-Inverse

In quantum information theory, linear operators are fundamental tools. While the *spectral theorem* applies to normal operators (i.e., operators that commute with their adjoints), it does not generalise to all operators. To analyse arbitrary operators, even those acting between different spaces, we use the *singular value theorem (SVT)*.

2.6.1 The Singular Value Theorem

Theorem 2.6.1. Let $A \in L(\mathcal{X}, \mathcal{Y})$ be a linear operator between complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$ with rank r. then,

$$A = \sum_{j=1}^{r} s_j y_j x_j^*. \tag{2.6.1}$$

where

- Positive real numbers $s_1, s_2, \ldots, s_r$ called (*singular values*),
- An orthonormal set of vectors $\{x_1, \ldots, x_r\} \subset X$ (*right singular vectors*),
- An orthonormal set of vectors $\{y_1, \ldots, y_r\} \subset Y$ (*left singular vectors*),

This decomposition is called the *Singular Value Decomposition (SVD)* of A.
Uniqueness and Ordering: The singular values $s_1, \ldots, s_r$ are uniquely determined up to order. Without loss of generality, we assume:

$$s_1 \geq s_2 \geq \cdots \geq s_r > 0. \tag{2.6.2}$$

For convenience, we define $s_k(A) = 0$ for $k > r$, allowing us to represent the singular value list as a full vector in $\mathbb{R}^n$.
Relation to the Spectral Theorem: The SVD relates closely to the spectral decomposition of the operators A^*A and AA^*:

$$s_k(A) = \sqrt{\lambda_k(A^*A)} = \sqrt{\lambda_k(AA^*)}, \quad 1 \leq k \leq r, \tag{2.6.3}$$

where λ_k denotes the k-th eigenvalue. The right singular vectors x_j are eigenvectors of A^*A, and the left singular vectors y_j are eigenvectors of AA^*.

Example 2.6.2. If A is a normal operator on $\mathcal{X}$, with spectral decomposition:

$$A = \sum_{j=1}^{n} \lambda_j x_j x_j^*, \tag{2.6.4}$$

where x_j are orthonormal eigenvectors, then the SVD is:

$$s_j = |\lambda_j|, \quad y_j = \frac{\lambda_j}{|\lambda_j|} x_j. \tag{2.6.5}$$

Hence, the singular values are the absolute values of the eigenvalues.

2.6.2 The Moore-Penrose Pseudo-Inverse

For non-invertible operators, we define a generalised inverse called the *Moore-Penrose pseudo-inverse*. For any $A \in L(\mathcal{X}, \mathcal{Y})$, its pseudo-inverse $A^+ \in L(\mathcal{Y}, \mathcal{X})$ is the unique operator satisfying:

1. $AA^+A = A$,
2. $A^+AA^+ = A^+$,
3. AA^+ and A^+A are Hermitian.

Construction via SVD: Given the SVD of A,

$$A = \sum_{j=1}^{r} s_j y_j x_j^*, \tag{2.6.6}$$

The pseudo-inverse is:

$$A^+ = \sum_{j=1}^{r} \frac{1}{s_j} x_j y_j^*. \tag{2.6.7}$$

Example 2.6.3 (Pseudo-Inverse of a Rectangular Matrix)**.** Let,

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \in \mathbb{C}^{2\times 3}, \tag{2.6.8}$$

then the pseudo-inverse is:

$$A^+ = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \in \mathbb{C}^{3\times 2}. \tag{2.6.9}$$

This reverses the action of A on its image.

Uniqueness: If two operators X, Y satisfy the pseudo-inverse conditions, one can show $X = Y$ by applying the identities and Hermitian properties, confirming the pseudo-inverse is unique.

2.7 Linear Mappings on Operator Algebras

Linear mappings of the form

$$\Phi : L(\mathcal{X}) \longrightarrow L(\mathcal{Y}), \tag{2.7.1}$$

where $\mathcal{X}$ and $\mathcal{Y}$ are complex Euclidean spaces, play a central role in quantum information theory. The set of all such mappings is denoted $L(\mathcal{X}, \mathcal{Y})$, or $L(\mathcal{X})$ when $\mathcal{X} = \mathcal{Y}$, and this set forms a vector space under the following operations,

1. Addition: For $\Phi, \Psi \in L(\mathcal{X}, \mathcal{Y})$, define:
$$(\Phi + \Psi)(A) = \Phi(A) + \Psi(A), \quad \forall A \in L(\mathcal{X}). \tag{2.7.2}$$
2. Scalar multiplication: For $\alpha \in \mathbb{C}$, define:
$$(\alpha\Phi)(A) = \alpha \cdot \Phi(A), \quad \forall A \in L(\mathcal{X}). \tag{2.7.3}$$

For any $\Phi \in L(\mathcal{X}, \mathcal{Y})$, the adjoint $\Phi^* \in L(\mathcal{Y}, \mathcal{X})$ is the unique map satisfying:
$$\langle \Phi^*(B), A \rangle = \langle B, \Phi(A) \rangle, \quad \forall A \in L(\mathcal{X}), B \in L(\mathcal{Y}). \tag{2.7.4}$$

Two standard examples of such mappings are:

Example 2.7.1.

- Transpose: $T : L(\mathcal{X}) \longrightarrow L(\mathcal{X})$, defined as $T(A) = A^T$.
- Trace: $\mathrm{Tr} : L(\mathcal{X}) \longrightarrow \mathbb{C}$, defined as $\mathrm{Tr}(A)$, identifying $L(\mathbb{C}) = \mathbb{C}$.

2.7.1 Tensor Products of Operators and Mappings

Tensor products of operators can be defined concretely via Kronecker products or abstractly using multilinear functions. Consider operators $A_1 \in L(\mathcal{X}_1, \mathcal{Y}_1), \ldots, A_n \in L(\mathcal{X}_n, \mathcal{Y}_n)$. We define:
$$A_1 \otimes \cdots \otimes A_n \in L(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n, \mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_n) \tag{2.7.5}$$
via the matrix rule:
$$(A_1 \otimes \cdots \otimes A_n)((a_1, \ldots, a_n), (b_1, \ldots, b_n)) = A_1(a_1, b_1) \cdots A_n(a_n, b_n). \tag{2.7.6}$$
This construction satisfies:
$$(A_1 \otimes \cdots \otimes A_n)(u_1 \otimes \cdots \otimes u_n) = (A_1 u_1) \otimes \cdots \otimes (A_n u_n), \tag{2.7.7}$$
for all $u_i \in \mathcal{X}_i$. Since $\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n$ is spanned by such elementary tensors, this identity uniquely determines the tensor product. The tensor product distributes over addition:
$$A_1 \otimes \cdots \otimes (A_k + B_k) \otimes \cdots \otimes A_n = A_1 \otimes \cdots \otimes A_k \otimes \cdots \otimes A_n + A_1 \otimes \cdots \otimes B_k \otimes \cdots \otimes A_n. \tag{2.7.8}$$
Composition also distributes:
$$(B_1 \otimes \cdots \otimes B_n)(A_1 \otimes \cdots \otimes A_n) = (B_1 A_1) \otimes \cdots \otimes (B_n A_n). \tag{2.7.9}$$
Spectral and singular value decompositions of tensor product operators follow from the individual decompositions:
$$\|A_1 \otimes \cdots \otimes A_n\|_p = \|A_1\|_p \cdots \|A_n\|_p. \tag{2.7.10}$$

Tensor Products of Linear Mappings

Let $\Phi_1 : L(\mathcal{X}_1) \longrightarrow L(\mathcal{Y}_1), \ldots, \Phi_n : L(\mathcal{X}_n) \longrightarrow L(\mathcal{Y}_n)$ be linear maps. Then their tensor product is the unique mapping:
$$\Phi_1 \otimes \cdots \otimes \Phi_n : L(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n) \longrightarrow L(\mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_n), \tag{2.7.11}$$
satisfying:
$$(\Phi_1 \otimes \cdots \otimes \Phi_n)(A_1 \otimes \cdots \otimes A_n) = \Phi_1(A_1) \otimes \cdots \otimes \Phi_n(A_n). \tag{2.7.12}$$

Example 2.7.2 (Partial Trace)**.** Let $\mathcal{X}$ be a complex Euclidean space. The trace $\mathrm{Tr} : L(\mathcal{X}) \longrightarrow \mathbb{C}$ can be viewed as $\mathrm{Tr} : L(\mathcal{X}) \longrightarrow L(\mathbb{C})$. For another space $\mathcal{Y}$, define:
$$\mathrm{Tr}_{\mathcal{X}} := \mathrm{Tr} \otimes \mathbb{1}_{L(\mathcal{Y})} : L(\mathcal{X} \otimes \mathcal{Y}) \longrightarrow L(\mathcal{Y}). \tag{2.7.13}$$
This satisfies:
$$(\mathrm{Tr} \otimes \mathbb{1})(A \otimes B) = \mathrm{Tr}(A)B, \quad \forall A \in L(\mathcal{X}), B \in L(\mathcal{Y}). \tag{2.7.14}$$
Alternatively, using any orthonormal basis $\{x_a : a \in \Sigma\}$ of $\mathcal{X}$, we can write:
$$\mathrm{Tr}_{\mathcal{X}}(A) = \sum_{a \in \Sigma} (x_a^* \otimes \mathbb{1}_{\mathcal{Y}}) A (x_a \otimes \mathbb{1}_{\mathcal{Y}}), \quad \forall A \in L(\mathcal{X} \otimes \mathcal{Y}). \tag{2.7.15}$$
A similar expression holds for the partial trace over $\mathcal{Y}$.

2.8 Schmidt Decomposition

Theorem 2.8.1. Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces. For any nonzero vector $u \in \mathcal{X} \otimes \mathcal{Y}$, the mapping ensures a bijection between bipartite vectors and operators. Therefore, there exists a unique operator $A \in L(\mathcal{Y}, \mathcal{X})$ such that

$$u = \operatorname{vec}(A) \tag{2.8.1}$$

By applying the singular value decomposition (SVD) to the operator A, we write:

$$A = \sum_{k=1}^{r} s_k x_k y_k^* \tag{2.8.2}$$

where:

- $r = \operatorname{rank}(A)$,
- $s_1, \ldots, s_r > 0$ are the singular values of A,
- $\{x_1, \ldots, x_r\} \subset \mathcal{X}$ is an orthonormal set,
- $\{y_1, \ldots, y_r\} \subset \mathcal{Y}$ is also an orthonormal set.

Applying the linearity of the `vec` mapping and using the identity $\operatorname{vec}(xy^*) = x \otimes y$, we obtain:

$$u = \operatorname{vec}(A) = \sum_{k=1}^{r} s_k x_k \otimes y_k \tag{2.8.3}$$

This yields a decomposition of u into a sum of tensor products of orthonormal vectors. Therefore, every nonzero vector $u \in \mathcal{X} \otimes \mathcal{Y}$ can be expressed as:

$$u = \sum_{k=1}^{r} s_k x_k \otimes z_k \tag{2.8.4}$$

for positive real coefficients $s_1, \ldots, s_r$, and orthonormal sets $\{x_1, \ldots, x_r\} \subset \mathcal{X}$, $\{z_1, \ldots, z_r\} \subset \mathcal{Y}$. An expression of this form is known as the *Schmidt decomposition* of the vector u. The Schmidt decomposition shows that any bipartite vector u can be represented as a sum of tensor products of orthonormal vectors, scaled by positive coefficients. This is extremely useful in quantum information, especially for analyzing entanglement.

Example 2.8.2. Consider the maximally entangled Bell state in $\mathbb{C}^2 \otimes \mathbb{C}^2$:

$$u = \frac{1}{\sqrt{2}} \left(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle \right) \tag{2.8.5}$$

This can be written in Schmidt form as:

$$u = \sum_{k=1}^{2} \frac{1}{\sqrt{2}} x_k \otimes z_k \tag{2.8.6}$$

with:

- $x_1 = |0\rangle, \quad z_1 = |0\rangle$
- $x_2 = |1\rangle, \quad z_2 = |1\rangle$
- $s_1 = s_2 = \frac{1}{\sqrt{2}}$

Here, the Schmidt rank is $r = 2$, and all coefficients are equal, indicating maximal entanglement.

2.9 Schatten Norms of Operators

Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces. A norm $\|\cdot\|$ on the operator space $L(\mathcal{X},\mathcal{Y})$ is a function $\|A\|$ satisfying the following:

1. Positive definiteness: $\|A\| \geq 0$ for all $A \in L(\mathcal{X},\mathcal{Y})$, and $\|A\| = 0$ iff $A = 0$.
2. Scalability: For all scalars $\alpha \in \mathbb{C}$, $\|\alpha A\| = |\alpha|\|A\|$.
3. Triangle inequality: $\|A + B\| \leq \|A\| + \|B\|$ for all $A, B \in L(\mathcal{X},\mathcal{Y})$.

A particularly important family of operator norms is the *Schatten p-norms*, defined for $p \in [1,\infty]$ and $A \in L(\mathcal{X},\mathcal{Y})$ by:

$$\|A\|_p = \left(\mathrm{Tr}\left((A^*A)^{p/2}\right)\right)^{1/p}. \tag{2.9.1}$$

The Schatten ∞-norm is defined as the largest singular value:

$$\|A\|_\infty = \max\{\|Au\| : u \in \mathcal{X}, \|u\| \leq 1\}. \tag{2.9.2}$$

Alternatively, all Schatten norms can be characterized in terms of the singular values $s(A)$ of A:

$$\|A\|_p = \|s(A)\|_p. \tag{2.9.3}$$

2.9.1 Properties of Schatten Norms

1. Monotonicity: If $1 \leq p \leq q \leq \infty$, then
$$\|A\|_p \geq \|A\|_q. \tag{2.9.4}$$
2. Inequalities: For any nonzero $A \in L(\mathcal{X},\mathcal{Y})$:
$$\|A\|_1 \leq (\mathrm{rank}(A))^{1-\frac{1}{p}}\|A\|_p, \tag{2.9.5}$$
$$\|A\|_q \leq (\mathrm{rank}(A))^{1/q}\|A\|_\infty. \tag{2.9.6}$$
3. Unitary invariance: For all unitary operators $U \in \mathcal{U}(\mathcal{Y},\mathcal{Z})$, $V \in \mathcal{U}(\mathcal{X},\mathcal{W})$:
$$\|A\|_p = \|UAV^*\|_p. \tag{2.9.7}$$
4. Duality: For $\frac{1}{p} + \frac{1}{p^*} = 1$,
$$\|A\|_p = \max_{B\in L(\mathcal{X},\mathcal{Y}),\ \|B\|_{p^*}\leq 1} |\langle B, A\rangle|. \tag{2.9.8}$$
This leads to the *Hölder inequality*:
$$|\langle B, A\rangle| \leq \|A\|_p\|B\|_{p^*}. \tag{2.9.9}$$
5. Submultiplicativity: For operators $A \in L(\mathcal{Z},\mathcal{W})$, $B \in L(\mathcal{Y},\mathcal{Z})$, $C \in L(\mathcal{X},\mathcal{Y})$,
$$\|ABC\|_p \leq \|A\|_\infty\|B\|_p\|C\|_\infty. \tag{2.9.10}$$
Hence,
$$\|AB\|_p \leq \|A\|_p\|B\|_p. \tag{2.9.11}$$
6. Invariance under adjoint and transpose:
$$\|A\|_p = \|A^*\|_p = \|A^T\|_p. \tag{2.9.12}$$

2.9.2 Common Schatten Norms

1. Spectral Norm ($p = \infty$):
$$\|A\| = \sup_{\|u\| \leq 1} \|Au\| = \sigma_{\max}(A). \tag{2.9.13}$$
Moreover,
$$\|A^*A\|_\infty = \|AA^*\|_\infty = \|A\|^2. \tag{2.9.14}$$

2. Frobenius Norm ($p = 2$):
$$\|A\|_2 = \sqrt{\mathrm{Tr}(A^*A)} = \sqrt{\langle A, A \rangle} = \|\mathrm{vec}(A)\|. \tag{2.9.15}$$

3. Trace Norm ($p = 1$):
$$\|A\|_1 = \mathrm{Tr}(\sqrt{A^*A}) = \sum_k \sigma_k(A). \tag{2.9.16}$$
For any unitary $U \in \mathcal{U}(\mathcal{X})$:
$$\|X\|_1 = \max_{U \in \mathcal{U}(\mathcal{X})} |\langle U, X \rangle|. \tag{2.9.17}$$
And for any $X \in L(\mathcal{X} \otimes \mathcal{Y})$:
$$\|\mathrm{Tr}_{\mathcal{Y}}(X)\|_1 \leq \|X\|_1. \tag{2.9.18}$$

2.9.3 Useful Identity

For unit vectors u, v and non-negative reals α, β, define:

$$\|\alpha uu^* - \beta vv^*\|_1 = \sqrt{(\alpha + \beta)^2 - 4\alpha\beta |\langle u, v \rangle|^2}. \tag{2.9.19}$$

This follows from the spectral decomposition of a rank-2 Hermitian operator:

$$\text{Eigenvalues: } \frac{\alpha - \beta}{2} \pm \frac{1}{2}\sqrt{(\alpha + \beta)^2 - 4\alpha\beta |\langle u, v \rangle|^2}. \tag{2.9.20}$$

In particular, when $\alpha = \beta = 1$,

$$\|uu^* - vv^*\|_1 = 2\sqrt{1 - |\langle u, v \rangle|^2}, \tag{2.9.21}$$

which quantifies the trace distance between pure states.

2.10 Review of Mathematical Analysis and Convexity

This section provides a concise review of essential notions from real and complex analysis, as well as convex analysis, which are crucial for understanding results in quantum information theory. While our focus is on finite-dimensional (Euclidean) spaces, these ideas generalize to more abstract settings.

2.10.1 Basic Concepts in Analysis

Let $\mathcal{V}$ be a real or complex Euclidean space, equipped with a norm $\|\cdot\|$. While we may assume the Euclidean norm, the results remain valid under any fixed norm due to the equivalence of norms in finite dimensions.

Open and Closed Sets

For $u \in \mathcal{V}$ and $r > 0$, define:

$$B_r(u) = \{v \in \mathcal{V} : \|u - v\| < r\} \quad \text{(open ball)} \tag{2.10.1}$$

$$S_r(u) = \{v \in \mathcal{V} : \|u - v\| = r\} \quad \text{(sphere)} \tag{2.10.2}$$

$$\overline{B}_r(u) = B_r(u) \cup S_r(u) \quad \text{(closed ball)} \tag{2.10.3}$$

A subset $A \subseteq \mathcal{V}$ is called *open* if for each $u \in A$, there exists $\varepsilon > 0$ such that $B_\varepsilon(u) \subseteq A$. It is called *closed* if its complement is open.

Relative Openness and Closure: If $B \subseteq A \subseteq \mathcal{V}$, we say B is open (resp. closed) *relative to* A if there exists an open (resp. closed) set $U \subseteq \mathcal{V}$ such that $B = A \cap U$. The *closure* of B relative to A is the smallest closed subset of A that contains B. B is said to be *dense in* A if its relative closure equals A.

Continuity

Let $f : A \longrightarrow \mathcal{W}$ be a function from a subset $A \subseteq \mathcal{V}$ into another Euclidean space $\mathcal{W}$. We say f is *continuous at* $u \in A$ if for every $\varepsilon > 0$, there exists $\delta > 0$ such that:

$$\|f(v) - f(u)\| < \varepsilon \quad \text{for all } v \in B_\delta(u) \cap A \tag{2.10.4}$$

Equivalently, the continuity condition can be expressed as:

$$(\forall \varepsilon > 0)(\exists \delta > 0) : f(B_\delta(u) \cap A) \subseteq B_\varepsilon(f(u)) \tag{2.10.5}$$

Topological Characterization: A function f is continuous on A if the preimage of every open set in $\mathcal{W}$ is open relative to A, or equivalently, the preimage of every closed set is closed relative to A.

Sequences and Convergence

A *sequence* in A is a function $s : \mathbb{N} \longrightarrow A$, usually denoted by $(u_n)_{n \in \mathbb{N}}$. We say (u_n) *converges* to $u \in \mathcal{V}$ if:

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ such that } \|u_n - u\| < \varepsilon \quad \text{for all } n \geq N \tag{2.10.6}$$

A *subsequence* (v_n) of (u_n) is formed by selecting a strictly increasing index sequence (k_n) such that $v_n = u_{k_n}$ for all n.

2.10.2 Compact Sets

A set $A \subseteq \mathcal{V}$ is said to be *compact* if every sequence in A has a convergent subsequence whose limit is also in A.

Heine-Borel Theorem: In Euclidean spaces, a set is compact if and only if it is closed and bounded:

$$A \text{ compact} \iff A \text{ is closed and bounded} \tag{2.10.7}$$

Useful Properties of Compact Sets:

- If A is compact and $f : A \longrightarrow \mathbb{R}$ is continuous, then f attains its maximum and minimum on A.
- If $A \subseteq \mathcal{V}$ is compact and $f : \mathcal{V} \longrightarrow \mathcal{W}$ is continuous, then $f(A)$ is also compact.

2.10.3 Convex Sets and Convexity

Convexity plays a central role in quantum information. This section reviews core ideas and results.

Definitions and Examples

Let $\mathcal{V}$ be a real Euclidean space. A set $A \subseteq \mathcal{V}$ is called *convex* if:

$$\forall u, v \in A, \forall \lambda \in [0,1] : \lambda u + (1-\lambda)v \in A \tag{2.10.8}$$

This means the line segment connecting any two points in A remains entirely within A. A point $w \in A$ is an *extreme point* of A if whenever:

$$w = \lambda u + (1-\lambda)v, \quad \lambda \in (0,1), \quad u, v \in A \tag{2.10.9}$$

it follows that $u = v = w$. A set A is a *cone* if $\forall u \in A, \lambda \geq 0 \Rightarrow \lambda u \in A$. A *convex cone* is a set that is both a cone and convex.

Example 2.10.1.

- The set $\text{Pos}(\mathcal{X})$ of positive semidefinite operators is a convex cone. It is closed under addition and scalar multiplication by non-negative reals.
- The set $\mathcal{D}(\mathcal{X})$ of density matrices is convex but not a cone. Its extreme points are pure states uu^* where u is a unit vector.

Probability Vectors and Convex Combinations: Let Σ be a finite, nonempty set. A vector $p \in \mathbb{R}^{\Sigma}$ is a *probability vector* if:

$$p(a) \geq 0 \, \forall a \in \Sigma, \quad \sum_{a \in \Sigma} p(a) = 1 \tag{2.10.10}$$

A *convex combination* of vectors $\{u_a : a \in \Sigma\} \subseteq A$ is:

$$\sum_{a \in \Sigma} p(a) u_a \tag{2.10.11}$$

The *convex hull* of A, denoted $\text{conv}(A)$, is the set of all convex combinations of points in A.

Important Theorems in Convex Analysis

Carathéodory's Theorem:

Theorem 2.10.2. Let $A \subseteq \mathbb{R}^d$. Every point $u \in \text{conv}(A)$ can be written as a convex combination of at most $d+1$ points from A.

Sion's Minimax Theorem:

Theorem 2.10.3. Let $A, B \subseteq \mathbb{R}^d$ be compact convex sets. Then:

$$\min_{u \in A} \max_{v \in B} \langle u, v \rangle = \max_{v \in B} \min_{u \in A} \langle u, v \rangle \tag{2.10.12}$$

Separating Hyperplane Theorem:

Theorem 2.10.4. Let $A \subseteq \mathbb{R}^d$ be closed and convex, and let $u \notin A$. Then there exists $v \in \mathbb{R}^d$ such that:

$$\langle v, w \rangle > \langle v, u \rangle \quad \text{for all } w \in A \tag{2.10.13}$$

Krein-Milman Theorem: If A is compact and convex in $\mathbb{R}^d$, then $A = \text{conv}(\text{Ext}(A))$, where $\text{Ext}(A)$ denotes the set of extreme points of A.

Chapter 3

Theory of Quantum Information

3.1 Quantum Information Theory

Quantum information theory delves into the behavior of idealized physical systems called registers. This field establishes the fundamental concepts of states (describing a register's condition), measurements (extracting classical information from states), and channels (transforming one register's state into another's). These definitions collectively form the core model for understanding quantum information.

3.1.1 Registers and Classical states

A *register* X is defined inductively as either:

1. A finite set Σ, called a *simple register*, or
2. An n-tuple $X = (Y_1, \ldots, Y_n)$ where each Y_k is itself a register; such X is called a *compound register*.

Notation:

Registers are denoted using capital sans-serif letters (e.g., $\mathsf{X}, \mathsf{Y}_1, \mathsf{Z}$), and often indexed to represent structured compositions.

Intuition:

This definition reflects the compositional nature of registers—multiple registers can be viewed as components of a larger one, forming a tree-like structure. Each leaf in this tree corresponds to a simple register.

Example 3.1.1 (Composite Register Tree)**.** Define the following registers:

$$\mathsf{X} = (\mathsf{Y}_0, \mathsf{Y}_1) \tag{3.1.1}$$

$$\mathsf{Y}_0 = \{1, 2, 3, 4\} \tag{3.1.2}$$

$$\mathsf{Y}_1 = (\mathsf{Z}_1, \mathsf{Z}_2, \mathsf{Z}_3) \tag{3.1.3}$$

$$\mathsf{Z}_1 = \mathsf{Z}_2 = \mathsf{Z}_3 = \{0, 1\} \tag{3.1.4}$$

Then X is a compound register whose structure is naturally represented as a tree. Its subregisters include: X, Y_0, Y_1, and each of $\mathsf{Z}_1, \mathsf{Z}_2, \mathsf{Z}_3$.

Classical State Sets

Let X be a register. Its *classical state set*, denoted Σ_X, is defined as follows:

1. If $X = \Sigma$ is a simple register, then $\Sigma_X = \Sigma$.

2. If $X = (Y_1, \dots, Y_n)$ is a compound register, then

$$\Sigma_X = \Sigma_{Y_1} \times \cdots \times \Sigma_{Y_n} \tag{3.1.5}$$

Each element $a \in \Sigma_X$ is called a *classical state* of the register X.

Interpretation:

A classical state of a compound register assigns a classical state to each of its components.

Trivial Register:

A register is said to be *trivial* if its classical state set contains only one element. Such registers are allowed, although they carry no useful information. Registers with *empty* state sets are disallowed.

State Reductions and Subregisters

Let $X = (Y_1, \dots, Y_n)$ be a compound register with classical state set

$$\Sigma_X = \Sigma_{Y_1} \times \cdots \times \Sigma_{Y_n} \tag{3.1.6}$$

Then any classical state $a = (b_1, \dots, b_n) \in \Sigma_X$ uniquely determines the classical state $b_k \in \Sigma_{Y_k}$ of each subregister Y_k. If $Z = (Y_{k_1}, \dots, Y_{k_m})$ for $1 \leq k_1 < \cdots < k_m \leq n$, then the classical state of Z corresponding to $a = (b_1, \dots, b_n) \in \Sigma_X$ is:

$$(b_{k_1}, \dots, b_{k_m}) \in \Sigma_{Y_{k_1}} \times \cdots \times \Sigma_{Y_{k_m}} = \Sigma_Z \tag{3.1.7}$$

Key Property:

The classical state of any register is uniquely determined by the classical states of its simple (leaf) subregisters. Conversely, a full state of a compound register induces consistent states on all subregisters via projection.

3.1.2 Quantum Registers

Quantum states generalize classical probabilistic states by embedding them into a richer mathematical framework rooted in linear algebra and operator theory. We begin by recalling how classical information is represented probabilistically, and then move on to the quantum generalization.

Probabilistic States of a Register

Let X be a classical register that takes values in a finite set Σ, referred to as the *alphabet* of X. A probabilistic state of X is a probability distribution over Σ. This is a vector:

$$p \in \mathsf{P}(\Sigma), \tag{3.1.8}$$

where $p(a) \geq 0$ for all $a \in \Sigma$ and $\sum_{a \in \Sigma} p(a) = 1$. The entry $p(a)$ is the probability that register X holds value $a \in \Sigma$.

Quantum States: The Mathematical Model

A *quantum state* is a density operator on a complex Euclidean space. These operators encode all information about a quantum system, including superposition, coherence, and mixedness. For now, we treat quantum states as purely mathematical objects, abstracting away physical interpretation.

Complex Euclidean Space of a Register

Given an alphabet Σ, define:

$$\mathcal{X} = \mathbb{C}^{\Sigma}, \tag{3.1.9}$$

i.e., the Hilbert space with orthonormal basis indexed by Σ. This serves as the state space for register X. If $X = (Y_1, Y_2, \ldots, Y_n)$, with Y_i having alphabet Γ_i, then:

$$\mathcal{X} = \mathbb{C}^{\Gamma_1 \times \cdots \times \Gamma_n} = \mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_n, \tag{3.1.10}$$

where $\mathcal{Y}_i = \mathbb{C}^{\Gamma_i}$. The tensor product represents composite systems.

Definition of Quantum States

A quantum state of register X is a density operator:

$$\rho \in \mathsf{D}(\mathcal{X}), \tag{3.1.11}$$

i.e., $\rho \in \mathsf{Pos}(\mathcal{X})$ with $\mathrm{Tr}(\rho) = 1$.

Convex Mixtures of Quantum States

Given states $\{\rho_a\}_{a\in\Gamma} \subset \mathsf{D}(\mathcal{X})$ and a distribution $p \in \mathsf{P}(\Gamma)$, the convex combination:

$$\rho = \sum_{a\in\Gamma} p(a)\rho_a \tag{3.1.12}$$

is also a quantum state. This models a probabilistic mixture.

Ensembles of Quantum States

Alternatively, define:

$$\eta : \Gamma \to \mathsf{Pos}(\mathcal{X}) \tag{3.1.13}$$

such that:

$$\sum_{a\in\Gamma} \mathrm{Tr}(\eta(a)) = 1. \tag{3.1.14}$$

Then $\rho = \sum_a \eta(a)$ is a valid density operator. The normalized components are:

$$\rho_a = \frac{\eta(a)}{\mathrm{Tr}(\eta(a))} \quad \text{if } \eta(a) \neq 0. \tag{3.1.15}$$

Pure States

A density operator $\rho \in \mathsf{D}(\mathcal{X})$ is *pure* if:

$$\rho = uu^*, \tag{3.1.16}$$

for unit vector $u \in \mathcal{X}$. Global phases don't change the state:

$$(\alpha u)(\alpha u)^* = |\alpha|^2 uu^* = uu^*, \quad \text{for } |\alpha| = 1. \tag{3.1.17}$$

Flat States

Let $\Pi \in \mathsf{Proj}(\mathcal{X})$ be a nonzero projection. Then the flat state is:

$$\rho = \frac{\Pi}{\mathrm{Tr}(\Pi)}. \tag{3.1.18}$$

If $V \subseteq \mathcal{X}$, the flat state on V is:

$$\omega_V = \frac{\Pi_V}{\mathrm{Tr}(\Pi_V)}. \tag{3.1.19}$$

The completely mixed state is:

$$\omega = \frac{\mathbb{1}_{\mathcal{X}}}{\dim(\mathcal{X})}. \tag{3.1.20}$$

Classical States as Quantum States

A classical value $a \in \Sigma$ corresponds to:

$$\rho = E_{a,a}, \tag{3.1.21}$$

where $E_{a,a}$ is the projection onto $e_a \in \mathcal{X}$. For distribution $p \in \mathsf{P}(\Sigma)$, the quantum state is:

$$\rho = \sum_{a \in \Sigma} p(a) E_{a,a} = \mathrm{Diag}(p). \tag{3.1.22}$$

Product and Correlated States

Given quantum registers $Y_1, \ldots, Y_n$ with $\mathcal{Y}_i$, a product state is:

$$\rho = \sigma_1 \otimes \cdots \otimes \sigma_n, \quad \sigma_i \in \mathsf{D}(\mathcal{Y}_i). \tag{3.1.23}$$

If $\rho \in \mathsf{D}(\mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_n)$ is not a tensor product, the registers are correlated (or entangled).

Example: Product vs Correlated States

Let Y, Z be qubit registers with alphabet $\{0, 1\}$.

Product state:

$$\rho = \left(\frac{1}{2} E_{0,0} + \frac{1}{2} E_{1,1} \right) \otimes \left(\frac{1}{2} E_{0,0} + \frac{1}{2} E_{1,1} \right). \tag{3.1.24}$$

Classically correlated state:

$$\sigma = \frac{1}{2} E_{0,0} \otimes E_{0,0} + \frac{1}{2} E_{1,1} \otimes E_{1,1}. \tag{3.1.25}$$

Entangled state:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle), \quad \rho = |\Phi^+\rangle \langle \Phi^+| . \tag{3.1.26}$$

3.1.3 Quantum Channels

Quantum channels are mathematical models for idealized physical processes that transform quantum states from one system (or register) to another. Formally, a quantum channel from a complex Euclidean space (register) $\mathcal{X}$ to another $\mathcal{Y}$ is a linear map:

$$\Phi : L(\mathcal{X}) \to L(\mathcal{Y}) \tag{3.1.27}$$

that satisfies two key properties:

1. Trace-preservation: the trace of any operator is preserved under the action of Φ, and
2. Complete positivity: the map remains positive even when extended by tensoring with an identity operation on an arbitrary auxiliary system.

These two requirements ensure that a quantum channel transforms density operators into valid density operators, not just in isolation, but even when part of a larger entangled system. When such a channel Φ acts on a register $\mathcal{X}$, we conceptually think of $\mathcal{X}$ as being replaced by $\mathcal{Y}$, whose state becomes $\Phi(\rho)$ if the input state was $\rho \in D(\mathcal{X})$. It's worth noting that nothing prevents $\mathcal{X} = \mathcal{Y}$; in that case, the channel maps a register to itself, possibly altering its state. A basic example is the *identity channel* $\mathbf{1}_{L(\mathcal{X})}$, which leaves all operators unchanged. This represents an idealized noiseless system or memory. Just as quantum states and measurements extend to multipartite systems via tensor products, channels also compose in this way. Suppose we have registers $\mathcal{X}_1, \ldots, \mathcal{X}_n$ and $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$, and channels:

$$\Phi_1 : L(\mathcal{X}_1) \to L(\mathcal{Y}_1), \quad \ldots, \quad \Phi_n : L(\mathcal{X}_n) \to L(\mathcal{Y}_n) \tag{3.1.28}$$

Then the *product channel* is defined as:

$$\Phi_1 \otimes \cdots \otimes \Phi_n : L(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n) \to L(\mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_n) \tag{3.1.29}$$

This product channel corresponds to applying each Φ_k independently on its respective subsystem. Now let's examine why complete positivity and trace preservation are necessary. If $\Phi : L(\mathcal{X}) \to L(\mathcal{Y})$ is a valid quantum channel, and $\rho \in D(\mathcal{X})$ is any input density operator, then $\Phi(\rho) \in D(\mathcal{Y})$ must be a valid output density operator. But this must also hold when Φ acts as part of a larger system. That is, for any auxiliary register $\mathcal{Z}$, the extended map $\Phi \otimes \mathbf{1}_{L(\mathcal{Z})}$ must satisfy:

$$(\Phi \otimes \mathbf{1}_{L(\mathcal{Z})})(\rho) \in D(\mathcal{Y} \otimes \mathcal{Z}) \quad \text{for all } \rho \in D(\mathcal{X} \otimes \mathcal{Z}) \tag{3.1.30}$$

This condition ensures that Φ is *completely positive*, meaning that positivity is preserved under all such extensions. Moreover, trace preservation means:

$$\mathrm{Tr}(\Phi(X)) = \mathrm{Tr}(X) \quad \text{for all } X \in L(\mathcal{X}) \tag{3.1.31}$$

Once complete positivity is enforced, it's straightforward to show that any tensor product of such channels will again map density operators to density operators. This stems from the composability of such maps. Specifically, the product channel:

$$\Phi_1 \otimes \cdots \otimes \Phi_n \tag{3.1.32}$$

can be decomposed into a sequence of tensor-product operations with identity maps as:

$$\Phi_1 \otimes \cdots \otimes \Phi_n = (\Phi_1 \otimes \mathbf{1}_{L(\mathcal{X}_2)} \otimes \cdots \otimes \mathbf{1}_{L(\mathcal{X}_n)}) \cdots (\mathbf{1}_{L(\mathcal{Y}_1)} \otimes \cdots \otimes \mathbf{1}_{L(\mathcal{Y}_{n-1})} \otimes \Phi_n) \tag{3.1.33}$$

Each step in this composition is a tensor product of a channel and identity maps. Since both complete positivity and trace preservation are preserved under composition, the full product channel also transforms density operators into valid density operators.

Kraus Representation

Every quantum channel $\Phi : L(\mathcal{X}) \to L(\mathcal{Y})$ admits a Kraus (operator-sum) representation:

$$\Phi(X) = \sum_{k=1}^{r} A_k X A_k^\dagger, \qquad \text{for all } X \in L(\mathcal{X}) \tag{3.1.34}$$

where $A_k \in L(\mathcal{X}, \mathcal{Y})$ satisfy the normalization condition:

$$\sum_{k=1}^{r} A_k^\dagger A_k = \mathbb{1}_{\mathcal{X}} \tag{3.1.35}$$

This representation ensures that Φ is completely positive and trace-preserving.

Stinespring Representation

The Stinespring dilation theorem states that any quantum channel $\Phi : L(\mathcal{X}) \to L(\mathcal{Y})$ can be written in the form:

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}(VXV^\dagger) \tag{3.1.36}$$

for some isometry $V \in L(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ and some auxiliary space $\mathcal{Z}$.

Choi Representation

The Choi matrix of a linear map $\Phi : L(\mathcal{X}) \to L(\mathcal{Y})$ is defined as:

$$J(\Phi) = (\Phi \otimes \mathbb{1}_{L(\mathcal{X})})(|\Omega\rangle \langle\Omega|) \tag{3.1.37}$$

where $|\Omega\rangle = \sum_i |i\rangle \otimes |i\rangle$ is the unnormalized maximally entangled state. Φ is completely positive if and only if $J(\Phi) \geq 0$, and Φ is trace-preserving if and only if $\mathrm{Tr}_{\mathcal{Y}}(J(\Phi)) = \mathbb{1}_{\mathcal{X}}$.

Some Standard Quantum Channels

(a) Identity Channel

$$\Phi(X) = X \tag{3.1.38}$$

This channel leaves the quantum state completely unchanged. It represents an ideal quantum memory or a noiseless communication channel.

(b) Depolarizing Channel

Let $\rho \in D(\mathbb{C}^d)$, $p \in [0,1]$:

$$\Phi(\rho) = (1-p)\rho + p\frac{\mathbb{1}}{d} \tag{3.1.39}$$

This channel replaces the state ρ with the maximally mixed state $\mathbb{1}/d$ with probability p, and otherwise leaves the state unchanged. It models uniform noise over all possible pure states.

(c) Bit-Flip Channel (Qubit)

$$\begin{aligned} A_0 &= \sqrt{1-p}\,\mathbb{1}, && (3.1.40)\\ A_1 &= \sqrt{p}\,X && (3.1.41)\\ \Phi(\rho) &= (1-p)\rho + pX\rho X && (3.1.42) \end{aligned}$$

This channel flips the state of a qubit from $|0\rangle$ to $|1\rangle$ and vice versa with probability p, simulating classical bit-flip noise.

(d) Amplitude Damping Channel

$$\begin{aligned} A_0 &= \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, && (3.1.43)\\ A_1 &= \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}, && (3.1.44)\\ \Phi(\rho) &= A_0\rho A_0^\dagger + A_1\rho A_1^\dagger && (3.1.45) \end{aligned}$$

This channel models energy dissipation, such as spontaneous emission of a photon by an excited atom. It represents irreversible decay from $|1\rangle$ to $|0\rangle$.

(e) Erasure Channel

Let $\mathcal{X} = \mathbb{C}^d$, $\mathcal{Y} = \mathbb{C}^d \oplus \mathbb{C}$:

$$\Phi(\rho) = (1-p)\rho + p\,|\text{err}\rangle\,\langle\text{err}| \tag{3.1.46}$$

This channel erases the input state and replaces it with a flag $|\text{err}\rangle$ orthogonal to the original space, with probability p. The receiver is notified that erasure has occurred.

(f) Phase Damping Channel

Describes pure dephasing:

$$\begin{aligned} A_0 &= \sqrt{1-p}\,\mathbb{1}, && (3.1.47)\\ A_1 &= \sqrt{p}\,Z && (3.1.48)\\ \Phi(\rho) &= A_0\rho A_0^\dagger + A_1\rho A_1^\dagger && (3.1.49) \end{aligned}$$

This channel destroys quantum coherence between $|0\rangle$ and $|1\rangle$ without changing the population of states. It models loss of phase information due to environmental noise.

3.1.4 Quantum Measurement

Measurement is a fundamental concept in quantum information theory. It describes the process by which information is extracted from a quantum system, and typically results in a probabilistic outcome along with a modified post-measurement state.

Projective Measurements

Let $\mathcal{X}$ be a complex Euclidean space associated with a quantum system. A *projective measurement* is described by a collection of orthogonal projection operators $\{\Pi_a : a \in \Sigma\} \subset \mathrm{Pos}(\mathcal{X})$ satisfying:

$$\sum_{a\in\Sigma} \Pi_a = \mathbb{1}_{\mathcal{X}}, \qquad \Pi_a \Pi_b = \delta_{ab} \Pi_a \tag{3.1.50}$$

Given a state $\rho \in D(\mathcal{X})$, the probability of obtaining outcome $a \in \Sigma$ upon measuring is:

$$\Pr[a] = \mathrm{Tr}(\Pi_a \rho) \tag{3.1.51}$$

and the post-measurement state, conditioned on obtaining outcome a, is:

$$\rho_a = \frac{\Pi_a \rho \Pi_a}{\mathrm{Tr}(\Pi_a \rho)} \tag{3.1.52}$$

General Measurements (POVMs)

More generally, a *quantum measurement* is described by a collection of positive semidefinite operators $\{M_a : a \in \Sigma\} \subset \mathrm{Pos}(\mathcal{X})$, known as a *Positive Operator-Valued Measure* (POVM), satisfying:

$$\sum_{a\in\Sigma} M_a = \mathbb{1}_{\mathcal{X}} \tag{3.1.53}$$

The probability of outcome a upon measuring $\rho \in D(\mathcal{X})$ is given by:

$$\Pr[a] = \mathrm{Tr}(M_a \rho) \tag{3.1.54}$$

For POVMs that arise from a physical process, there exists an associated set of measurement operators $\{K_a : a \in \Sigma\} \subset L(\mathcal{X})$, such that:

$$M_a = K_a^\dagger K_a \tag{3.1.55}$$

and the post-measurement state, conditioned on outcome a, is:

$$\rho_a = \frac{K_a \rho K_a^\dagger}{\mathrm{Tr}(K_a^\dagger K_a \rho)} = \frac{K_a \rho K_a^\dagger}{\Pr[a]} \tag{3.1.56}$$

Measurements on Composite Systems and the Partial Trace

Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, and let $\rho \in D(\mathcal{X} \otimes \mathcal{Y})$ be the state of a bipartite system. The *partial trace* over $\mathcal{Y}$ is a linear map:

$$\mathrm{Tr}_{\mathcal{Y}} : L(\mathcal{X} \otimes \mathcal{Y}) \to L(\mathcal{X}) \tag{3.1.57}$$

which satisfies the identity:

$$\mathrm{Tr}(A\, \mathrm{Tr}_{\mathcal{Y}}(\rho)) = \mathrm{Tr}((A \otimes \mathbb{1}_{\mathcal{Y}})\rho) \quad \text{for all } A \in L(\mathcal{X}) \tag{3.1.58}$$

The reduced state of subsystem $\mathcal{X}$ is given by:

$$\rho_{\mathcal{X}} = \mathrm{Tr}_{\mathcal{Y}}(\rho) \tag{3.1.59}$$

Measurements on a Subsystem

Suppose a projective measurement $\{\Pi_a\} \subset \mathrm{Pos}(\mathcal{X})$ is applied to the $\mathcal{X}$-part of $\rho \in D(\mathcal{X} \otimes \mathcal{Y})$. The probability of outcome a is:

$$\Pr[a] = \mathrm{Tr}((\Pi_a \otimes \mathbb{1}_{\mathcal{Y}})\rho) \tag{3.1.60}$$

The corresponding post-measurement state (conditioned on outcome a) is:

$$\rho_a = \frac{(\Pi_a \otimes \mathbb{1}_{\mathcal{Y}})\rho(\Pi_a \otimes \mathbb{1}_{\mathcal{Y}})}{\Pr[a]} \tag{3.1.61}$$

The unconditioned post-measurement state (after discarding measurement outcomes) is:

$$\rho' = \sum_{a \in \Sigma} (\Pi_a \otimes \mathbb{1}_{\mathcal{Y}})\rho(\Pi_a \otimes \mathbb{1}_{\mathcal{Y}}) \tag{3.1.62}$$

3.2 Quantum Information and Entropy

3.2.1 Quantum Entropy

Suppose Alice prepares a quantum system A in a state described by a density operator $\rho_A \in \mathcal{D}(\mathcal{H}_A)$. The *entropy* of this state, also known as the *von Neumann entropy* or *quantum entropy*, is defined by the expression:

$$\mathrm{H}(A)_\rho := -\mathrm{Tr}(\rho_A \log \rho_A). \tag{3.2.1}$$

This quantity, which we often refer to simply as the entropy of the system, is typically denoted as $\mathrm{H}(A)_\rho$ or $\mathrm{H}(\rho_A)$ to indicate its dependence on the density operator ρ_A. The quantum entropy is closely related to the spectrum (i.e., the eigenvalues) of the state ρ_A.

Mathematical Properties of Quantum Entropy

We now present several key properties of the quantum (von Neumann) entropy, including its non-negativity, extremal values, invariance under isometries, and concavity. Many of these results reflect analogous properties from classical information theory, due to the fact that the von Neumann entropy of a state depends only on the eigenvalues of its density operator.

1. Non-Negativity For any density operator $\rho \in \mathcal{D}(\mathcal{H})$, the entropy is non-negative:

$$\mathrm{H}(\rho) \geq 0. \tag{3.2.2}$$

Proof. This follows from the non-negativity of Shannon entropy, since $\mathrm{H}(\rho)$ is equal to the Shannon entropy of the eigenvalue distribution of ρ. □

2. Minimum Value The quantum entropy achieves its minimum value of zero if and only if ρ is a pure state:

$$\mathrm{H}(\rho) = 0 \iff \rho = |\psi\rangle\langle\psi| \text{ for some } |\psi\rangle \in \mathcal{H}. \tag{3.2.3}$$

Proof. The entropy vanishes when all the weight of the eigenvalue distribution is concentrated on a single eigenstate, implying that ρ is rank-one (i.e., pure). Although quantum systems can exhibit uncertainty, a known pure state carries no uncertainty from the perspective of an observer who knows the preparation. For instance, if Alice prepares a known state $|\phi\rangle$, Bob can verify it by measuring using the projective measurement $\{|\phi\rangle\langle\phi|, I - |\phi\rangle\langle\phi|\}$, always obtaining the same result. Hence, the entropy is zero. □

3. Maximum Value Let $\mathcal{H}$ be a Hilbert space of dimension d. The quantum entropy of a state $\rho \in \mathcal{D}(\mathcal{H})$ satisfies

$$\mathrm{H}(\rho) \leq \log d, \tag{3.2.4}$$

with equality if and only if $\rho = \frac{1}{d}I$, the maximally mixed state.

Proof. This parallels the classical case, where entropy is maximized when the probability distribution is uniform. In the quantum case, this occurs when all eigenvalues of ρ are equal to $1/d$. □

4. Concavity Let $\{\rho_x \in \mathcal{D}(\mathcal{H})\}$ be a collection of density operators and let $\{p_X(x)\}$ be a probability distribution. Then the entropy is a concave function of the density operator:

$$\mathrm{H}\left(\sum_x p_X(x)\rho_x\right) \geq \sum_x p_X(x)\,\mathrm{H}(\rho_x). \tag{3.2.5}$$

Proof. This follows from the concavity of the Shannon entropy and the joint convexity of relative entropy. Physically, this expresses that mixing quantum states cannot decrease entropy. □

5. Invariance under Isometries Let $\rho \in \mathcal{D}(\mathcal{H})$, and let $U : \mathcal{H} \to \mathcal{H}'$ be an isometry. Then:

$$\mathrm{H}(\rho) = \mathrm{H}(U\rho U^\dagger). \tag{3.2.6}$$

Proof. Isometries preserve the eigenvalue spectrum of the density operator. If $\rho = \sum_x p_X(x)|x\rangle\langle x|$, then

$$U\rho U^\dagger = \sum_x p_X(x)\,U|x\rangle\langle x|U^\dagger = \sum_x p_X(x)|\phi_x\rangle\langle\phi_x|, \tag{3.2.7}$$

where $|\phi_x\rangle = U|x\rangle$ forms an orthonormal set. Thus, $U\rho U^\dagger$ has the same eigenvalues as ρ, and the entropy remains unchanged. □

3.2.2 Joint Quantum Entropy

The joint quantum entropy of a bipartite state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is defined by

$$H(AB)_\rho := -\mathrm{Tr}\left\{\rho_{AB}\log\rho_{AB}\right\}. \tag{3.2.8}$$

If $\rho_{ABC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ is a tripartite state, then the entropy $H(AB)_\rho$ refers to the entropy of the reduced state $\rho_{AB} = \mathrm{Tr}_C\{\rho_{ABC}\}$.

Marginal Entropies of a Pure Bipartite State

A fundamental distinction between quantum and classical entropy arises in the case of pure bipartite states. In classical information theory, the joint entropy is never less than the marginal entropies. However, this does not hold in the quantum setting.

Theorem 3.2.1 (Marginal Entropies of a Pure Bipartite State)**.** Let $|\phi\rangle_{AB}$ be a pure bipartite state. Then the entropy of each marginal is equal:

$$H(A)_\phi = H(B)_\phi, \tag{3.2.9}$$

while the joint entropy vanishes:

$$H(AB)_\phi = 0. \tag{3.2.10}$$

Proof. Using the Schmidt decomposition, any pure bipartite state can be expressed as

$$|\phi\rangle_{AB} = \sum_i \sqrt{\lambda_i}\,|i\rangle_A\,|i\rangle_B\,, \tag{3.2.11}$$

where $\{\lambda_i\}$ form a probability distribution, and $\{|i\rangle_A\}$, $\{|i\rangle_B\}$ are orthonormal sets. The reduced states are then

$$\rho_A = \sum_i \lambda_i\,|i\rangle\,\langle i|_A\,, \tag{3.2.12}$$

$$\rho_B = \sum_i \lambda_i\,|i\rangle\,\langle i|_B\,. \tag{3.2.13}$$

Since both have identical spectra, we have $H(\rho_A) = H(\rho_B)$. Furthermore, as $|\phi\rangle_{AB}$ is pure, $H(\rho_{AB}) = 0$. □

This property extends to larger systems under bipartite cuts. For a pure state $|\phi\rangle_{ABCDE}$, one obtains relations like:

$$\begin{aligned} H(A)_\phi &= H(BCDE)_\phi, && (3.2.14)\\ H(AB)_\phi &= H(CDE)_\phi, && (3.2.15)\\ H(ABC)_\phi &= H(DE)_\phi, && (3.2.16)\\ H(ABCD)_\phi &= H(E)_\phi. && (3.2.17) \end{aligned}$$

There is no exact classical analogue of this behavior. For instance, copying a classical variable X yields $H(X\hat{X}) = H(X)$, not zero.

Additivity of Quantum Entropy

Let $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ and $\sigma_B \in \mathcal{D}(\mathcal{H}_B)$. Then

$$H(\rho_A \otimes \sigma_B) = H(\rho_A) + H(\sigma_B). \tag{3.2.18}$$

This follows from the fact that the eigenvalues of the tensor product are products of eigenvalues of the individual density operators, and the logarithm of a product splits additively.

Joint Entropy of a Classical–Quantum State

Let ρ_{XB} be a classical–quantum (CQ) state of the form

$$\rho_{XB} = \sum_x p_X(x) \, |x\rangle \, \langle x|_X \otimes \rho_B^x, \tag{3.2.19}$$

where X is classical and B is quantum.

Theorem 3.2.2. The joint entropy of a CQ state is given by

$$H(XB)_\rho = H(X) + \sum_x p_X(x) H(\rho_B^x), \tag{3.2.20}$$

where $H(X)$ is the Shannon entropy of p_X.

Proof. Start from

$$H(XB) = -\mathrm{Tr}\left\{\rho_{XB} \log \rho_{XB}\right\}. \tag{3.2.21}$$

Observe that

$$\log \rho_{XB} = \sum_x |x\rangle \, \langle x|_X \otimes \log(p_X(x)\rho_B^x) = \sum_x |x\rangle \, \langle x|_X \otimes [\log p_X(x) I + \log \rho_B^x]. \tag{3.2.22}$$

Hence,

$$\begin{aligned} H(XB)_\rho &= -\sum_x p_X(x) \mathrm{Tr}\left\{\rho_B^x \left[\log p_X(x) I + \log \rho_B^x\right]\right\} \\ &= -\sum_x p_X(x) \left[\log p_X(x) + \mathrm{Tr}\{\rho_B^x \log \rho_B^x\}\right] \\ &= H(X) + \sum_x p_X(x) H(\rho_B^x). \end{aligned} \tag{3.2.23}$$

□

3.2.3 Conditional Quantum Entropy

Given a bipartite quantum state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the conditional quantum entropy of subsystem A given B is defined analogously to the classical case:

$$H(A|B)_\rho := H(AB)_\rho - H(B)_\rho. \tag{3.2.24}$$

This formulation is both practical and conceptually aligned with classical information theory, enabling various structural properties to carry over naturally to the quantum setting.

Theorem 3.2.3 (Conditioning Does Not Increase Entropy)**.** For any bipartite state ρ_{AB}, conditioning reduces entropy:

$$H(A)_\rho \geq H(A|B)_\rho. \tag{3.2.25}$$

This reflects the intuition that access to additional information (here, system B) can only decrease uncertainty about system A.

Conditional Entropy in Classical–Quantum States

Consider a classical–quantum state of the form

$$\rho_{XB} = \sum_x p_X(x) \left|x\right\rangle \left\langle x\right|_X \otimes \rho_B^x, \tag{3.2.26}$$

where X is a classical register and ρ_B^x are quantum states indexed by x. The conditional entropy of B given X is then:

$$\begin{aligned} H(B|X)_\rho &= H(XB)_\rho - H(X)_\rho & (3.2.27)\\ &= H(X) + \sum_x p_X(x) H(\rho_B^x) - H(X) & (3.2.28)\\ &= \sum_x p_X(x) H(\rho_B^x). & (3.2.29) \end{aligned}$$

This expression resembles the classical conditional entropy and applies whenever the conditioning system is classical.

Negative Conditional Quantum Entropy

One of the distinctly quantum features of conditional entropy is that it can be negative, a phenomenon with no classical analogue. For instance, consider the maximally entangled Bell state:

$$\left|\Phi^+\right\rangle_{AB} = \frac{1}{\sqrt{2}}(\left|00\right\rangle + \left|11\right\rangle). \tag{3.2.30}$$

The joint state is pure, so $H(AB) = 0$. However, the reduced state on either subsystem is maximally mixed, giving $H(A) = H(B) = 1$. Applying the definition:

$$H(A|B) = H(AB) - H(B) = 0 - 1 = -1. \tag{3.2.31}$$

This negative entropy reflects strong quantum correlations (entanglement) and has no classical analogue.

3.2.4 Coherent Information and Conditional Entropy Bounds

Coherent Information

Given a bipartite state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the coherent information quantifies how much quantum information from A can be transmitted to B through a quantum channel. It is defined as:

$$I(A\rangle B)_\rho := H(B)_\rho - H(AB)_\rho. \tag{3.2.32}$$

This quantity is particularly significant in quantum communication theory and is intimately related to the quantum capacity of a channel.

Theorem 3.2.4 (Duality of Conditional Entropy)**.** The duality of conditional entropy in quantum information theory, specifically for a tripartite pure state $|\psi_{ABC}\rangle$, is expressed by the relation

$$H(A|B) = -H(A|C) \tag{3.2.33}$$

Bounds on Conditional Quantum Entropy

The conditional quantum entropy $H(A|B)_\rho$ of a bipartite state is bounded in absolute value by the logarithm of the dimension of system A. This reflects the limited amount of quantum information that can be shared or conditioned upon.

Theorem 3.2.5 (Bounds on Conditional Quantum Entropy)**.** Let $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Then the conditional entropy satisfies:

$$|H(A|B)_\rho| \leq \log \dim(\mathcal{H}_A). \tag{3.2.34}$$

This bound is tight, i.e., it can be achieved for specific states.

Proof. We first establish the upper bound. By the fact that conditioning doesn't increase entropy (Theorem 3.2.3), we have:

$$H(A|B)_\rho \leq H(A)_\rho. \tag{3.2.35}$$

Since the von Neumann entropy of ρ_A is maximized when ρ_A is the maximally mixed state $\pi_A = \frac{I_A}{\dim(\mathcal{H}_A)}$, it follows that:

$$H(A)_\rho \leq \log \dim(\mathcal{H}_A), \tag{3.2.36}$$

so altogether,

$$H(A|B)_\rho \leq \log \dim(\mathcal{H}_A). \tag{3.2.37}$$

To prove the lower bound, consider a purification $|\psi\rangle_{EAB}$ of ρ_{AB}. Then by the duality of conditional entropy, we have:

$$H(A|B)_\rho = -H(A|E)_\psi. \tag{3.2.38}$$

Again applying the conditioning-reduces-entropy principle:

$$-H(A|E)_\psi \geq -H(A)_\rho, \tag{3.2.39}$$

and using the fact that entropy is non-negative and upper-bounded by $\log \dim(\mathcal{H}_A)$:

$$H(A)_\rho \leq \log \dim(\mathcal{H}_A), \tag{3.2.40}$$

we obtain:

$$H(A|B)_\rho \geq -\log \dim(\mathcal{H}_A). \tag{3.2.41}$$

This completes the proof. □

Tightness of the Bound

The bounds are achieved in the following cases:

- When $\rho_{AB} = \pi_A \otimes \sigma_B$, where π_A is the maximally mixed state on system A, and σ_B is arbitrary. In this case, $H(A|B) = \log \dim(\mathcal{H}_A)$.
- When ρ_{AB} is a maximally entangled pure state (e.g., a Bell state), the joint entropy is zero, and each marginal is maximally mixed. Thus, $H(B) = \log \dim(\mathcal{H}_B)$ and $H(AB) = 0$, so $H(A|B) = -\log \dim(\mathcal{H}_A)$.

3.2.5 Coherent Information and Quantum Mutual Information

Given a bipartite state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the coherent information and mutual information are related by:

$$I(A;B)_\rho = H(A)_\rho + I(A\rangle B)_\rho, \tag{3.2.42}$$

$$I(A;B)_\rho = H(B)_\rho + I(B\rangle A)_\rho. \tag{3.2.43}$$

These identities show that mutual information not only captures the total correlations between subsystems A and B, but also decomposes naturally into entropy and directional coherent information.

Theorem 3.2.6 (Non-Negativity of Quantum Mutual Information)**.** Let ρ_{AB} be a bipartite quantum state. Then the quantum mutual information is always non-negative:

$$I(A;B)_\rho \geq 0. \tag{3.2.44}$$

It reflects the fact that total correlations—whether classical or quantum—cannot be less than zero. Quantum mutual information serves as a general measure of correlation, reducing to classical mutual information when the state is classical.

3.2.6 Holevo Information and Accessible Information

Suppose Alice prepares a classical–quantum ensemble

$$\mathcal{E} = \{p_X(x), \rho_B^x\}, \tag{3.2.45}$$

and sends it to Bob, who does not know the classical label x. From Bob's point of view, the ensemble appears as the average density operator

$$\rho_B = \mathbb{E}_X[\rho_B^X] = \sum_x p_X(x)\rho_B^x. \tag{3.2.46}$$

Bob attempts to learn the classical variable X by performing some measurement $\{\Lambda_y\}$ on system B. The *accessible information* is defined as:

$$I_{\text{acc}}(\mathcal{E}) = \max_{\{\Lambda_y\}} I(X;Y), \tag{3.2.47}$$

where Y is the outcome of the measurement. The *Holevo information* $\chi(\mathcal{E})$ of the ensemble $\mathcal{E}$ is defined as:

$$\chi(\mathcal{E}) = H(\rho_B) - \sum_x p_X(x)H(\rho_B^x). \tag{3.2.48}$$

The Holevo information quantifies the total amount of information about the classical variable X that is, in principle, encoded in the quantum system B, before any measurement is made.

Theorem 3.2.7 (Holevo Bound)**.** For any classical–quantum ensemble $\mathcal{E}$, the accessible information is bounded from above by the Holevo information:

$$I_{\text{acc}}(\mathcal{E}) \leq \chi(\mathcal{E}). \tag{3.2.49}$$

3.2.7 Conditional Quantum Mutual Information (CQMI)

Let $\rho_{ABC} \in D(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ be a tripartite quantum state. The *conditional quantum mutual information (CQMI)* is defined as

$$I(A;B \mid C)_\rho = H(A \mid C)_\rho + H(B \mid C)_\rho - H(AB \mid C)_\rho. \tag{3.2.50}$$

CQMI quantifies how much mutual information systems A and B share, given access to a third system C. It generalises the classical notion of conditional mutual information to the quantum setting.

Theorem 3.2.8 (Chain Rule for Quantum Mutual Information)**.** The quantum mutual information satisfies the chain rule:

$$I(A;BC)_\rho = I(A;B)_\rho + I(A;C \mid B)_\rho. \tag{3.2.51}$$

This property allows us to understand the total correlation between A and BC in terms of simpler two-party correlations. An important exercise using the chain rule is:

$$I(A;BC)_\rho = I(AC;B)_\rho + I(A;C)_\rho - I(B;C)_\rho. \tag{3.2.52}$$

Theorem 3.2.9 (Non-Negativity of CQMI / Strong Subadditivity)**.** Let $\rho_{ABC} \in D(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ be any tripartite quantum state. Then the conditional quantum mutual information is non-negative:

$$I(A;B \mid \mathrm{C})_\rho \geq 0. \tag{3.2.53}$$

This fundamental inequality is known as the *strong subadditivity* of quantum entropy. It plays a central role in quantum information theory, underpinning several important results such as the data-processing inequality, the Holevo bound, and more.

3.2.8 Quantum Relative Entropy

Kernel and Support

The kernel of an operator $A \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ is defined as:

$$\ker(A) \equiv \{|\psi\rangle \in \mathcal{H} : A|\psi\rangle = 0\}. \tag{3.2.54}$$

The support of A is the orthogonal complement of the kernel:

$$\mathrm{supp}(A) \equiv \ker(A)^\perp \equiv \{|\psi\rangle \in \mathcal{H} : \langle\psi|\phi\rangle = 0 \;\forall|\phi\rangle \in \ker(A)\}. \tag{3.2.55}$$

If A is Hermitian with spectral decomposition $A = \sum_{i:a_i \neq 0} a_i |i\rangle\langle i|$, then:

$$\mathrm{supp}(A) = \mathrm{span}\{|i\rangle : a_i \neq 0\}. \tag{3.2.56}$$

The projection onto the support of A is:

$$\Pi_A \equiv \sum_{i:a_i \neq 0} |i\rangle\langle i|. \tag{3.2.57}$$

Quantum Relative Entropy

Let $\rho \in \mathcal{D}(\mathcal{H})$ be a density operator and $\sigma \in \mathcal{L}(\mathcal{H})$ be positive semi-definite. The quantum relative entropy between ρ and σ is defined as:

$$D(\rho\|\sigma) \equiv \mathrm{Tr}\left\{\rho\left[\log\rho - \log\sigma\right]\right\}, \tag{3.2.58}$$

if $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$. Otherwise, $D(\rho\|\sigma) = +\infty$.

Theorem 3.2.10 (Monotonicity of Quantum Relative Entropy)**.** Let $\rho \in \mathcal{D}(\mathcal{H})$, $\sigma \in \mathcal{L}(\mathcal{H})$ be positive semi-definite, and $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ be a quantum channel. Then:

$$D(\rho\|\sigma) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)). \tag{3.2.59}$$

Theorem 3.2.11 (Non-Negativity of Quantum Relative Entropy)**.** Let $\rho \in \mathcal{D}(\mathcal{H})$ and $\sigma \in \mathcal{L}(\mathcal{H})$ be positive semi-definite such that $\mathrm{Tr}\{\sigma\} \leq 1$. Then:

$$D(\rho\|\sigma) \geq 0, \tag{3.2.60}$$

and $D(\rho\|\sigma) = 0$ if and only if $\rho = \sigma$.

Let $\rho \in \mathcal{D}(\mathcal{H})$ and $\sigma, \sigma' \in \mathcal{L}(\mathcal{H})$ be positive semi-definite. Suppose $\sigma \leq \sigma'$. Then:

$$D(\rho\|\sigma') \leq D(\rho\|\sigma). \tag{3.2.61}$$

Theorem 3.2.12 (Isometric Invariance)**.** Let $U : \mathcal{H} \longrightarrow \mathcal{H}'$ be an isometry. Then for all positive semi-definite $\rho, \sigma \in \mathcal{L}(\mathcal{H})$,

$$D(\rho\|\sigma) = D(U\rho U^\dagger \| U\sigma U^\dagger). \tag{3.2.62}$$

Theorem 3.2.13 (Additivity of Quantum Relative Entropy)**.** Let $\rho_1 \in \mathcal{D}(\mathcal{H}_1), \rho_2 \in \mathcal{D}(\mathcal{H}_2)$ and σ_1, σ_2 positive semi-definite. Then:

$$D(\rho_1 \otimes \rho_2 \| \sigma_1 \otimes \sigma_2) = D(\rho_1 \| \sigma_1) + D(\rho_2 \| \sigma_2). \tag{3.2.63}$$

Moreover,

$$D(\rho^{\otimes n} \| \sigma^{\otimes n}) = nD(\rho \| \sigma). \tag{3.2.64}$$

Theorem 3.2.14 (Quantum Relative Entropy for Classical–Quantum States)**.** Let

$$\rho_{XB} = \sum_x p(x) |x\rangle\langle x|_X \otimes \rho_B^x, \tag{3.2.65}$$

$$\sigma_{XB} = \sum_x q(x) |x\rangle\langle x|_X \otimes \sigma_B^x. \tag{3.2.66}$$

Then,

$$D(\rho_{XB} \| \sigma_{XB}) = \sum_x p(x) D(\rho_B^x \| \sigma_B^x) + D(p \| q). \tag{3.2.67}$$

Theorem 3.2.15 (Scaling Property of Relative Entropy)**.** Let $a, b > 0$, $\rho \in \mathcal{D}(\mathcal{H})$, and $\sigma \in \mathcal{L}(\mathcal{H})$ be positive semi-definite. Then:

$$D(a\rho \| b\sigma) = a \left[D(\rho \| \sigma) + \log \left(\frac{a}{b} \right) \right]. \tag{3.2.68}$$

3.3 Quantum Entropy Inequalities

Theorem 3.3.1 (Subadditivity)**.** Let $\rho_{AB} \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a bipartite quantum state. Then the von Neumann entropy satisfies the subadditivity inequality:

$$H(AB)_\rho \leq H(A)_\rho + H(B)_\rho. \tag{3.3.1}$$

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

Theorem 3.3.2 (Araki–Lieb Inequality)**.** For any bipartite quantum state $\rho_{AB} \in D(\mathcal{H}_A \otimes \mathcal{H}_B)$, the following inequality holds:

$$|H(A)_\rho - H(B)_\rho| \leq H(AB)_\rho. \tag{3.3.2}$$

Theorem 3.3.3 (Strong Subadditivity)**.** Let $\rho_{ABC} \in D(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ be a tripartite state. Then the entropy satisfies:

$$H(ABC)_\rho + H(B)_\rho \leq H(AB)_\rho + H(BC)_\rho. \tag{3.3.3}$$

Equivalently, the conditional quantum mutual information $I(A; C|B)_\rho \geq 0$.

Theorem 3.3.4 (Concavity of Entropy)**.** Let $\{p_i, \rho_i\}$ be an ensemble of density operators. Then the von Neumann entropy is concave:

$$H\left(\sum_i p_i \rho_i\right) \geq \sum_i p_i H(\rho_i). \tag{3.3.4}$$

Theorem 3.3.5 (Monotonicity of Quantum Relative Entropy)**.** Let $\rho, \sigma \in D(\mathcal{H})$, and let $\mathcal{N} : \mathcal{L}(\mathcal{H}) \longrightarrow \mathcal{L}(\mathcal{H}')$ be a quantum channel. Then:

$$D(\rho \| \sigma) \geq D(\mathcal{N}(\rho) \| \mathcal{N}(\sigma)). \tag{3.3.5}$$

Theorem 3.3.6 (Non-Negativity of Relative Entropy)**.** Let $\rho \in D(\mathcal{H})$, $\sigma \in \mathcal{L}(\mathcal{H})$ be positive semi-definite and $\text{Tr}[\rho] = 1$, $\text{Tr}[\sigma] \leq 1$. Then:

$$D(\rho \| \sigma) \geq 0, \tag{3.3.6}$$

with equality if and only if $\rho = \sigma$.

Theorem 3.3.7 (Quantum Data Processing Inequality)**.** Let ρ_{AB} be a bipartite state and let $\mathcal{N}$ be a quantum channel acting on system B. Then:

$$I(A; B)_\rho \geq I(A; B')_{(\mathbb{I} \otimes \mathcal{N})(\rho)}, \tag{3.3.7}$$

where B' is the output system of $\mathcal{N}$.

Chapter 4

Entanglement and Quantum Correlation

Quantum correlations are central to nearly all nonclassical phenomena observed in systems composed of two or more quantum subsystems. Both theoretical insights and experimental advancements in quantum information and computation have highlighted the crucial role these correlations play in various quantum information processing tasks. Among the different forms of quantum correlations, quantum entanglement [47],[32],[56] is the key resources in the field of quantum information science.

Applications such as quantum teleportation[9], dense coding[12], state merging[31], remote state preparation[11], quantum cryptography[24],[10], and quantum key distribution[7],[3],[16],[35] rely fundamentally on quantum correlations to function effectively.

To fully harness the potential of quantum correlations in such applications, it is essential to understand their nature and behaviour. This includes both a conceptual understanding of their characteristics and a rigorous examination of their quantitative measures. In recent years, substantial research has been directed toward identifying and formalising various ways to characterize and quantify quantum correlations. While the problem of quantifying correlations in pure bipartite systems is largely resolved, the situation becomes considerably more complex for multipartite systems, where even the analysis of pure states presents unresolved challenges. Furthermore, several quantification methods have also been developed for mixed states. This chapter addresses key issues related to the characterization and quantification of quantum correlations. It begins with an overview of quantum entanglement and its essential properties, followed by a review of several well-established entanglement measures found in the literature. The chapter ends with classical correlations and mutual information—two fundamental concepts for understanding quantum correlations.[23]

4.1 Quantum Entanglement

Quantum entanglement is the most extensively explored form of quantum correlation and has attracted significant attention from both theoretical and experimental communities. It has been pivotal in deepening our understanding of the fundamental principles of quantum mechanics, and in the realm of quantum information and computation, entanglement is regarded as a key resource. First introduced by Schrödinger and famously referred to as "spooky action at a distance" by Einstein, Podolsky, and Rosen, entanglement was initially viewed as a qualitative hallmark that starkly contrasted quantum theory with classical expectations.

The introduction of Bell's inequalities later provided a way to quantitatively distinguish this nonclassical feature, enabling its empirical investigation. Over the past few decades, entanglement has emerged as the most profound and operationally meaningful form of quantum correlation, playing a vital role in quantum technologies. As a result, substantial research efforts have been directed toward understanding, characterizing, and quantifying entanglement. These efforts will be discussed in detail in the upcoming sections, with particular emphasis on bipartite entanglement. The multipartite case will also be briefly outlined to provide broader context.

4.1.1 Bipartite Entanglement

Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be Hilbert spaces corresponding to two subsystems A and B, and let $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. A pure state $|\psi\rangle_{AB} \in \mathcal{H}_{AB}$ is said to be *entangled* if it cannot be written as a product state of the form

$$|\psi\rangle_{AB} = |\phi\rangle_A \otimes |\chi\rangle_B \,. \tag{4.1.1}$$

If such a decomposition is possible, the state is called *separable*. Any bipartite pure state $|\psi\rangle_{AB}$ admits a Schmidt decomposition:

$$|\psi\rangle_{AB} = \sum_{i=1}^{r} \sqrt{\alpha_i}\, |i\rangle_A \otimes |i\rangle_B \,, \tag{4.1.2}$$

where $r \leq \min(\dim(\mathcal{H}_A), \dim(\mathcal{H}_B))$, the coefficients $\alpha_i > 0$ satisfy $\sum_{i=1}^{r} \alpha_i = 1$, and $\{|i\rangle_A\}$ and $\{|i\rangle_B\}$ are orthonormal sets in $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. The values $\alpha_1, \ldots, \alpha_r$ are called the *Schmidt coefficients*, and r is referred to as the *Schmidt rank* of the state. A pure state is separable if and only if its Schmidt rank is one. Otherwise, it is entangled. For instance, in the case where $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = 2$ and the standard basis $\{|0\rangle, |1\rangle\}$ is used for both subsystems, a general pure state takes the form:

$$|\psi\rangle_{AB} = \sqrt{\alpha_1}\, |0\rangle_A \otimes |0\rangle_B + \sqrt{\alpha_2}\, |1\rangle_A \otimes |1\rangle_B \,, \tag{4.1.3}$$

where $\alpha_1, \alpha_2 \geq 0$ and $\alpha_1 + \alpha_2 = 1$. This state is entangled if both α_1 and α_2 are strictly positive. In practice,

quantum systems are often subject to environmental interactions, making it necessary to consider *mixed states* instead of pure ones. A mixed bipartite state is described by a density operator $\rho_{AB} \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Such a state is called *separable* if it can be written as a convex combination of product states:

$$\rho_{AB} = \sum_{i=1}^{k} p_i\, \rho_A^{(i)} \otimes \rho_B^{(i)}, \tag{4.1.4}$$

where $p_i \geq 0$, $\sum_{i=1}^{k} p_i = 1$, and $\rho_A^{(i)} \in \mathrm{D}(\mathcal{H}_A)$, $\rho_B^{(i)} \in \mathrm{D}(\mathcal{H}_B)$. If no such decomposition exists, the state is entangled.

It is essential to determine whether a given state is separable or entangled. For pure states, this is straightforward: compute the reduced density matrix $\rho_A = \mathrm{Tr}_B(\rho_{AB})$ (or $\rho_B = \mathrm{Tr}_A(\rho_{AB})$) and check its rank. A rank-one reduced state corresponds to a separable pure state; higher rank implies entanglement.

4.2 Properties of Entanglement

Entanglement is a fundamental feature of composite quantum systems and serves as a key resource in quantum information processing. This section outlines some important properties of entanglement, particularly in the context of bipartite systems.

4.2.1 Local Operations and Classical Communication (LOCC)

Entanglement cannot be increased under LOCC. This leads to the concept of entanglement as a non-increasing quantity under operations that involve only local quantum operations on subsystems, coordinated via classical communication. Formally, if ρ_{AB} is transformed into σ_{AB} by LOCC, then:

$$E(\rho_{AB}) \geq E(\sigma_{AB}), \tag{4.2.1}$$

for any valid entanglement measure E.

4.2.2 Separable States

A state $\rho_{AB} \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is said to be separable if it can be written as a convex combination of product states:

$$\rho_{AB} = \sum_{i} p_i\, \rho_A^{(i)} \otimes \rho_B^{(i)}, \tag{4.2.2}$$

where $p_i \geq 0$, $\sum_i p_i = 1$, and $\rho_A^{(i)}, \rho_B^{(i)}$ are density operators on $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Separable states do not exhibit entanglement.

4.2.3 Monogamy of Entanglement

Entanglement is a monogamous resource. If system A is maximally entangled with system B, it cannot be entangled with a third system C. This is quantified by the Coffman-Kundu-Wootters (CKW) inequality for three qubits:

$$\mathcal{C}^2_{A|BC} \geq \mathcal{C}^2_{AB} + \mathcal{C}^2_{AC}, \tag{4.2.3}$$

where $\mathcal{C}$ denotes the concurrence, a measure of bipartite entanglement.

4.2.4 Entanglement Measures

A valid entanglement measure $E(\rho_{AB})$ should satisfy the following properties:

- *Non-negativity:* $E(\rho_{AB}) \geq 0$, with equality if and only if ρ_{AB} is separable.
- *Monotonicity under LOCC:* E does not increase under LOCC.
- *Convexity:* E is a convex function, i.e., for any ensemble $\{p_i, \rho_{AB}^{(i)}\}$,

$$E\left(\sum_i p_i \rho_{AB}^{(i)}\right) \leq \sum_i p_i E(\rho_{AB}^{(i)}). \tag{4.2.4}$$

- *Normalization:* For maximally entangled states $|\Phi_d\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle_A |i\rangle_B$, $E(|\Phi_d\rangle) = \log d$.

4.2.5 Schmidt Decomposition and Entanglement of Pure States

Every pure state $|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ admits a Schmidt decomposition:

$$|\psi\rangle_{AB} = \sum_{i=1}^{r} \sqrt{\alpha_i}\, |i\rangle_A \otimes |i\rangle_B\,, \tag{4.2.5}$$

where $\{\alpha_i\}$ are the Schmidt coefficients. The state is separable if and only if the Schmidt rank $r = 1$. For pure states, the entropy of entanglement, defined as the von Neumann entropy of the reduced state, is a valid entanglement measure:

$$E(|\psi\rangle_{AB}) = S(\rho_A) = -\mathrm{Tr}(\rho_A \log \rho_A), \tag{4.2.6}$$

where $\rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|_{AB})$.

4.2.6 Mixed-State Entanglement and the PPT Criterion

For mixed states, determining entanglement is generally hard. A widely used necessary condition for separability is the *Positive Partial Transpose (PPT)* criterion:

$$\rho_{AB}^{T_B} \geq 0 \quad \Rightarrow \quad \rho_{AB} \text{ is separable (for } 2 \otimes 2 \text{ or } 2 \otimes 3\text{)}. \tag{4.2.7}$$

Here, T_B denotes the partial transpose with respect to subsystem B.

4.2.7 Distillability and Bound Entanglement

An entangled mixed state ρ_{AB} is said to be *distillable* if, by LOCC, one can extract a maximally entangled state from many copies of ρ_{AB}. Some entangled states that are not distillable are known as *bound entangled* states. All bound entangled states are necessarily PPT.

4.2.8 Activation and Superadditivity

Entanglement may exhibit superadditive behavior: two copies of a state may exhibit more entanglement together than the sum of entanglement in each individually. Also, entanglement can be activated: a bound entangled state combined with another entangled state may allow distillation.

4.3 Bell Nonlocality

Bell's theorem was proposed by John Stewart Bell. He proved that quantum theory is incompatible with local hidden variable theory and dispelled the Einstein–Podolsky–Rosens (EPR) paradox. Bell theorem looks something like this: "No physical theory based on local hidden variables can reproduce all the predictions of quantum mechanics". The kernel of that and all subsequent inequalities is founded on this suggestion regarding local realism: Local realism denies the existence of measurable properties when not measured, and disallows any faster-than-light travel of information (no signaling theorem). On the other hand, quantum mechanics predicts and experimentally confirms the violation of such inequalities, which would involve abandonment of either or both locality and realism. Following the special theory of relativity, no signal can travel faster than light therefore quantum mechanics vividly violates the reality that physical properties exist independent of measurement and measurement only reveals them. The first crude experiment specifically designed to test the Bell theorem was carried out by John Clauser and Stuart Freedman. In this case, it was in 1969 when John Clauser, Michael Horne, Abner Shimony, and Richard Holt referred to a widely common test, the CHSH test, to show the validity of Bell theorem that is Quantum mechanics is incompatible with local hidden variable theory.

Mathematical formulation of the CHSH Inequality

Consider two spatially separated observers, Alice and Bob, each measuring one of two possible observables on a shared entangled state. Let the measurement settings be represented by binary-valued observables A_1, A_2 for Alice and B_1, B_2 for Bob, with eigenvalues ± 1. A local hidden variable theory assumes pre-existing values determined by some hidden variable λ, such that the measurement outcomes satisfy deterministic functions $A_i(\lambda), B_j(\lambda) \in \{-1, 1\}, \quad i, j \in \{1, 2\}$. The CHSH operator is defined as

$$S = \langle A_1 B_1 \rangle + \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle - \langle A_2 B_2 \rangle \tag{4.3.1}$$

where $\langle A_i B_j \rangle$ corresponds to the expectation values of A_i and B_j. In a local hidden variable theory, the correlation function can be written as an average over the hidden variable distribution $\rho(\lambda)$:

$$\langle A_i B_j \rangle = \int d\lambda \rho(\lambda) A_i(\lambda) B_j(\lambda). \tag{4.3.2}$$

Using algebraic manipulations, we can show that for any deterministic assignment $S \leq 2$. This constitutes the Bell-CHSH inequality.In order to show that quantum mechanics is incompatible with local hidden variable theory,we show the violation of the CHSH inequality we consider an entangled Bell state:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{4.3.3}$$

Let us define the observables simply as pauli observables and linear combinations of the same $A_1 = \sigma_x, \quad A_2 = \sigma_z, \quad B_1 = \frac{1}{\sqrt{2}}(\sigma_x + \sigma_z), \quad B_2 = \frac{1}{\sqrt{2}}(\sigma_x - \sigma_z)$. The effect of the observables on the given state can be evaluated using trace norm formulation

$$\langle A_X B_Y \rangle = \mathrm{Tr}\left(|\psi\rangle\langle\psi|, (A_X \otimes B_Y)\right) \tag{4.3.4}$$

Evaluating everything and putting back to the equation we have

$$S = 2\sqrt{2} \approx 2.828. \tag{4.3.5}$$

The expectation value of CHSH operator $2\sqrt{2} > 2$ suggest that quantum mechanics rules out local hidden variable models.

4.4 Classical and Quantum Correlations

1. *Classical Correlations:* These are correlations that can be fully described by separable states and admit a local hidden variable (LHV) model. Consequently, they obey all Bell-type inequalities.

   (a) Mathematical Definition: A bipartite quantum state $\rho_{AB} \in \mathrm{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is classically correlated if it is separable, i.e.,

   $$\rho_{AB} = \sum_i p_i \, \rho_i^A \otimes \rho_i^B, \tag{4.4.1}$$

   where each ρ_i^A, ρ_i^B is a valid density operator on $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively, $p_i \geq 0$, and $\sum_i p_i = 1$.

   (b) Operational Characterization: For all sets of local observables $\{A_x\}$ and $\{B_y\}$, the correlations satisfy the Bell inequalities. For instance, the CHSH inequality holds:

   $$|\langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle| \leq 2, \tag{4.4.2}$$

   where $\langle A_x B_y \rangle = \mathrm{Tr}\left(\rho_{AB}(A_x \otimes B_y)\right)$.

   (c) Classical Joint Distribution: The joint probability distribution for outcomes a, b given measurement settings x, y factorizes as:

   $$P(a, b|x, y) = \sum_i p_i \, P_i(a|x) \, P_i(b|y), \tag{4.4.3}$$

   indicating the existence of a local hidden variable model with shared classical randomness.

2. *Quantum Correlations:* These are correlations that cannot be described by separable states and violate Bell inequalities. They arise due to entanglement and demonstrate quantum nonlocality.

   (a) Mathematical Definition: A bipartite state ρ_{AB} exhibits quantum correlations if it is entangled, i.e., it cannot be written in the separable form:

   $$\rho_{AB} \neq \sum_i p_i \, \rho_i^A \otimes \rho_i^B. \tag{4.4.4}$$

   (b) Operational Characterization: There exist local measurements $\{A_x\}$ and $\{B_y\}$ for which the state ρ_{AB} violates a Bell inequality. For example, the Bell state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ achieves:

   $$\text{CHSH value} = 2\sqrt{2} > 2, \tag{4.4.5}$$

   for appropriate Pauli observables:

   $$A_0 = \sigma_z, \quad A_1 = \sigma_x, \quad B_0 = \frac{\sigma_z + \sigma_x}{\sqrt{2}}, \quad B_1 = \frac{\sigma_z - \sigma_x}{\sqrt{2}}. \tag{4.4.6}$$

   (c) Nonclassical Joint Distribution: For entangled states, the joint probability distribution is not factorizable as in the classical case:

   $$P(a, b|x, y) = \mathrm{Tr}(\rho_{AB} \, A_x \otimes B_y) \neq \sum_i p_i \, P_i(a|x) \, P_i(b|y). \tag{4.4.7}$$

   In general, quantum correlations can involve interference effects due to non-commutativity of observables.

   (d) Total Correlations (Mutual Information): The total correlations—both classical and quantum—in a bipartite state ρ_{AB} are captured by the mutual information:

   $$I(\rho_{AB}) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}), \tag{4.4.8}$$

   where $S(\rho) = -\mathrm{Tr}(\rho \log \rho)$ is the von Neumann entropy. For a maximally entangled two-qubit state, $I(\rho_{AB}) = 2 \log 2$.

(e) Quantum Discord: Quantum discord quantifies the quantum part of the correlations that go beyond classical correlations. For a bipartite state ρ_{AB}, it is defined as:

$$D_A(\rho_{AB}) = I(\rho_{AB}) - \max_{\{B_y\}} J(\rho_{AB}|\{B_y\}), \tag{4.4.9}$$

where $J(\rho_{AB}|\{B_y\})$ is the classical mutual information obtained by performing a complete measurement $\{B_y\}$ on subsystem B. Discord may be nonzero even for some separable states.

Chapter 5

Device Independent Quantum Key Distribution

5.1 Ekert 91 Protocol

The genesis of DIQKD can be traced back to Ekert's protocol[24] where he had used any of the four Bell states and distributed them among Alice and Bob. They locally chose a set of three observables, viz. $\{A_i\}$ and $\{B_i\}$, where

$$\begin{aligned} A_i &= \cos\phi_i^A \sigma_z + \sin\phi_i^A \sigma_x, \\ B_i &= \cos\phi_i^B \sigma_z + \sin\phi_i^B \sigma_x, \end{aligned} \tag{5.1.1}$$

And the values of the corresponding angles are specified as $\{\phi_1^A = 0, \quad \phi_2^A = \frac{\pi}{2}, \quad \phi_3^A = \frac{\pi}{4}\}$ (for Alice) and $\{\phi_1^B = 0, \quad \phi_2^B = -\frac{\pi}{4}, \quad \phi_3^B = \frac{\pi}{4}\}$ (for Bob)

They announce their chosen observables for each pair before the key generation and CHSH security test. Alice and Bob use a portion of all nine combinations of observables for key generation and some portion for the CHSH Inequality check based on the Bell theorem.

In particular, they choose the results of the action of (A_1, B_1) and (A_2, B_2) on their corresponding singlets for key generation, where they get completely correlated or anticorrelated depending on the Bell pair.

The CHSH test is performed on the results of the action of the observables. $(A_1, B_3), (A_1, B_2), (A_2, B_3)$, and (A_2, B_2)on the corresponding singlets using the equation

$$S = \langle A_1 B_1 \rangle + \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle - \langle A_2 B_2 \rangle \tag{5.1.2}$$

where $\langle A_i B_j \rangle$ corresponds to the expectation values of A_i and B_j.
Classically, any observables can take values ±1, so from the above equation, the upper bound for the CHSH inequality is 2. For quantum observables, the expectation value is given by

$$\langle A_i B_j \rangle = \mathrm{Tr}(A_i \otimes B_j \, \rho) \tag{5.1.3}$$

And thus the value of the CHSH inequality can violate the classical upper bound of 2. For a maximally entangled bipartite state, it is upper-bounded by $2\sqrt{2}$.

The violation of the CHSH inequality is necessary and sufficient to verify that Alice and Bob are sharing an entangled state, rather than any classically correlated state.

5.2 DI QKD against collective attack

The first work on Device Independent Quantum Key Distribution was done by "Aci ́n et al." [2]. The protocol being proposed is a modified version of the *Ekert 1992 QKD protocol*[24]. It is designed to ensure

security in a device-independent manner, meaning it does not rely on specific assumptions about the devices used by Alice and Bob. Traditional QKD protocols, like BB84, assume that the devices used by Alice and Bob are perfectly controlled and behave as expected. However, this assumption is not always realistic. Device-Independent QKD removes these assumptions and guarantees security based solely on the laws of quantum mechanics and the observed correlations in the measurement outcomes.
Essential requirements for a QKD system are independence and privacy of the measurement settings and privacy of the measurement outcomes. In addition to these two essential requirements, a dependent QKD system also holds a third requirement, which is being eliminated by DIQKD, perfect control over the state preparation and measurement devices.
The main problem addressed in this paper is to prove the security of a QKD protocol against "collective attacks" by an eavesdropper, Eve, in a device-independent scenario. A collective attack is where Eve applies the same strategy to each particle transmitted between Alice and Bob, but can perform a coherent measurement on all the quantum states she holds at a later time. The protocol being proposed is based on a source that distributes an entangled particle between Alice and Bob.

5.2.1 Proposed Protocol

Alice and Bob share a quantum channel through which a source (potentially controlled by Eve) distributes entangled pairs of particles to both parties.

Measurement Settings: Alice can choose from three measurement settings: A_0, A_1, A_2. Bob can choose from two measurement settings: B_1, B_2. All measurements yield binary outcomes, $a_i, b_j \in \{-1, +1\}$.

Steps of the Protocol: A source generates pairs of entangled particles and sends them to Alice and Bob. For each pair of particles, Alice randomly selects one of her measurement settings (A_0, A_1, A_2).Bob randomly selects one of his measurement settings (B_1, B_2). Both record the outcomes of their measurements. The raw key bits are extracted from the measurements with settings A_0 and B_1.
Alice and Bob publicly announce a subset of their measurement settings and outcomes to estimate the quantum bit error rate (QBER), $Q = \Pr(a_0 \neq b_1)$. Alice and Bob compute the CHSH polynomial using the outcomes of settings A_1, A_2, B_1, B_2,

$$S = \langle A_1 B_1 \rangle + \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle - \langle A_2 B_2 \rangle. \tag{5.2.1}$$

A violation of the CHSH inequality (i.e., $S > 2$) indicates nonlocal quantum correlations, certifying that Eve cannot have full knowledge of the key.

Post Processing: If the CHSH violation S and QBER Q are within acceptable bounds: *Error Correction:* Alice and Bob perform error correction to align their raw keys. *Privacy Amplification:* They reduce Eve's information about the final key using privacy amplification techniques.After error correction and privacy amplification, Alice and Bob obtain a final shared secret key that is secure against any eavesdropping attempts, assuming the correctness of quantum mechanics and no leakage of unwanted information.

5.2.2 Key Result

The authors derive a tight bound on the Holevo information, which quantifies the maximum amount of information Eve can obtain about the secret key as a function of the violation of a Bell-type inequality (specifically, the CHSH inequality).

$$\chi(B1 : E) \leq h\left(\frac{1 + \sqrt{(S/2)^2 - 1}}{2}\right), \tag{5.2.2}$$

where: $\chi(B1 : E)$ is the Holevo information between one of the legitimate parties (Bob) and Eve. S represents the violation of the CHSH inequality. $h(x) = -x \log_2(x) - (1 - x) \log_2(1 - x)$ is the binary entropy function.

5.2.3 Key Rate and Eve's Uncertainity

The key rate, r, quantifies the number of secure key bits that can be distilled per communication attempt. It is lower-bounded by:

$$r \geq I(A_0 : B_1) - \chi(B_1 : E). \tag{5.2.3}$$

For correlations satisfying $S = 2\sqrt{2}(1 - 2Q)$, the key rate is computed as a function of the QBER, Q. Even with imperfect devices, a positive key rate can be achieved if the CHSH violation is sufficiently large.

5.3 DI QKD using Random Key Basis

Device Independent Quantum Key Distribution using a random key basis was done by [48] and is being inspired from the work of [2]. Instead of using only a single pair of measurement settings here the key generation basis is randomly chosen before each round.

5.3.1 Proposed Protocol

The protocol consists of five main stages: measurements, sifting, parameter estimation, one-way error correction and verification, and privacy amplification. The measurement stage is executed over an asymptotically large number of rounds, denoted by N. In each round, Alice and Bob independently select their measurement settings $X \in \{0, 1\}$ and $Y \in \{0, 1, 2, 3\}$, respectively, according to the following probability distributions:

$$P(X = 0) = p, \quad P(X = 1) = 1 - p, \quad P(Y = 0) = qp, \quad P(Y = 1) = q(1 - p), \quad P(Y = 2) = P(Y = 3) = \frac{1 - q}{2}, \tag{5.3.1}$$

where $0 \leq p, q \leq 1$. After selecting their settings, Alice and Bob perform measurements with their respective devices and obtain outcomes $A_X \in \{0, 1\}$ and $B_Y \in \{0, 1\}$.

In the second stage, they announce their chosen settings over an authenticated classical communication channel. They retain only those rounds in which $Y \in \{0, 1\}$ and $X = Y$. The corresponding outcomes are used to generate a raw key of approximate size $\sim q(p^2 + (1 - p)^2)N$. The data associated with the settings $X \in \{0, 1\}$ and $Y \in \{2, 3\}$ is reserved for parameter estimation, yielding a dataset of size $\sim (1 - q)N$. All other data are discarded.

The third stage is devoted to estimating the CHSH parameter. To achieve this, Alice and Bob reveal their measurement outcomes from the parameter estimation dataset. The CHSH value is given by

$$S = \max\{0, C_{12} - C_{02} - C_{03} - C_{13}\}, \tag{5.3.2}$$

where the correlation terms C_{XY} are defined as

$$C_{XY} = P(A_X = B_Y \mid X, Y) - P(A_X \neq B_Y \mid X, Y), \tag{5.3.3}$$

quantifying the strength of correlations between the outcomes A_X and B_Y. This step is crucial in determining whether to proceed with the protocol. If the condition $S > S_{tol}$ is met, where S_{tol} is a predetermined threshold for the violation of the CHSH inequality, Alice and Bob continue with the subsequent stages. Otherwise, the protocol is aborted. A violation exceeding S_{tol} ensures that the raw keys are sufficiently secure.

Upon successful parameter estimation, Alice and Bob proceed with one-way error correction and verification to reconcile discrepancies in their raw keys. Finally, privacy amplification is applied to produce a shared secret key that is secure against any adversary.

5.3.2 Key Result

5.3.3 Asymptotic Key Rate and Eve's Uncertainty

The author derives the asymptotic secret key rate as

$$K_\infty = p_s r_\infty \tag{5.3.4}$$

where $p_s := p^2 + (1-p)^2$ denotes the probability that Alice and Bob select matching key-generation bases in the limit $q \to 1$. This quantity represents the ratio of the length of the extractable secret key to the total number of measurement rounds N, which is taken to approach infinity, i.e., $N \to \infty$. The second factor in Eq. (1), denoted by r_∞, is known as the *secret fraction* [48], and it can be expressed in terms of entropic functions as

$$\begin{aligned} r_\infty &:= \lambda H(A_0|E) + (1-\lambda)H(A_1|E) \\ &\quad -\lambda h(Q_{A_0B_0}) - (1-\lambda)h(Q_{A_1B_1}) \end{aligned} \tag{5.3.5}$$

where $\lambda := \frac{p^2}{p_s}$, and $h(x) := -x\log(x) - (1-x)\log(1-x)$ denotes the binary entropy function. The argument of the binary entropy function, $Q_{A_XB_Y} := P(A_X \neq B_Y|X,Y)$, quantifies the quantum bit error rate (QBER) for measurement choices X, Y. The variable E refers to the quantum side information accessible to Eve immediately before the error correction phase. The first line of above equation comprises the conditional von Neumann entropy terms, which measure the degree of uncertainty Eve has about Alice's outcomes conditioned on her side information. In contrast, the second line accounts for the potential information leakage to Eve during the error correction stage through her ability to exploit correlations. If we define

$$U = \lambda H(A_0|E) + (1-\lambda)H(A_1|E), \tag{5.3.6}$$

then Alice and Bob can achieve a positive secret key rate, provided they are able to establish a reliable lower bound on U solely from the observed CHSH violation. While this task is non-trivial, it has been shown that using a family of device-independent entropic uncertainty relations [13][21], one can guarantee [48]

$$\lambda H(A_0|E) + (1-\lambda)H(A_1|E) \geq C^\star(S) \tag{5.3.7}$$

where $C^\star$ is a function dependent on the observed CHSH value S. Using U as a figure of merit, it has been demonstrated that Eve's uncertainty in this protocol exceeds that in the original protocol for all $S \in (2, 2\sqrt{2}]$ [48]. Notably, when $\lambda = \frac{1}{2}$, the lower bound on U nearly saturates its optimal value, representing the fundamental upper bound on Eve's uncertainty [48]. The objective of this work is to reduce the computational complexity involved in bounding Eve's uncertainty without degrading the achievable secret key rate.

5.4 DI QKD using Noisy Preprocessing

The experimental demonstration of Device Independent Quantum Key Distribution poses a serious challenge. The photonic realisation comes with a serious bottleneck that is detection efficiency η. Detection efficiency refer to the probability that the signal sent over the quantum channel are being successfully received. The work by "Ho et al." [28]. In this work a new variant of DIQKD protocol had been proposed to significantly relaxed the detection efficiency without comprising the security. An artificial noise had been added to the to the raw key followed by a error correction step to finally get the secure key. The artificial noise is local and thus can't be controlled by the adversary.

5.4.1 Proposed Protocol

A source generates pairs of entangled photons and distributes them such that one photon from each pair is sent to Alice and the other to Bob. Alice randomly selects a measurement setting $x \in \{0, 1, 2\}$. She uses $x = 0$ Used for key generation and the rest for CHSH test. The Bob measurement settings $y \in \{1, 2\}$ where he uses both for key generation as well as CHSH test. The measurement outcome for alice and bob are denoted as A_x and B_y respectively for chosen setting x and y. $A_x, B_y \in \{-1, +1\}$. Alice uses the outcome from the measurement setting $x = 0$ for key generation. This setting is chosen such that it is in the *key generation basis*, which is aligned to minimize the conditional entropy $H(B_1'|A_0)$. Similarly bob uses the outcome from the measurement setting $y = 1$, which is chosen to maximize correlation with Alice's outcome when she uses $x = 0$.

Alice and Bob publicly exchange a subset of their measurement results (particularly A_1, A_2, B_1 and B_2) to estimate the CHSH score S:

$$S = \langle A_1B_1\rangle + \langle A_1B_2\rangle + \langle A_2B_1\rangle - \langle A_2B_2\rangle, \tag{5.4.1}$$

where:

$$\langle A_x B_y \rangle = p(A_x = B_y \,|\, x, y) - p(A_x \neq B_y \,|\, x, y). \tag{5.4.2}$$

The CHSH score determines the level of quantum entanglement and thus the security of the protocol. The next crucial step is noisy preprocessing. It's followed by Bob applies noisy preprocessing by flipping each of his raw key bits B_1 with a certain probability p. This step is crucial for reducing Eve's information about the key. Alice and Bob then perform an error correction protocol to reconcile their keys. Alice learns Bob's new (noisy) raw key B'_1. Essentially bob only shares the index where he performed the bit flip. The probability p is supposed to be already shared between alice and bob before the start of the protocol. Alice and Bob apply a hash function to their reconciled key to distill the final secure key, removing any information Eve may have.

5.4.2 Key Result

The main key result of this work is in lowering the detection efficiency threshold to 83.2% from 92.7% in [2].

5.4.3 Key rate

The secret key generation rate r is a function of the entropy terms:

$$r = H(B'_1|E) - H(B'_1|A_0) \tag{5.4.3}$$

where: $H(B'_1|E)$ is the von Neumann entropy representing the uncertainty Eve has about Bob's noisy key bits B'_1. $H(B'_1|A_0)$ represents the amount of uncertainty remaining after Alice's measurement results A_0 are known to Bob. The expanded expression for $H(B'_1|E)$ can be rewritten as:

$$H(B'_1|E) = H(B'_1) - \Delta, \tag{5.4.4}$$

where:

$$\Delta = \left(H(\rho_E) - \sum_b p_b H(\rho_{E|b}) \right). \tag{5.4.5}$$

Interpretation of Δ:

- Δ represents the reduction in Eve's uncertainty due to her knowledge of Bob's measurement outcome b.
- $H(\rho_E)$ is the total entropy of Eve's state, without any knowledge of Bob's key bits.
- $\sum_b p_b H(\rho_{E|b})$ is the weighted sum of Eve's entropy over different possible outcomes b on Bob's side. This term averages the entropy of Eve's state given each possible outcome of Bob's key bits.
- The difference $H(\rho_E) - \sum_b p_b H(\rho_{E|b})$ quantifies the amount of information Eve gains about Bob's key after observing his measurement outcomes.

The key rate r is lower-bounded by:

$$r \geq 1 - I_p(S) - H(B'_1|A_0) \tag{5.4.6}$$

where $I_p(S)$ is defined as:

$$I_p(S) = h\left(\frac{1 + \sqrt{(S/2)^2 - 1}}{2}\right) - h\left(\frac{1 + \sqrt{1 - p(1-p)(8 - S^2)}}{2}\right) \tag{5.4.7}$$

and $h(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the binary entropy function. $I_p(S)$ captures the effect of noisy preprocessing on Eve's knowledge about the key.
The term $I_p(S)$ is composed of two parts. The first term

$$h\left(\frac{1+\sqrt{(S/2)^2-1}}{2}\right) \tag{5.4.8}$$

This term represents the maximal amount of information Eve could have about the key in the absence of noisy preprocessing, considering the quantum correlations represented by S. The expression inside the entropy function:

$$\frac{1+\sqrt{(S/2)^2-1}}{2} \tag{5.4.9}$$

relates to the probability distribution of the outcomes that Eve might have access to. The square root term comes from the use of the Bell parameter or a similar measure of correlation.
The second term:

$$h\left(\frac{1+\sqrt{1-p(1-p)(8-S^2)}}{2}\right) \tag{5.4.10}$$

This term represents the reduced amount of information Eve has about the key after the noisy preprocessing characterized by parameter p. The term:

$$\frac{1+\sqrt{1-p(1-p)(8-S^2)}}{2} \tag{5.4.11}$$

shows how the noisy preprocessing affects the correlations that Eve can exploit. The factor $p(1-p)(8-S^2)$ quantifies the impact of noise on Eve's accessible information.

5.5 DI QKD using Random Postselection

The practical realisation of device-independent quantum key distribution is quite challenging due to low noise tolerance. The practical realisation requires a loophole-free bell test [27][46]. The critical bottleneck remains the same as in any photonic realisation that's restricted by detection efficiency [58]. Restricted detection efficiency refers to the inability of the emitted photons to be detected due to a loss in either transmission or the detector or both.[58] Detection efficiency is directly related to the Bell violation and to the key rate. Photonic realisation of loophole-free Bell test [20][49][26] has achieved $\eta \approx 80\%$. Conventional security proof of DIQKD requires $\eta > 90\%$ [42],[37],[44],[54],[5]. The required detection efficiency is far beyond the current technology. The work of [58] proposed an approach to lower this threshold by considering the key from a post selected smaller string instead of the entire data.

5.5.1 Proposed protocol

The two legitimate users, Alice and Bob, use a source to generate entangled photon pairs that are shared. The measurement settings are $x \in \{1,2\}$ and $y \in \{1,2,3\}$ for Alice and Bob respectively. The measurement outcomes are a, b for Alice and Bob, respectively where $a, b \in \{0,1\}$. The key is being generated from setting $x = 1$ and $y = 3$. CHSH test is done from all the rest settings that are $x \in \{1,2\}$ and $y \in \{1,2,3\}$. Now, for the imperfect detectors, there are four possibilities to be observed: none of the detectors is being clicked, one of the detectors is being clicked, and finally, both of them are being clicked. Outcome 0 is assigned only when the first detector is clicked for both Alice and Bob, and 1 is assigned for the remaining three cases. The following table summarises this The post-selection event is done on the outcome corresponding to the setting $x = 1$ and $y = 3$. Alice and Bob will keep all the 0 outcomes corresponding to the first detector click, as it will contain genuine quantum correlation in principle. The event where both detectors click or don't click contain little correlation, there outcome 1 will be randomly and independently be retain and discarded with probability p and $(1-p)$ respectively. The probability p is being local and predetermined by Alice and Bob, Eve has no control over this.

Table 5.1: Recorded outcome corresponding to imperfect detector

Detector 1	Detector 2	Recorded outcome
NO	NO	1
YES	NO	0
NO	YES	1
YES	YES	1

5.5.2 Key Result

The main key result of this work is in lowering the detection efficiency threshold to 68.5% from 82.6% in [28] and 92.7% in [2].

5.5.3 Key Rate

Let $\mathcal{H}_A$, $\mathcal{H}_B$ and $\mathcal{H}_E$ be the Hilbert space associated with the alice's, bob's and eve's devices respectively and ρ_{ABE} be the state shared tripartite state. Then joint probability distribution after the measurement with respect to the settings[28] will be

$$P(a,b|x,y) = \mathrm{Tr}[(A_{a|x} \otimes B_{b|y} \otimes \mathbb{I})_{\rho_{ABE}}] \tag{5.5.1}$$

where $(A_{a|x}, B_{b|y})$ are the corresponding POVMs. Now let ν_p be set of post selected event, $\nu_p = \{ab|ab = 00, 01, 10, 11\}$ and let ω_{ab} be the weight of each correspond post selected event. In the work [58] author has chosen the weight as $\omega_{00} = 1$, $\omega_{01} = \omega_{10} = p$ and $\omega_{11} = p^2$. The probability that a bit pair in ν will be kept is defined as

$$p\nu_p = \sum_{ab \in \nu} \omega_{ab} \cdot P(a,b|x,y) \tag{5.5.2}$$

The probability distribution of post selected event is defined as[58],

$$\hat{P}(a,b|x,y,\nu_p) = P(a,b|x,y) \cdot \frac{\omega_{ab}}{p\nu_p} \tag{5.5.3}$$

The asymptotic key rate r with one way optimal error correction is shown to be

$$r \geq p_{\nu_p} \left[H_{\min}(A_x|E,\nu_p) - H(A_x|B_y,\nu_p) \right], \tag{5.5.4}$$

where $H_{\min}(A_x|E,\nu_p)$ is the conditional min-entropy and $H(A_x|B_y,V_p)$ is the one-way error correction cost.

Table 5.2: Comparison Table

Protocol	Measurement Settings	Key Generating	CHSH Test	Key Rate	Detection Efficiency Threshold
Acín et al.	Alice: A_0, A_1, A_2, Bob: B_1, B_2	A_0, B_1	A_1, A_2, B_1, B_2	$r \geq I(A_0 : B_1) - \chi(B_1 : E)$	92.7%
Random Key Basis	Alice: A_0, A_1, Bob: B_0, B_1, B_2, B_3	A_0, A_1, B_0, B_1	A_0, A_1, B_2, B_3	$K_\infty = p_s r_\infty$	–
Noisy Preprocessing	Alice: A_0, A_1, A_2, Bob: B_1, B_2	A_0, B_1	A_1, A_2, B_1, B_2	$r \geq 1 - I_p(S) - H(B_1' \mid A_0)$	83.2%
Random Post Selection	Alice: A_1, A_2, Bob: B_1, B_2, B_3	A_1, B_3	A_1, A_2, B_1, B_2, B_3	$r \geq p_{\nu_p} \left[H_{\min}(A_x \mid E, \nu_p) - H(A_x \mid B_y, \nu_p)\right]$	68.5%

Chapter 6

Survey of Security Proofs in Quantum Key Distribution

Quantum Key Distribution (QKD) stands as a groundbreaking paradigm in cryptography, offering an unprecedented level of security for communication that is rooted in the very fabric of quantum mechanics. Unlike classical cryptographic methods, whose security hinges on the computational difficulty of certain mathematical problems—making them vulnerable to future advancements in computing, including the advent of quantum computers—QKD's robustness derives from fundamental physical laws. Specifically, its security is guaranteed by principles such as the "no-cloning theorem," which postulates that an unknown quantum state cannot be perfectly replicated. This inherent property means that any attempt by an unauthorized third party to intercept or measure the quantum bits (qubits) used for key exchange will inevitably alter their state, thereby immediately revealing the presence of an eavesdropper. This immediate detection mechanism forms the cornerstone of QKD's promise: if an intrusion is detected, the legitimate users can simply abort the key generation process, ensuring that no information is compromised.

6.1 Renner's Security Definition

Let $\rho_{KE} \in \mathcal{S}(KE)$ be the joint quantum state of the final key K and the adversary's quantum system E. A QKD protocol is called *ε-secure* if it satisfies the trace distance criterion[45]:

$$\frac{1}{2} \|\rho_{KE} - \tau_K \otimes \rho_E\|_1 \leq \varepsilon, \tag{6.1.1}$$

where:

- $\|\cdot\|_1$ denotes the trace norm,
- $\tau_K = \frac{1}{|K|} \sum_k |k\rangle\langle k|$ is the fully mixed (ideal) key state,
- $\rho_E = \mathrm{Tr}_K[\rho_{KE}]$ is the reduced state of Eve.

This implies that the real state is ε-close to the ideal key—uniform and independent of the adversary. The parameter ε captures the maximum failure probability and ensures composable security.

6.2 Tools for Security Analysis

6.2.1 Distance Measures Between Quantum States

Trace Distance: For quantum states $\rho, \sigma \in \mathcal{S}(A)$,

$$\Delta(\rho, \sigma) := \frac{1}{2} \|\rho - \sigma\|_1 = \sup_{0 \leq P \leq \mathbb{I}} \mathrm{Tr}[P(\rho - \sigma)]. \tag{6.2.1}$$

Operationally, if $\Delta(\rho, \sigma) = \epsilon$, the maximum probability of distinguishing them is $\frac{1}{2}(1+\epsilon)$.

Generalized for Subnormalized States: For $\hat{\rho}, \hat{\sigma}$ with Tr ≤ 1,

$$\Delta(\hat{\rho}, \hat{\sigma}) = \frac{1}{2}\|\hat{\rho} - \hat{\sigma}\|_1 + \frac{1}{2}|\mathrm{Tr}(\hat{\rho} - \hat{\sigma})|. \tag{6.2.2}$$

Purified Distance: For subnormalized states,

$$D_{\mathrm{P}}(\rho, \sigma) := \sqrt{1 - F(\rho, \sigma)}, \tag{6.2.3}$$

where the generalized fidelity is

$$F(\rho, \sigma) := \left(\mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}} + \sqrt{(1 - \mathrm{Tr}\rho)(1 - \mathrm{Tr}\sigma)}\right)^2. \tag{6.2.4}$$

Relation:

$$\|\rho - \sigma\|_1 \leq 2D_{\mathrm{P}}(\rho, \sigma)^2 \leq 2\|\rho - \sigma\|_1. \tag{6.2.5}$$

Properties: Both trace and purified distances satisfy:

- Non-negativity and identity of indiscernibles,
- Symmetry,
- Triangle inequality,
- Monotonicity under trace-non-increasing CP maps $\mathcal{M}$:

$$\Delta(\mathcal{M}(\rho), \mathcal{M}(\sigma)) \leq \Delta(\rho, \sigma). \tag{6.2.6}$$

6.2.2 Classical-Quantum (cq) States

A cq-state takes the form:

$$\rho_{AE} = \sum_x p(x)\, |x\rangle\langle x|_A \otimes \rho_{E|x}, \tag{6.2.7}$$

where $\{|x\rangle\}$ is an orthonormal basis for classical system A, and $\rho_{E|x}$ are conditional quantum states.

6.2.3 Entropies

Shannon and von Neumann Entropies

Shannon Entropy Let X be a discrete random variable taking values in a finite set $\mathcal{X}$, with probability distribution $\{p(x)\}_{x\in\mathcal{X}}$. The *Shannon entropy* of X is defined as

$$H(X) = -\sum_{x\in\mathcal{X}} p(x)\log p(x), \tag{6.2.8}$$

where the logarithm is usually taken in base 2, and the unit is *bits*.

- $H(X) \geq 0$, with equality if and only if X is deterministic (i.e., $p(x) = 1$ for some x).
- The maximum entropy is $\log|\mathcal{X}|$, achieved when $p(x) = 1/|\mathcal{X}|$ for all $x \in \mathcal{X}$ (uniform distribution).

von Neumann Entropy Let $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ be a density operator on a finite-dimensional Hilbert space $\mathcal{H}_A$. The *von Neumann entropy* of the quantum system A in state ρ_A is defined as

$$S(A)_\rho = -\text{Tr}[\rho_A \log \rho_A]. \tag{6.2.9}$$

This definition is analogous to the Shannon entropy, and in fact, if ρ_A has spectral decomposition

$$\rho_A = \sum_i \lambda_i |i\rangle\langle i|, \tag{6.2.10}$$

then the von Neumann entropy reduces to

$$S(\rho_A) = -\sum_i \lambda_i \log \lambda_i = H(\{\lambda_i\}), \tag{6.2.11}$$

i.e., the Shannon entropy of its eigenvalue spectrum.

- $S(\rho_A) = 0$ if and only if ρ_A is a pure state.
- $S(\rho_A) \leq \log d$, where $d = \dim \mathcal{H}_A$, with equality if $\rho_A = \frac{I}{d}$ is the maximally mixed state.

Classical Conditional Entropy. Given two classical random variables X and Y with joint distribution $p(x,y)$, the conditional Shannon entropy of X given Y is:

$$H(X|Y) := -\sum_{x,y} p(x,y) \log p(x|y) = H(X,Y) - H(Y). \tag{6.2.12}$$

It quantifies the uncertainty remaining about X when Y is known.

Quantum Conditional Entropy. Let $\rho_{AE} \in \mathcal{S}(AE)$ be a bipartite quantum state. The (von Neumann) entropy of system A conditioned on E is defined as:

$$S(A|E)_\rho := S(AE)_\rho - S(E)_\rho, \tag{6.2.13}$$

where $S(X)_\rho = -\text{Tr}[\rho_X \log \rho_X]$ is the von Neumann entropy.

Properties of Quantum Conditional Entropy. Let $\rho_{AB}, \rho_{ABX} \in \mathcal{S}(ABX)$. The conditional von Neumann entropy satisfies the following important properties:

1. **Positivity for separable states:** If ρ_{AB} is separable, then

$$S(A|B)_\rho \geq 0.[53] \tag{6.2.14}$$

2. **Data Processing Inequality:** If $\tau_{AB'} = \mathbb{I}_A \otimes \mathcal{E}_B(\rho_{AB})$ for a CPTP map $\mathcal{E}$, then

$$S(A|B)_\rho \leq S(A|B')_\tau. \tag{6.2.15}$$

3. **Additivity:** For product states $\rho_{AB} \otimes \tau_{A'B'}$,

$$S(AA'|BB')_{\rho\otimes\tau} = S(A|B)_\rho + S(A'|B')_\tau. \tag{6.2.16}$$

4. **Conditioning on Classical Information (cq-states):** If $\rho_{ABX} = \sum_x p(x)\, |x\rangle\langle x|_X \otimes \rho_{AB|x}$, then

$$S(A|BX)_\rho = \sum_x p(x) S(A|B)_{\rho|x}. \tag{6.2.17}$$

5. **Removing Classical Register:** If X is classical in ρ_{ABX}, then

$$S(A|XB)_\rho \geq S(A|B)_\rho - \log |X|. \tag{6.2.18}$$

These properties are crucial for analyzing the information held by the adversary and for bounding leakage during error correction.

Relevance to QKD. In quantum key distribution, the quantity $S(A|E)_\rho$ captures how uncertain Eve is about Alice's raw key bits. The smaller this entropy, the more Eve knows. After privacy amplification, the goal is to ensure that the final key K_A is nearly uniform and independent of E, i.e.,

$$S(K_A|E) \approx \ell, \tag{6.2.19}$$

where ℓ is the key length. This motivates the use of smooth entropies in one-shot settings and of $S(A|E)$ in asymptotic analyses.

Min-Entropy and Guessing Probability

Given cq-state $\rho_{AE} = \sum_a p(a) \, |a\rangle\langle a| \otimes \rho_{E|a}$,

$$p_{\text{guess}}(A|E) = \sup_{\{M_a\}} \sum_a p(a) \text{Tr}[M_a \rho_{E|a}], \tag{6.2.20}$$

$$H_{\min}(A|E) := -\log p_{\text{guess}}(A|E). \tag{6.2.21}$$

Smooth Entropies

$$H^\varepsilon_{\min}(A|E)_\rho = \max_{\tilde{\rho} \in \mathcal{B}^\varepsilon(\rho)} H_{\min}(A|E)_{\tilde{\rho}}, \tag{6.2.22}$$

$$H^\varepsilon_{\max}(A|E)_\rho = \min_{\tilde{\rho} \in \mathcal{B}^\varepsilon(\rho)} H_{\max}(A|E)_{\tilde{\rho}}, \tag{6.2.23}$$

where $\mathcal{B}^\varepsilon(\rho) = \{\tilde{\rho} : D_{\text{P}}(\rho, \tilde{\rho}) \leq \varepsilon\}$.

Duality for Pure States:

$$H^\varepsilon_{\max}(A|B) = -H^\varepsilon_{\min}(A|C), \quad \text{if } \rho_{ABC} \text{ is pure.} \tag{6.2.24}$$

6.2.4 Leftover Hash Lemma (Privacy Amplification)

Theorem 6.2.1 (Leftover Hashing Lemma)**.** Let $\rho_{A^n E}$ be a cq-state and $\mathcal{F}$ a 2-universal hash family. Then:

$$\|\rho_{K_A F E} - \tau_{K_A} \otimes \rho_{FE}\|_1 \leq \frac{1}{2} \cdot 2^{-\frac{1}{2}(H^\varepsilon_{\min}(A^n|E)_\rho - \ell)} + 2\varepsilon, \tag{6.2.25}$$

where ℓ is the key length and τ_{K_A} is the uniform distribution over K_A.

Key length selection:

$$\ell = H^\varepsilon_{\min}(A^n|E)_\rho - 2\log\left(\frac{1}{2\epsilon_{\text{PA}}}\right) \Rightarrow \epsilon_{\text{sec}} = \epsilon_{\text{PA}} + 2\varepsilon. \tag{6.2.26}$$

6.2.5 Information Reconciliation

Leakage due to public communication affects Eve's knowledge:

$$H^\varepsilon_{\min}(A^n|E_T)_\rho \geq H^\varepsilon_{\min}(A^n|E)_\rho - \text{leak}_{\text{IR}} \tag{6.2.27}$$

6.3 Security Definition for QKD Protocols

Correctness: A QKD protocol is ϵ_{corr}-correct if

$$\Pr[K_A \neq K_B] \leq \epsilon_{\text{corr}}. \tag{6.3.1}$$

Secrecy: Let Ω be the non-abort event with probability $p(\Omega)$. The protocol is ϵ_{sec}-secret if

$$p(\Omega) \cdot \|\rho_{K_A E|\Omega} - \tau_{K_A} \otimes \rho_{E|\Omega}\|_1 \leq \epsilon_{\text{sec}}. \tag{6.3.2}$$

Overall Security: A protocol is ϵ_{QKD}-secure if

$$\epsilon_{\text{QKD}} \geq \epsilon_{\text{corr}} + \epsilon_{\text{sec}}. \tag{6.3.3}$$

Key Rate:

$$r = \frac{\ell}{n} \quad \text{bits/round}, \quad r_{\text{gen}} = \tau \cdot r \quad \text{bits/sec}. \tag{6.3.4}$$

6.4 Semidefinite Programming in the Security Proof of Quantum Key Distribution

Semidefinite programming (SDP) plays a fundamental role in the modern security analysis of Quantum Key Distribution (QKD), particularly in device-independent and finite-key settings. SDP provides a powerful convex optimization framework to analyze the set of quantum states and measurements consistent with observed data, and to rigorously bound the information an adversary (Eve) may obtain about the final key.

6.4.1 Motivation

In QKD, the security is quantified by the amount of information that an adversary can gain about the key. To prove security, we must:

1. Model all quantum states and measurement operators compatible with observed statistics.
2. Quantify the eavesdropper's information using entropic quantities.
3. Optimize over the space of all possible adversarial strategies consistent with quantum mechanics.

These steps naturally lead to semidefinite programs, where the optimization is over quantum states (positive semidefinite matrices) subject to linear constraints.

6.4.2 Semidefinite Programming Basics

A semidefinite program is an optimization problem of the form:

$$\begin{aligned} \text{maximize:} \quad & \operatorname{Tr}(CX) && (6.4.1)\\ \text{subject to:} \quad & \operatorname{Tr}(A_i X) = b_i \quad \forall i \in \{1, \dots, m\} && (6.4.2)\\ & X \succeq 0, && (6.4.3) \end{aligned}$$

where $X \in \mathbb{C}^{n \times n}$ is a Hermitian matrix variable, $A_i, C \in \mathbb{C}^{n \times n}$, and $X \succeq 0$ denotes that X is positive semidefinite.

6.4.3 Security Criterion

Following Renner's composable security definition [45], a QKD protocol is ε-secure if:

$$\frac{1}{2} \left\| \rho_{KE} - \tau_K \otimes \rho_E \right\|_1 \leq \varepsilon, \tag{6.4.4}$$

where ρ_{KE} is the joint state of the key K and the adversary's system E, and τ_K is the uniform distribution on the key space. The key rate r is typically bounded as:

$$r \geq H^{\varepsilon}_{\min}(K|E) - \text{leak}_{\text{EC}}, \tag{6.4.5}$$

where $H^{\varepsilon}_{\min}(K|E)$ is the smooth min-entropy of the key given Eve's information. Computing or bounding $H^{\varepsilon}_{\min}(K|E)$ involves optimization over all quantum states consistent with measurement data — an SDP problem.

6.4.4 SDP in Device-Independent QKD: Example with CHSH

In device-independent QKD (DI-QKD), we do not assume knowledge of the internal workings of the devices. Instead, security is certified via observed violation of a Bell inequality. A prominent example is the CHSH inequality:

$$\langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle \leq 2. \tag{6.4.6}$$

Quantum correlations can violate this bound up to $2\sqrt{2}$. Suppose Alice and Bob observe a CHSH value $S \in [2, 2\sqrt{2}]$. We want to bound Eve's knowledge about Alice's raw key bits. This can be formalized as the following SDP:

$$\begin{aligned} \text{maximize:} \quad & H(K|E) && (6.4.7) \\ \text{subject to:} \quad & \rho_{AB} \succeq 0, \quad \mathrm{Tr}(\rho_{AB}) = 1 && (6.4.8) \\ & \mathrm{Tr}(\rho_{AB} \cdot \mathcal{B}_{\mathrm{CHSH}}) = S && (6.4.9) \\ & \text{Other observed statistics constraints.} && (6.4.10) \end{aligned}$$

Here, $\mathcal{B}_{\mathrm{CHSH}}$ denotes the CHSH Bell operator:

$$\mathcal{B}_{\mathrm{CHSH}} = A_0 \otimes B_0 + A_0 \otimes B_1 + A_1 \otimes B_0 - A_1 \otimes B_1. \tag{6.4.11}$$

Solving this SDP yields a lower bound on $H(K|E)$, and hence on the achievable key rate.

6.4.5 SDP and the NPA Hierarchy

To model the quantum behaviors in a device-independent scenario, the Navascués–Pironio–Acín (NPA) hierarchy [39] provides a sequence of SDP relaxations that approximate the quantum set of correlations. For a given level of the hierarchy, one constructs a moment matrix Γ subject to:

$$\begin{aligned} & \Gamma \succeq 0, && (6.4.12) \\ & \text{Linear constraints from commutation relations and observed statistics.} && (6.4.13) \end{aligned}$$

The feasibility of Γ provides certificates of whether the observed statistics could arise from quantum systems.

6.4.6 Finite-Key Analysis

In practical QKD implementations, only a finite number of rounds are executed. Thus, statistical fluctuations must be taken into account. This leads to the use of hypothesis testing entropy and the Entropy Accumulation Theorem (EAT). SDP is used to bound the conditional entropy at each round by optimizing over all adversarial strategies consistent with the finite statistics.

Chapter 7

Semi-definite Programming

7.1 Definition and Standard Form

A *semidefinite program (SDP)* is an optimization problem that aims to maximize (or minimize) a linear objective function over a matrix variable subject to affine constraints, where the matrix variable is constrained to be positive semidefinite.[50]
The standard form of an SDP is expressed as:

$$\begin{aligned} \alpha = \max \quad & \langle A, X \rangle \\ \text{subject to} \quad & \Phi(X) = B, \\ & X \in \mathrm{Pos}(\mathcal{X}) \end{aligned} \tag{7.1.1}$$

Here:

- $\mathcal{X}, \mathcal{Y}$ are finite-dimensional complex vector spaces.
- $A \in \mathrm{Herm}(\mathcal{X})$ and $B \in \mathrm{Herm}(\mathcal{Y})$ are Hermitian matrices.
- $\Phi : \mathrm{Herm}(\mathcal{X}) \longrightarrow \mathrm{Herm}(\mathcal{Y})$ is a linear map.
- $X \in \mathrm{Pos}(\mathcal{X})$ means X is a positive semidefinite operator.
- $\langle A, X \rangle := \mathrm{Tr}(A^* X)$ is the Hilbert-Schmidt inner product.

Theorem 7.1.1. If $B \neq 0$, then the constraint $\Phi(X) = B$ defines an affine space. Otherwise, it defines a linear subspace.

The *feasible region* is the set:

$$\mathcal{A} := \{X \in \mathrm{Pos}(\mathcal{X}) : \Phi(X) = B\} \tag{7.1.2}$$

For a given feasible region $\mathcal{A}$:

- If $\mathcal{A} = \emptyset$, the SDP is *infeasible*, and $\alpha = -\infty$.
- If $\mathcal{A} \neq \emptyset$, the SDP is *feasible*. Then α is finite or $+\infty$.
- A matrix $X \in \mathcal{A} \cap \mathrm{Pd}(\mathcal{X})$ (i.e., strictly positive definite) is *strictly feasible*.
- If $\alpha = +\infty$, the SDP is *unbounded*.

A matrix $X \in \mathcal{A}$ satisfying $\langle A, X \rangle = \alpha$ is called an *optimal solution*.

Lemma 7.1.2. Even if α is finite, an optimal solution need not exist; the infimum may not be attained.

7.1.1 Minimization Form and Nomenclature

The SDP can equivalently be expressed in minimization form:

$$\begin{aligned} \alpha = \min \quad & \langle A, X\rangle \\ \text{subject to} \quad & \Phi(X) = B, \\ & X \in \mathrm{Pos}(\mathcal{X}) \end{aligned} \tag{7.1.3}$$

In this case:

- The SDP is *unbounded* if $\alpha = -\infty$.
- It is *infeasible* if $\alpha = +\infty$.
- All other notions, such as feasibility and strict feasibility, carry over naturally.

7.1.2 Non-Intuitive Properties

Theorem 7.1.3. Let $A = \begin{bmatrix} t & b \\ b^* & s \end{bmatrix} \in \mathrm{Pos}(\mathbb{C}^2)$. If $s = 0$, then $b = 0$.

Proof. Suppose $s = 0$. The determinant is $\det(A) = ts - |b|^2 = -|b|^2$. Since A is positive semidefinite, we must have $\det(A) \geq 0$, implying $b = 0$. □

This reasoning generalizes: if any diagonal entry of a PSD matrix is zero, then the entire corresponding row and column must be zero.

Theorem 7.1.4. If $A = \begin{bmatrix} t & b \\ b^* & s \end{bmatrix} \in \mathrm{Pos}(\mathbb{C}^2)$, then:

- $t \geq 0, s \geq 0$, and $ts \geq |b|^2$.

Proof. From PSD conditions: $\begin{bmatrix} 1 \\ 0 \end{bmatrix}^* A \begin{bmatrix} 1 \\ 0 \end{bmatrix} = t \geq 0$, similarly $s \geq 0$. Also, $\det(A) = ts - |b|^2 \geq 0 \Rightarrow ts \geq |b|^2$. □

Theorem 7.1.5. The converse holds: If $t, s \geq 0$ and $ts \geq |b|^2$, then $A \in \mathrm{Pos}(\mathbb{C}^2)$.

Proof. Since $ts \geq |b|^2$, the determinant ≥ 0. The eigenvalues are non-negative because their sum $t + s \geq 0$ and product $ts - |b|^2 \geq 0$. □

7.1.3 A Feasible, Bounded SDP without Optimal Solution

Consider the following SDP:

$$\alpha = \min s\text{subject to} \quad \begin{bmatrix} t & 1 \\ 1 & s \end{bmatrix} \in \mathrm{Pos}(\mathbb{C}^2) \tag{7.1.4}$$

Let us analyze feasibility:

- Taking $(s, t) = (1, 1)$, the matrix is PSD, so the problem is feasible. Hence, $\alpha \leq 1$.
- From earlier facts, for PSD, $s > 0$, so $\alpha \geq 0$.
- Consider $s = \varepsilon, t = 1/\varepsilon$ with $\varepsilon > 0$: the matrix remains PSD, and $s \longrightarrow 0$ as $\varepsilon \longrightarrow 0$.

Therefore, $\alpha = 0$, but this value is not attained by any feasible solution (since $s = 0$ is not allowed). Thus:

- The problem is feasible,
- Bounded below,
- Has a finite infimum, but
- No optimal solution exists.

7.2 Duality in Semidefinite Programming

In semidefinite programming (SDP), each optimization problem, referred to as the *primal problem*, is associated with another optimization problem known as the *dual problem*. The dual serves multiple purposes. It offers theoretical insights into the primal problem. It can provide upper or lower bounds on the primal optimal value. It helps identify feasible and optimal solutions.
The general primal form of an SDP is:

$$\begin{aligned} \text{maximize:} \quad & \langle A, X\rangle \\ \text{subject to:} \quad & \Phi(X) = B, \\ & X \in \mathrm{Pos}(\mathcal{X}), \end{aligned} \tag{7.2.1}$$

where $\langle A, X\rangle = \mathrm{Tr}(A^*X)$ denotes the Hilbert–Schmidt inner product, Φ is a linear map, and $\mathrm{Pos}(\mathcal{X})$ denotes the cone of positive semidefinite operators. The dual problem can be expressed in two equivalent formulations:
Dual Form 1:

$$\begin{aligned} \text{minimize:} \quad & \langle B, Y\rangle \\ \text{subject to:} \quad & \Phi^*(Y) \geq A, \\ & Y \in \mathrm{Herm}(\mathcal{Y}), \end{aligned} \tag{7.2.2}$$

Dual Form 2 (slack variable Z):

$$\begin{aligned} \text{minimize:} \quad & \langle B, Y\rangle \\ \text{subject to:} \quad & \Phi^*(Y) - Z = A, \\ & Y \in \mathrm{Herm}(\mathcal{Y}), \quad Z \in \mathrm{Pos}(\mathcal{X}), \end{aligned} \tag{7.2.3}$$

Here, Φ^* is the adjoint map of Φ, defined by the relation $\langle \Phi(X), Y\rangle = \langle X, \Phi^*(Y)\rangle$ for all X, Y. Both dual forms are equivalent, and they rely on the same data A, B, Φ. Consider the primal-dual pair: *Primal Problem:*

$$\begin{aligned} \alpha = \text{maximize:} \quad & \langle H, X\rangle \\ \text{subject to:} \quad & \mathrm{Tr}(X) = 1, \\ & X \in \mathrm{Pos}(\mathcal{X}), \end{aligned} \tag{7.2.4}$$

Dual Problem:

$$\begin{aligned} \beta = \text{minimize:} \quad & y \\ \text{subject to:} \quad & yI - Z = H, \\ & Z \in \mathrm{Pos}(\mathcal{X}), \quad y \in \mathbb{R}. \end{aligned} \tag{7.2.5}$$

To compute $\Phi^*(y)$ where $\Phi = \mathrm{Tr}$, observe:

$$\mathrm{Tr}(X) \cdot y = \langle I, X\rangle \cdot y = \langle X, yI\rangle, \tag{7.2.6}$$

which implies $\Phi^*(y) = yI$. where, α is the optimal value of the primal,β is the optimal value of the dual, the sets of constraints define their respective feasible regions.

7.2.1 Weak Duality

Theorem 7.2.1 (Weak Duality)**.** Let α and β denote the optimal values of the primal and dual semidefinite programs, respectively. Then:

$$\alpha \leq \beta. \tag{7.2.7}$$

Proof. 1. *Case 1: The primal problem is infeasible.*
In this case, there exists no X satisfying both $\Phi(X) = B$ and $X \in \mathrm{Pos}(\mathcal{X})$, hence by convention $\alpha = -\infty$. Since $\beta \in \mathbb{R} \cup \{+\infty\}$, it follows trivially that $\alpha \leq \beta$.

2. *Case 2: The dual problem is infeasible.*
   Here, no pair $(Y, Z) \in \text{Herm}(\mathcal{Y}) \times \text{Pos}(\mathcal{X})$ satisfies $\Phi^*(Y) - Z = A$. Thus, by definition, $\beta = +\infty$, and again the inequality $\alpha \leq \beta$ holds trivially.

3. *Case 3: Both the primal and dual problems are feasible.*
   Let $X \in \text{Pos}(\mathcal{X})$ satisfy $\Phi(X) = B$, and let $(Y, Z) \in \text{Herm}(\mathcal{Y}) \times \text{Pos}(\mathcal{X})$ satisfy $\Phi^*(Y) - Z = A$. Then, observe the following:

$$\begin{aligned}\beta - \alpha &= \inf_{(Y,Z)} \langle B, Y\rangle - \sup_X \langle A, X\rangle \\ &= \inf_{X,Y,Z} \left(\langle Y, \Phi(X)\rangle - \langle A, X\rangle\right) \quad (\text{since } \Phi(X) = B) \\ &= \inf_{X,Y,Z} \left(\langle \Phi^*(Y), X\rangle - \langle A, X\rangle\right) \quad (\text{by adjoint property}) \\ &= \inf_{X,Y,Z} \langle \Phi^*(Y) - A, X\rangle \\ &= \inf_{X,Y,Z} \langle Z, X\rangle \quad (\text{since } \Phi^*(Y) - A = Z).\end{aligned} \tag{7.2.8}$$

   Since both X and Z are constrained to lie in $\text{Pos}(\mathcal{X})$, the inner product $\langle Z, X\rangle \geq 0$. Therefore,

$$\beta - \alpha \geq 0 \quad \Rightarrow \quad \alpha \leq \beta. \tag{7.2.9}$$

This concludes the proof. □

7.2.2 Spectral Example

Let H be a Hermitian matrix with spectral decomposition:

$$H = \sum_i \lambda_i v_i v_i^*, \tag{7.2.10}$$

where λ_i are the real eigenvalues of H, and $\{v_i\}$ are the corresponding orthonormal eigenvectors. Define $y = \lambda_{\max}(H) = \max_i \lambda_i$, the largest eigenvalue of H. We now consider the matrix $yI - H$, and compute:

$$yI - H = \sum_i y v_i v_i^* - \sum_i \lambda_i v_i v_i^* = \sum_i (y - \lambda_i) v_i v_i^*. \tag{7.2.11}$$

Since $y = \lambda_{\max}(H) \geq \lambda_i$ for all i, we have $y - \lambda_i \geq 0$. Therefore, each term $(y - \lambda_i) v_i v_i^*$ is positive semidefinite, and the sum of positive semidefinite matrices is again positive semidefinite. Hence,

$$yI - H \succeq 0, \tag{7.2.12}$$

which shows that the constraint $yI - Z = H$, with $Z = yI - H \in \text{Pos}(\mathcal{X})$, is satisfied. This confirms that $y \in \mathbb{R}$ and $Z \in \text{Pos}(\mathcal{X})$ together define a feasible solution to the dual problem. Since the objective function in the dual is to minimize y, and we have found a specific feasible value $y = \lambda_{\max}(H)$, it follows that:

$$\beta \leq \lambda_{\max}(H). \tag{7.2.13}$$

Combining this with the weak duality result $\alpha \leq \beta$, we conclude:

$$\alpha \leq \beta \leq \lambda_{\max}(H). \tag{7.2.14}$$

This upper bound on the primal optimal value is particularly useful, as it holds even when the primal problem is difficult to solve directly.

7.2.3 Optimality Condition

To demonstrate that the upper bound $\lambda_{\max}(H)$ is tight, we construct a specific feasible solution to the primal problem that achieves this bound. Let v be a unit eigenvector of the Hermitian matrix H corresponding to its largest eigenvalue $\lambda_{\max}(H)$, i.e.,

$$Hv = \lambda_{\max}(H) \cdot v, \quad \text{with} \quad \|v\| = 1. \tag{7.2.15}$$

Define $X = vv^*$, which is a rank-one projector (outer product of v with itself). Since $vv^* \succeq 0$ and $\mathrm{Tr}(vv^*) = \|v\|^2 = 1$, the matrix X is a feasible point for the primal problem. Now, compute the primal objective value:

$$\langle H, X \rangle = \langle H, vv^* \rangle = \mathrm{Tr}(Hvv^*) = \mathrm{Tr}(v^* Hv) = v^* Hv. \tag{7.2.16}$$

Using the eigenvalue equation $Hv = \lambda_{\max}(H)v$, we get:

$$v^* Hv = v^*(\lambda_{\max}(H)v) = \lambda_{\max}(H) \cdot v^* v = \lambda_{\max}(H). \tag{7.2.17}$$

Thus, the feasible matrix $X = vv^*$ attains the value $\langle H, X \rangle = \lambda_{\max}(H)$, which matches the upper bound previously derived from the dual solution. Hence, this solution is not only feasible but also optimal, and we conclude:

$$\alpha = \lambda_{\max}(H), \tag{7.2.18}$$

with the optimal solution explicitly given by $X = vv^*$. This construction demonstrates strong duality in this case, where the primal and dual optimal values coincide and optimal solutions exist for both problems.

7.2.4 Optimality Certificate

In semidefinite programming, the weak duality theorem guarantees that the primal optimal value α is always less than or equal to the dual optimal value β, i.e.,

$$\alpha \leq \beta. \tag{7.2.19}$$

This inequality holds for any pair of feasible solutions X for the primal and (Y, Z) for the dual. However, under certain conditions, this inequality becomes an equality, which is an indication that both X and (Y, Z) are not just feasible, but in fact optimal. Suppose we have a primal feasible matrix $X \in \mathrm{Pos}(\mathcal{X})$ satisfying $\Phi(X) = B$, and a dual feasible pair $(Y, Z) \in \mathrm{Herm}(\mathcal{Y}) \times \mathrm{Pos}(\mathcal{X})$ satisfying $\Phi^*(Y) - Z = A$, such that:

$$\langle A, X \rangle = \langle B, Y \rangle. \tag{7.2.20}$$

This equality has significant consequences. From weak duality, we already know:

$$\langle A, X \rangle \leq \langle B, Y \rangle. \tag{7.2.21}$$

So, the reverse inequality holding as well implies that this is the tightest possible bound. This can only happen when both values are equal to the optimum:

$$\alpha = \langle A, X \rangle, \quad \beta = \langle B, Y \rangle, \quad \text{and} \quad \alpha = \beta. \tag{7.2.22}$$

Therefore, this condition certifies that:

- X is a solution that achieves the primal optimum α,
- (Y, Z) achieves the dual optimum β,
- and there is no duality gap, i.e., strong duality holds.

In conclusion, if the primal and dual feasible solutions satisfy the condition:

$$\langle A, X \rangle = \langle B, Y \rangle, \tag{7.2.23}$$

then we can certify optimality and write:

$$\alpha = \langle A, X \rangle = \langle B, Y \rangle = \beta. \tag{7.2.24}$$

This serves as a practical optimality certificate, allowing us to confirm that both solutions are optimal without having to search over the entire feasible region.

7.2.5 Duality Gap

The *duality gap* in semidefinite programming refers to the difference between the optimal values of the dual and the primal problems. It is defined as:

$$\text{Duality gap} = \beta - \alpha. \tag{7.2.25}$$

By the principle of weak duality, this quantity is always nonnegative:

$$\beta - \alpha \geq 0. \tag{7.2.26}$$

In the ideal case—when certain regularity or constraint qualifications are satisfied—strong duality holds, meaning the gap is exactly zero: $\beta = \alpha$. However, in practice, there are important scenarios where the gap may be strictly positive or even infinite. These are typically due to one or both problems being infeasible or pathological in nature. Several key observations can be made:

- If the primal problem is unbounded above, i.e., $\alpha = +\infty$, then no finite upper bound exists. Since the dual problem always upper bounds the primal, this implies that no feasible solution to the dual problem can exist. Therefore, the dual is infeasible.
- Conversely, if the dual problem is unbounded below, i.e., $\beta = -\infty$, this suggests there is no lower bound on the dual objective. In such cases, the primal problem must be infeasible, otherwise it would contradict weak duality.

These edge cases highlight that a strictly positive or infinite duality gap arises precisely when either the primal or the dual lacks a feasible solution or one of the problems is unbounded. In summary, the duality gap is a fundamental diagnostic tool for analyzing feasibility and optimality in SDPs. While a zero gap certifies optimality under strong duality, a nonzero gap signals potential issues in problem structure, such as unboundedness or constraint inconsistency.

7.3 Slater's Theorem and Optimality Conditions

The strong duality concept in semi-definite programming is concerned with the exploring the condition under which primal and dual optimal values are equal.

7.3.1 Strong Duality via Strict Feasibility

Consider the primal-dual SDP pair.
Primal:

$$\begin{aligned} \text{maximize:} \quad & \langle A, X \rangle \\ \text{subject to:} \quad & \Phi(X) = B, \\ & X \in \mathrm{Pos}(\mathcal{X}). \end{aligned} \tag{7.3.1}$$

Dual:

$$\begin{aligned} \text{minimize:} \quad & \langle B, Y \rangle \\ \text{subject to:} \quad & \Phi^*(Y) \succeq A, \\ & Y \in \mathrm{Herm}(\mathcal{Y}). \end{aligned} \tag{7.3.2}$$

If the primal problem admits a *strictly feasible* solution—meaning there exists an $X \succ 0$ such that $\Phi(X) = B$—then strong duality holds: $\alpha = \beta$, and the dual optimal value is actually attained (assuming $\alpha < \infty$).

Sketch of the Argument

To establish this, define:

$$Q := \operatorname{image}(\Phi) \subseteq \operatorname{Herm}(\mathcal{Y}), \tag{7.3.3}$$

and consider the set:

$$M := \{(S, V, t) \in \operatorname{Herm}(\mathcal{X}) \times Q \times \mathbb{R} \mid \exists X \succeq S,\ \Phi(X) = B - V,\ \langle A, X\rangle \geq t\}, \tag{7.3.4}$$

which represents feasible data under relaxed constraints. Let us define:

$$N := \{(0, 0, s) \mid s > \alpha\}. \tag{7.3.5}$$

The sets M and N are convex and disjoint by construction. By the separating hyperplane theorem, there exists a non-zero linear functional—represented by a triple (Z, Y, λ)—and a constant $C \in \mathbb{R}$ such that:

$$\begin{aligned} \forall (S, V, t) \in M: &\quad \langle Z, S\rangle + \langle Y, V\rangle + \lambda t \leq C, \\ \forall (0, 0, s) \in N: &\quad \lambda s \geq C. \end{aligned} \tag{7.3.6}$$

The second inequality implies $\lambda \geq 0$ and $C \leq \lambda\alpha$. Assuming $\lambda = 0$ leads to a contradiction, as we would obtain $Z = 0$ and $Y = 0$ by analyzing the implications on strictly feasible X. Thus, $\lambda > 0$ must hold. We can rescale the inequality and rewrite:

$$\langle Z, S\rangle + \langle Y, V\rangle + t \leq \alpha. \tag{7.3.7}$$

By analyzing this inequality over arbitrary Hermitian S, we conclude that the constraint $Z - \Phi^*(Y) + A = 0$ must hold; otherwise, one could choose S to violate the bound. Therefore, the pair (Y, Z) satisfies the dual feasibility condition. Moreover, the bound gives:

$$\langle B, Y\rangle \leq \alpha. \tag{7.3.8}$$

Combining this with weak duality $\alpha \leq \beta \leq \langle B, Y\rangle$ gives:

$$\alpha = \beta = \langle B, Y\rangle, \tag{7.3.9}$$

establishing both strong duality and dual attainment.

7.3.2 Consequences of Duality Symmetry

Since the dual of the dual is the primal, the argument is symmetric. If the dual problem is strictly feasible (i.e., $\Phi^*(Y) \succ A$), then strong duality holds and the primal optimum is also attained, assuming $\beta < \infty$. If both primal and dual problems admit strictly feasible solutions, then strong duality holds and both optima are achieved.

Example 7.3.1. Consider the following SDP pair: *Primal:*

$$\text{maximize:} \quad \langle H, X\rangle \quad \text{subject to } X \in \mathcal{D}(\mathcal{X}), \tag{7.3.10}$$

where $\mathcal{D}(\mathcal{X})$ denotes the set of density operators (i.e., positive semidefinite with trace one). *Dual:*

$$\text{minimize:} \quad y \quad \text{subject to } yI \succeq H. \tag{7.3.11}$$

A dual feasible solution is $y = \lambda_{\max}(H) + 1$, which is strictly greater than all eigenvalues of H, hence ensuring strict feasibility. Similarly, choosing $X = \frac{1}{\dim(\mathcal{X})} I$ gives a strictly feasible primal solution. Hence, both problems attain optimal values and satisfy $\alpha = \beta$, although the actual optimizers are not specified.

7.3.3 Orthogonality of PSD Matrices

Suppose $X \succeq 0$, $Y \succeq 0$. Then:

$$\langle X, Y \rangle = 0 \quad \text{if and only if} \quad XY = 0. \tag{7.3.12}$$

This result shows that inner product zero for PSD matrices implies operator orthogonality. It follows from the spectral decomposition:

$$X = \sum_i \lambda_i v_i v_i^*, \quad Y = \sum_j \mu_j w_j w_j^*, \tag{7.3.13}$$

and computing:

$$\langle X, Y \rangle = \sum_{i,j} \lambda_i \mu_j |\langle v_i, w_j \rangle|^2 = 0 \quad \Rightarrow \quad \langle v_i, w_j \rangle = 0 \, \forall i, j. \tag{7.3.14}$$

7.3.4 Complementary Slackness

Let X be primal feasible and (Y, Z) be dual feasible. Then the following are equivalent:

1. $\langle A, X \rangle = \langle B, Y \rangle$;
2. $\Phi^*(Y)X = AX$.

This equivalence expresses the condition of complementary slackness in SDP. From weak duality, equality of objective values implies $\langle \Phi^*(Y) - A, X \rangle = 0$, which in turn implies the matrix product $(\Phi^*(Y) - A)X = 0$, due to positivity.

7.3.5 Optimality Conditions

Suppose X and (Y, Z) satisfy the following:

- X is primal feasible;
- (Y, Z) is dual feasible;
- $\Phi^*(Y)X = AX$.

Then X and (Y, Z) are optimal solutions for their respective problems. Furthermore, if both the primal and the dual problems admit strictly feasible solutions, then such a pair (X, Y, Z) satisfying the above always exists.

7.3.6 Practical Strategy

To effectively analyze or solve a given semidefinite program, follow this procedure:

1. First, check whether the problem is trivially infeasible or unbounded.
2. Compute the dual formulation of the primal problem.
3. Identify strictly feasible solutions if possible. These help predict whether strong duality will hold, and provide useful bounds or guidance even in numerical approaches.

Chapter 8

Security Proof of Device Independent QKD using Random Key Basis

Security proof of any QKD protocol is crucial to detect any vulnerability and potential threat to a QKD protocol. This chapter address the same. It's an improvisation over the existing security proof of Device-Independent Quantum Key Distribution Protocol using Random Key Basis.
The first section address the framework of the protocol and hence the proof where we define objective function that we aim to optimize in term of devetak winter rate. The next section reformulate the change in entropy in term of accessible Alice's and Bob's subsystem. In the next section we worked on existing reduction of the problem to a two qubit space where we derive a closed form of angles from the respective eigenvalues. Moreover we also prove that such a mapping is unique. In this section we also classify the projector into two group showing classical and quantum correlation respectively which in turn will help in the further analysis. We also derived a closed form of the CHSH operator. The next section, Section 4 work on the existing analysis of transforming the problem in term of trace norm through a modified Pinsker inequality. In Section 5 we work on the existing SDP formulation of the objective fucntion for fixed angles. The next section, Section 6 deals with optimising alice and bob angles using ϵ-net approach. In the work [48] showed that an polytope optimisation is required for bob angle, here we show that ϵ-net approach suffice. We also derived an closed form of pessimistic error. In order to find a closed form of the pessimistic error. We analysis the effect of perturbation of alice and bob angle upon the CHSH operator and show that the perturbation of alice and bob angle by ϵ show the max deviation when Bob angle is being perturbated. We also studied effect of perturbation over the objective function and lowered the error induced due to pertubation in the objective fucntion. The final and last section deal with creating an convex hull from all the two qubit blocks resulting the CHSH value $S \in (2, 2\sqrt{2}$ using Jensen's inequality.

8.1 Framework

The aim is to establish a lower bound on the quantity $C^*(S)$ [48] where S is the observed CHSH value for the shared entangled state, such that

$$\lambda H(A_0|E)_{\rho_{A'BEA_0}} + (1-\lambda)H(A_1|E)_{\rho_{A'BEA_1}} \geq C^*(S), \tag{8.1.1}$$

where A_x is Alice's outcome for $X \in \{0,1\}$ and $H(A_X|E)_{\rho_{A'BEA_x}}$ is the conditional entropy given Eve's side channel information characterized by E. The protocol has a predefined set of settings for Alice and Bob observables. $\{0,1\}$ for Alice and $\{0,1,2,3\}$ for Bob. Let $\mathcal{X}$ and $\mathcal{Y}$ be the complex Euclidean space defining Alice's and Bob's subsystems, respectively.
Now consider $O_x^{\mathcal{X}} \in \text{Herm}(\mathcal{X})$ for $x \in \{0,1\}$ and $O_y^{\mathcal{Y}} \in \text{Herm}(\mathcal{Y})$ for $y \in \{0,1,2,3\}$ are the observables for respective sub-scripted settings. The observables can be explicitly written as $A = \{O_0^{\mathcal{X}}, O_1^{\mathcal{X}}\}$ for Alice and

$B = \{O_0^{\mathcal{Y}}, O_1^{\mathcal{Y}}, O_2^{\mathcal{Y}}, O_3^{\mathcal{Y}}\}$ for Bob Using spectral decomposition, one can get

$$O_x^{\mathcal{X}} = \sum_i \lambda_i^x P_i^x, \quad O_y^{\mathcal{Y}} = \sum_j \lambda_j^y P_j^y \tag{8.1.2}$$

where the eigenvalues λ_i are associated with the corresponding outcomes. Thus, each observation is associated with two projectors, and each projector corresponds to one particular outcome among $\{0,1\}$. Let us denote the projector as $P_V^{o|z}$ where $o \in \{0,1\}$ corresponds to the two outcomes, $z \in \{0,1,2,3\}$ corresponds to the particular observables and $V \in \mathcal{X}, \mathcal{Y}$ corresponds to a particular complex Euclidean space of Alice's or Bob's subsystem.
Since one is deriving the lower bound on the quantity $C^*(S)$ which is solely a function of the expected value of the CHSH operator, one can perform the analysis using only those observables that define the CHSH operator in our case viz. observables corresponds to $\{0,1\}$ settings for Alice and $\{2,3\}$ settings for Bob. One can define the correlation operators as

$$\begin{aligned} C^{O_0^{\mathcal{X}}} &= P_{\mathcal{X}}^{0|0} - P_{\mathcal{X}}^{1|0} \\ C^{O_1^{\mathcal{X}}} &= P_{\mathcal{X}}^{0|1} - P_{\mathcal{X}}^{1|1} \\ C^{O_2^{\mathcal{Y}}} &= P_{\mathcal{Y}}^{0|2} - P_{\mathcal{Y}}^{1|2} \\ C^{O_3^{\mathcal{Y}}} &= P_{\mathcal{Y}}^{0|3} - P_{\mathcal{Y}}^{1|3} \end{aligned} \tag{8.1.3}$$

Let us define a and b as the respective outcomes of Alice and Bob after the action of the observables from the chosen set, viz A and B respectively, as defined above on their respective subsystem of the Bell state. One can always define the joint probability of obtaining the outcomes a and b when Alice and Bob measure their respective subsystems using the measurement settings x and y as

$$P(a,b|x,y) = \mathrm{Tr}(\rho \cdot (P_{\mathcal{X}}^{a|x} \otimes P_{\mathcal{Y}}^{b|y})) \tag{8.1.4}$$

The joint probability of obtaining the same outcomes or different outcomes is defined using (8.1.4) as

$$\begin{aligned} P(a=b|x,y) &= \sum_{a=b} P(a,b|x,y) \\ P(a \neq b|x,y) &= \sum_{a \neq b} P(a,b|x,y) \end{aligned} \tag{8.1.5}$$

The probability of obtaining the same or a different outcome is given by

$$\begin{aligned} P(a=b) &= \sum_{x,y} P_x \cdot P_y \cdot P(a=b|x,y) \\ P(a \neq b) &= \sum_{x,y} P_x \cdot P_y \cdot P(a \neq b|x,y) \end{aligned} \tag{8.1.6}$$

where P_x and P_y are the probability of choosing the x and y setting for Alice and Bob respectively.
The correlation function is now defined as

$$C_{xy} = P(a=b|x,y) - P(a \neq b|x,y) \tag{8.1.7}$$

Now combing the results of (8.1.7), (8.1.5) and (8.1.4) One can get

$$C_{xy} = \mathrm{Tr}(\rho \cdot (P_{\mathcal{X}}^{a|x} \otimes P_{\mathcal{Y}}^{a|y})) - \mathrm{Tr}(\rho \cdot (P_{\mathcal{X}}^{a|x} \otimes P_{\mathcal{Y}}^{b|y})) \tag{8.1.8}$$

Now, one can define the CHSH value S as

$$S = max(2, C_{12} - C_{02} - C_{03} - C_{13}) \tag{8.1.9}$$

The underlying CHSH operator can be expressed in terms of correlation operators [48] as

$$\begin{aligned} CHSH \;=\;& C^{O_1^{\mathcal{X}}} \otimes C^{O_2^{\mathcal{Y}}} - C^{O_0^{\mathcal{X}}} \otimes C^{O_2^{\mathcal{Y}}} \\ -\;& C^{O_0^{\mathcal{X}}} \otimes C^{O_3^{\mathcal{Y}}} - C^{O_1^{\mathcal{X}}} \otimes C^{O_3^{\mathcal{Y}}} \end{aligned} \tag{8.1.10}$$

The relation between the CHSH operator expressed in (8.1.10) and the expected value as in (8.1.9) when it acts on our state, can easily be shown in some simple algebra.
The terms in the operator can be expressed as [57]

$$\begin{aligned} C^{O_x^{\mathcal{X}}} \otimes C^{O_y^{\mathcal{Y}}} &= (P_{\mathcal{X}}^{0|x} - P_{\mathcal{X}}^{1|x}) \otimes (P_{\mathcal{Y}}^{0|y} - P_{\mathcal{Y}}^{1|y}) \\ &= (P_{\mathcal{X}}^{0|x} \otimes P_{\mathcal{Y}}^{0|y}) - (P_{\mathcal{X}}^{0|x} \otimes P_{\mathcal{Y}}^{1|y}) - (P_{\mathcal{X}}^{1|x} \otimes P_{\mathcal{Y}}^{0|y}) + (P_{\mathcal{X}}^{1|x} \otimes P_{\mathcal{Y}}^{1|y}) \\ &= (P_{\mathcal{X}}^{0|x} \otimes P_{\mathcal{Y}}^{0|y} + P_{\mathcal{X}}^{1|x} \otimes P_{\mathcal{Y}}^{1|y}) - (P_{\mathcal{X}}^{0|x} \otimes P_{\mathcal{Y}}^{1|y} + P_{\mathcal{X}}^{1|x} \otimes P_{\mathcal{Y}}^{0|y}) \end{aligned} \tag{8.1.11}$$

Now, by applying the Born rule to find the expectation value of the CHSH operator on our given Bell state, one can easily get (8.1.9). On the assumption that the state shared between Alice and Bob is a mixed state ρ_{AB} in general, and Eve holds the purification of this state, the pure state is now described by a joint system of Alice, Bob, and Eve Ψ_{ABE}.A more particular description of the state Ψ_{ABEA_x} may include a fourth component that's Alice's outcome A_x, which is initially associated with a pure quantum state ψ. The dimension of Eve's system is unknown, and Eve has full control over all the devices, including Alice's and Bob's measurement devices. Eve has to encounter a certain amount of uncertainty due to the underlying principle of local realism.
Our objective function as explained in (8.1.1) is to establish a lower bound on the convex combination on terms involving conditional Von-Neuman entropies between Alice's outcome A_x and Eve's information E, one can analyze the same conditional entropies $H(A_x|E)_{\rho_{A'BEA_x}}$ in terms of mutual information and Holevo information which are more standard in context of quantum information theory[29]. Thereby, it has been shown the equivalency between the Devetak winter rate and the conditional Von Neumann entropy. The Conditional Von Neuman Entropy can be decomposed as $H(A_x|E)_{\rho_{A'BEA_x}} = H(A_x)_{\rho_{A'BEA_x}} - \chi(A_x;E)$ where $\chi(A_x;E)$ is the Holevo information between Alice (for setting x) and Eve. Similarly, the Von Neuman entropy can be expressed in terms of mutual information and conditional entropy as $H(A_x)_{\rho_{A'BEA_x}} = I(A_x;B)_{\rho_{A'BEA_x}} + H(A_x|B)_{\rho_{A'BEA_x}}$ where $H(A_x)_{\rho_{A'BEA_x}}$ is the Von Neuman entropy and $I(A_x;B)_{\rho_{A'BEA_x}}$ is the mutual information between Alice and Bob for a compatible Bob setting y. Now, by using the above two relations, one may get

$$H(A_x|E)_{\rho_{A'BEA_x}} = I(A_x;B)_{\rho_{A'BEA_x}} + H(A_x|B)_{\rho_{A'BEA_x}} - \chi(A_x;E). \tag{8.1.12}$$

The (8.1.1) can be rewritten using (8.1.12) as

$$\begin{aligned} C^*(S) \leq \lambda & \left[I(A_0;B)_{\rho_{A'BEA_0}} + H(A_0|B)_{\rho_{A'BEA_0}} - \chi(A_0;E)\right] \\ & + (1-\lambda)\left[I(A_1;B)_{\rho_{A'BEA_1}} + H(A_1|B)_{\rho_{A'BEA_1}} - \chi(A_1;E)\right] \end{aligned} \tag{8.1.13}$$

Under the assumption of optimal error correction ($H(A_x|B_y)_{\rho_{A'BEA_x}} \to 0$), the bound simplifies to:

$$C^*(S) \leq \lambda\left[I(A_0;B)_{\rho_{A'BEA_0}} - \chi(A_0;E)\right] + (1-\lambda)\left[I(A_1;B)_{\rho_{A'BEA_1}} - \chi(A_1;E)\right]. \tag{8.1.14}$$

The upper bound on $C^*(S)$ is now equivalent to the weighted sum of mutual information between Alice and Bob (for correlated settings) minus the weighted sum of Holevo information between Alice and Eve,

$$C^*(S) \leq \left[\lambda I(A_0;B)_{\rho_{A'BEA_0}} + (1-\lambda) I(A_1;B)_{\rho_{A'BEA_1}}\right] - \left[\lambda\chi(A_0;E) + (1-\lambda)\chi(A_1;E)\right]. \tag{8.1.15}$$

This shows the equivalency between the Devatak Winter rate regarding mutual information, Holevo information, and the conditional Von Neumann entropic relation. Now, after Alice applies her observables $O_x^{\mathcal{X}}$ in accordance to measurement settings x for $x \in \{0,1\}$ the quantum-classical state obtained is given by,

$$\rho_{A'BEA_x} = \sum_{j\in\{0,1\}} \rho_{\psi_j} \otimes \mathrm{Tr}_A((P_{\mathcal{X}}^{j|x} \otimes \mathbb{1}_{BE})\Psi_{ABE}(P_{\mathcal{X}}^{j|x} \otimes \mathbb{1}_{BE})^*) \tag{8.1.16}$$

where A', B, E, A_x are Alice's subsystem over $\mathcal{X}$ after she applied her chosen observables $O_x^{\mathcal{X}}$ for settings x, Bob's subsystem over $\mathcal{Y}$, Eve's subsystem whose vector space is unknown, and Alice's outcomes subsystem after the effect of her chosen observables associated with pure quantum state ψ respectively. $\rho_{\psi_j} \in D(\mathcal{O})$ for some chosen complex Euclidean Space $\mathcal{O}$ of the outcome registers A_x space of density operators are defined. The conditional Von Neumann entropic term $H(A_x|E)_{\rho_{A'BEA_x}}$ involving Alice's outcome conditioned by Eve's information, which essentially captures the change in entropy by following a simple relation,

$$\begin{aligned} H(A_x|E)_{\rho_{A'BEA_x}} &= H(A_xE)_{\rho_{A'BEA_x}} - H(E)_{\rho_{ABEA'_x}} \\ &= H(A'B)_{\rho_{A'BEA_x}} - H(AB)_{\rho_{ABEA'_x}} \\ &= \Delta H_x \end{aligned} \tag{8.1.17}$$

The optimization problem includes a constraint that the reduced density operator of the joint subsystem comprising Alice's and Bob's systems, denoted as ρ_{AB}, must be equal to the partial trace of the overall state $\rho_{ABEA'_x}$ over the environment (E) and Alice's ancilla (A'_x). Mathematically, this constraint is expressed as: $\rho_{AB} = \mathrm{Tr}_{EA'_x}(\rho_{ABEA'_x})$ Now one can formally define the optimization problem to establish the lower bound the quantity $C^*(S)$ defined in terms of conditional Von Neuman entropies as in (8.1.1) or equivalently the Devetak-Winter rate as in (8.1.14) as

$$\boxed{\begin{aligned} C^*(S) = \inf \quad & \lambda H(A_0|E)_{\rho_{A'BEA_0}} + (1-\lambda)H(A_1|E)_{\rho_{A'BEA_1}} \\ \text{s.t.} \quad & \mathrm{Tr}(\rho_{AB} \cdot \mathrm{CHSH}) = S \\ & \rho_{AB} \succeq 0 \\ & \mathrm{Tr}(\rho_{AB}) = 1 \\ \text{or equivalently,} & \\ C^*(S) = \inf \quad & \lambda \left[I(A_0;B)_{\rho_{A'BEA_0}} - \chi(A_0;E) \right] + (1-\lambda) \left[I(A_1;B)_{\rho_{A'BEA_1}} - \chi(A_1;E) \right] \\ \text{s.t.} \quad & \mathrm{Tr}(\rho_{AB} \cdot \mathrm{CHSH}) = S \\ & \rho_{AB} \succeq 0 \\ & \mathrm{Tr}(\rho_{AB}) = 1 \end{aligned}} \tag{8.1.18}$$

8.2 Reformulation of change in entropy ΔH term in of Alice and Bob subsystems

After formally defining the optimization problem, our first task is to reformulate the problem in terms of the Alice and Bob subsystems A and B. The problem defined in (8.1.18) is defined in terms of Alice's outcome A_x and Eve's subsystem E.This type of setting has a serious flaw: inaccessibility to the Eve subsystem E. Although one can access Alice's outcome A_x in the context of security analysis, access to the Eve subsystem is impractical and meaningless. The Eve subsystem generally consists of a vector space whose dimension is unknown, and if one has access to Eve's subsystem, one can configure the protocol accordingly to bypass Eve's intervention, which is not only practically infeasible but also meaningless in the Device Independent QKD framework.

The whole state ρ is a pure state; one may consider it as a bipartite system where A_x and E are considered one side of the bipartition and A and B are considered on the other side of the bipartition. Now, from the result of *Theorem 2.c* of [4], one can relate that the local von-Neumann entropy production ΔH_x on the two sides of the bipartition is essentially the same

$$\Delta H_x = H(A'B)_{\rho_{A'BEA_x}} - H(AB)_{\rho_{ABEA'_x}} \tag{8.2.1}$$

and thus from (8.1.17), one can perform the analysis in terms of accessible Alice and Bob subsystems.[48] The transformation can be seen as the transformation from the state $\rho_{ABEA'_x} \in \mathcal{R}$ where $\mathcal{R}$ defines the joint state describing Alice's system A, Bob's system B, Eve's system E, and the A'_x being Alice's outcomes subsystem before the effect of her chosen observables associated with pure quantum state ψ to $\rho_{A'BEA_x} \in \mathcal{H}$ where $\mathcal{H}$ denotes the same as a joint state describing Alice Bob and Eve system except for the fact that here

A' denotes the Alice subsystem after the transformation, and A_x is the space describing the register storing the outcomes of the effect of the observables after the given transformation. The work of [48] shows that this transformation can indeed be defined through pinching channels defined for each observable as,

$$\begin{aligned}\Lambda_0[\rho_{ABEA'_0}] = (P_{\mathcal{X}}^{0|0} \otimes \mathbb{1})\rho_{ABEA'_0}(P_{\mathcal{X}}^{0|0} \otimes \mathbb{1}) + (P_{\mathcal{X}}^{1|0} \otimes \mathbb{1})\rho_{ABEA'_0}(P_{\mathcal{X}}^{1|0} \otimes \mathbb{1}) = \rho_{A'BEA_0} \\ \Lambda_1[\rho_{ABEA'_1}] = (P_{\mathcal{X}}^{0|1} \otimes \mathbb{1})\rho_{ABEA'_1}(P_{\mathcal{X}}^{0|1} \otimes \mathbb{1}) + (P_{\mathcal{X}}^{1|1} \otimes \mathbb{1})\rho_{ABEA'_1}(P_{\mathcal{X}}^{1|1} \otimes \mathbb{1}) = \rho_{A'BEA_1}\end{aligned} \tag{8.2.2}$$

Now let's redefine the reduced joint state of Alice and Bob as,

$$\begin{aligned}Tr_{EA_x}(\Lambda[\rho_{ABEA'_x}]) = \Lambda[\rho_{ABEA'_x}]^{A'B} \\ Tr_{EA_x}(\rho_{ABEA'_x}) = \rho^{AB}_{ABEA'_x}\end{aligned} \tag{8.2.3}$$

8.2.1 Pinching Channel

The pinching channel Λ is defined as:

$$\Lambda[\rho] = \sum_i P_i \rho P_i \tag{8.2.4}$$

where $\{P_i\}$ is a set of orthogonal projectors satisfying the *Kronecker delta function* $P_iP_j = \delta_{ij}P_i$ for *orthogonality*, and $\sum_i P_i = \mathbb{1}$ for *completeness*. Here, $\rho \in$ is a density operator.

Properties of Pinching Channels

1. Idempotence: The pinching channel is idempotent, meaning:

$$\Lambda[\Lambda[\rho]] = \Lambda[\rho]. \tag{8.2.5}$$

Proof:

$$\Lambda[\Lambda[\rho]] = \sum_i P_i \left(\sum_j P_j \rho P_j \right) P_i. \tag{8.2.6}$$

Using the orthogonality of the projectors ($P_iP_j = \delta_{ij}P_i$), this simplifies to:

$$\Lambda[\Lambda[\rho]] = \sum_i P_i \rho P_i = \Lambda[\rho]. \tag{8.2.7}$$

Thus, Λ is idempotent.

2. Commutativity with Projectors: If ρ commutes with all P_i, then:

$$\Lambda[\rho] = \rho. \tag{8.2.8}$$

Proof: If ρ commutes with P_i, then $P_i\rho = \rho P_i$. Substituting into the definition of Λ:

$$\Lambda[\rho] = \sum_i P_i \rho P_i = \sum_i P_i P_i \rho = \sum_i P_i \rho = \rho, \tag{8.2.9}$$

where one can used $P_i^2 = P_i$ (since P_i is a projector) and $\sum_i P_i = I$.

3. Non-Increase of the Trace Distance: The trace distance between two states ρ and σ does not increase under the pinching channel:

$$\|\Lambda[\rho] - \Lambda[\sigma]\|_1 \leq \|\rho - \sigma\|_1. \tag{8.2.10}$$

Proof: The trace distance is defined as:

$$\|\rho - \sigma\|_1 = Tr|\rho - \sigma|, \tag{8.2.11}$$

where $|A| = \sqrt{A^*A}$. The pinching channel Λ is a completely positive trace-preserving (CPTP) map, and all CPTP maps are contractive concerning the trace norm. Thus:

$$\|\Lambda[\rho] - \Lambda[\sigma]\|_1 \leq \|\rho - \sigma\|_1. \tag{8.2.12}$$

4. Preservation of Diagonal Terms: For any state ρ, the diagonal terms of ρ in the basis of $\{P_i\}$ remain unchanged after applying Λ.
   Proof: The diagonal terms of ρ in the basis of $\{P_i\}$ are given by $\mathrm{Tr}(P_i\rho)$. Applying Λ:

$$\mathrm{Tr}(P_i\Lambda[\rho]) = \mathrm{Tr}\left(P_i \sum_j P_j \rho P_j\right) = \mathrm{Tr}(P_i\rho P_i) = \mathrm{Tr}(P_i\rho) \tag{8.2.13}$$

   one can use the trace's cyclic property and $P_iP_j = \delta_{ij}P_i$. Thus, the diagonal terms are preserved.

5. Entropy Properties: Pinching channels are doubly stochastic maps, meaning they do not decrease the von Neumann entropy:

$$S(\Lambda[\rho]) \geq S(\rho), \tag{8.2.14}$$

   where $S(\rho) = -\mathrm{Tr}(\rho \log_2 \rho)$ is the von Neumann entropy.
   Proof: The pinching channel Λ is unital ($\Lambda[I] = I$) and doubly stochastic. By the monotonicity of the von Neumann entropy under doubly stochastic maps, one can have:

$$S(\Lambda[\rho]) \geq S(\rho). \tag{8.2.15}$$

6. Majorization: The eigenvalues of $\Lambda[\rho]$ majorize by the eigenvalues of ρ, reflecting a redistribution of probabilities towards uniformity.
   Proof: Let $\lambda(\rho)$ and $\lambda(\Lambda[\rho])$ denote the vectors of eigenvalues of ρ and $\Lambda[\rho]$, respectively. The pinching channel Λ is an unital quantum channel, and by the Schur-Horn theorem, the eigenvalues of $\Lambda[\rho]$ majorize the eigenvalues of ρ:

$$\lambda(\Lambda[\rho]) \prec \lambda(\rho). \tag{8.2.16}$$

   This means that the eigenvalues of $\Lambda[\rho]$ are more uniformly distributed than those of ρ.

7. Commutativity with Certain Functions: For any measurable function f, the pinching channel satisfies:

$$f(\Lambda[\rho]) = \Lambda[f(\rho)]. \tag{8.2.17}$$

   Proof: The pinching channel Λ acts as a projection onto the diagonal basis defined by $\{P_i\}$. For any function f, the action of f on ρ commutes with the pinching operation because f acts on the eigenvalues of ρ, and Λ preserves the diagonal terms (eigenvalues) while removing off-diagonal terms. Thus:

$$f(\Lambda[\rho]) = \Lambda[f(\rho)]. \tag{8.2.18}$$

In quantum information theory, the dual (adjoint) of a map Φ, denoted Φ^*, is defined via the Hilbert-Schmidt inner product:

$$\mathrm{Tr}(A \cdot \Phi[\rho]) = \mathrm{Tr}(\Phi^*[A] \cdot \rho), \quad \forall A, \rho \in L(\mathcal{V}), \tag{8.2.19}$$

where A and ρ are Hermitian operators on a complex Euclidean space $\mathcal{V}$. More generally, for any linear map Φ, the adjoint satisfies:

$$\mathrm{Tr}(A \cdot \Phi[B]) = \mathrm{Tr}(\Phi^*[A] \cdot B), \quad \forall A, B \in L(\mathcal{V}), \tag{8.2.20}$$

where A and B are Hermitian operators on a complex Euclidean space $\mathcal{V}$. Now since pinching channel Λ are self adjoint from (8.2.20) one can have,

$$\mathrm{Tr}(A \cdot \Lambda[\rho]) = \mathrm{Tr}(\rho \cdot \Lambda^*[A]) = \mathrm{Tr}(\rho \cdot \Lambda[A]) \quad \forall A, \rho \in L(\mathcal{V}) \tag{8.2.21}$$

Now, one can find a mathematical expression for the production of entropy ΔH_x using (8.2.3)

$$\begin{aligned} \Delta H_x = H(A'B)_{\rho_{A'BEA_x}} - H(AB)_{\rho_{ABEA'_x}} \\ = H(Tr_{EA_x}(\Lambda[\rho_{ABEA'_x}])) - H(Tr_{EA'_x}(\rho_{ABEA'_x})) \\ = H(\Lambda[\rho_{ABEA'_x}]^{A'B}) - H(\rho^{AB}_{ABEA'_x}) \end{aligned} \tag{8.2.22}$$

Thus from (8.2.22) and (8.2.21) one can continue by putting $A = (\Lambda[\rho_{ABEA'_x}]^{A'B})$ and $B = \rho^{AB}_{ABEA'_x}$ are the reduced density operator representing the Alice and Bob subsystem after the application of the pinching channel Λ and before applying the pinching channel as,

$$\begin{aligned}
&H(\Lambda[\rho_{ABEA'_x}]^{A'B}) - H(\rho^{AB}_{ABEA'_x}) \\
&= -\mathrm{Tr}(\Lambda[\rho_{ABEA'_x}]^{A'B} \cdot \log_2(\Lambda[\rho_{ABEA'_x}]^{A'B})) + \mathrm{Tr}(\rho^{AB}_{ABEA'_x} \cdot \log_2(\rho^{AB}_{ABEA'_x})) \\
&= -\mathrm{Tr}(\rho^{AB}_{ABEA'_x} \cdot \Lambda^*(\log_2(\Lambda[\rho_{ABEA'_x}]^{A'B}))) + \mathrm{Tr}(\rho^{AB}_{ABEA'_x} \cdot \log_2(\rho^{AB}_{ABEA'_x}))
\end{aligned} \tag{8.2.23}$$

Now, from (8.2.5) one can have,

$$\begin{aligned}
&-\mathrm{Tr}(\rho^{AB}_{ABEA'_x} \cdot \Lambda(\log_2(\Lambda[\rho_{ABEA'_x}]^{A'B}))) + \mathrm{Tr}(\rho^{AB}_{ABEA'_x} \cdot \log_2(\rho^{AB}_{ABEA'_x})) \\
&= -\mathrm{Tr}(\rho^{AB}_{ABEA'_x} \cdot \log_2(\Lambda[\rho_{ABEA'_x}]^{A'B})) + \mathrm{Tr}(\rho^{AB}_{ABEA'_x} \cdot \log_2(\rho^{AB}_{ABEA'_x})) \\
&= \mathrm{Tr}(\rho^{AB}_{ABEA'_x} \cdot \log_2(\rho^{AB}_{ABEA'_x})) - Tr(\rho^{AB}_{ABEA'_x} \cdot \log_2(\Lambda[\rho_{ABEA'_x}]^{A'B})) \\
&= \mathrm{Tr}(\rho^{AB}_{ABEA'_x} \cdot (\log_2(\rho^{AB}_{ABEA'_x}) - \log_2(\Lambda[\rho_{ABEA'_x}]^{A'B}))) \\
&= \mathrm{Tr}(\rho^{AB}_{ABEA'_x}(\log_2(\rho^{AB}_{ABEA'_x}) - \log_2(\Lambda[\rho_{ABEA'_x}]^{A'B}))) \\
&= D((\rho^{AB}_{ABEA'_x})||(\Lambda[\rho_{ABEA'_x}]^{A'B}))
\end{aligned} \tag{8.2.24}$$

Here the relative entropy D is defined as $D(\rho||\sigma) = \mathrm{Tr}(\rho(log(\rho) - log(\sigma)))$ [57].
Thus, one can reformulate the entropy production in terms of relative von Neumann entropy as

$$\Delta H_x = D((\rho^{AB}_{ABEA'_x})||(\Lambda[\rho_{ABEA'_x}]^{A'B})) \tag{8.2.25}$$

Our main objective function in (8.1.18) is a convex combination of conditional von Neumann entropies. Using (8.1.17) and (8.2.25) the objective function can be reformulate as

$$\boxed{\begin{aligned}
C^*(S) = \inf \quad & \lambda D((\rho^{AB}_{ABEA'_0})||(\Lambda_0[\rho_{ABEA'_0}]^{A'B})) \\
& +(1-\lambda)D((\rho^{AB}_{ABEA'_1})||(\Lambda_1[\rho_{ABEA'_1}]^{A'B})) \\
\text{s.t.} \quad & \mathrm{Tr}(\rho_{AB} \cdot \mathrm{CHSH}) = S \\
& \rho_{AB} \succeq 0 \\
& \mathrm{Tr}(\rho_{AB}) = 1
\end{aligned}} \tag{8.2.26}$$

One can have reformulated the convex combination of conditional Von-Neuman entropies of Alice's outcomes and Eve's subsystem in terms of the relative entropy of Alice's and Bob's subsystems before and after the pinching channel Λ.

8.3 Reduction of the problem to two qubits space $(\mathbb{C}_{4\times4})$ of Alice and Bob

Let $\rho_{A'BEA_x}, \rho_{ABEA'_x}, \in \mathcal{D}(\mathcal{P})$ for some chosen complex Euclidean $\mathcal{P}$ space over which the space of density operators is defined. The dimension of $D(\mathcal{P})$ is unknown as the dimension of Eve's subsystem E is not known on the standard assumptions of security analysis. Analysing the subsystem of Alice and Bob in a larger, unknown dimension is not only impractical but also complicates the analysis, as larger systems are inherently more susceptible to attacks due to their increased complexity. The work of [48] shows that without loss of generality one can can reduce the problem to two-qubit space $(\mathbb{C}_{4\times4})$ of Alice's and Bob's subsystem and from the result of [18] one can decompose a projector of higher dimensions into projectors of dimension (2×2) acting on either Alice or Bob subsystems space along with a commuting part.

Therefore, the four pairs of projectors as in (8.1.3) can be decomposed accordingly. Let L_S^z be the set of inter-commuting projectors where $S \in \{A, B\}$ denotes the subsystem of either Alice or Bob and $z \in \{0, 1, 2, 3\}$ corresponds to particular observables as stated earlier. Specifically $L_S^z = \{L_A^0, L_A^1, L_B^2, L_B^3\}$.
Now consider the following projector of dimension (2×2) parameterized by angle θ

$$Q(\theta) = \begin{pmatrix} \cos^2\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) \\ \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) & \sin^2\left(\frac{\theta}{2}\right) \end{pmatrix} \tag{8.3.1}$$

Now one can decompose the $P_{\mathcal{X}}^{0|0}, P_{\mathcal{X}}^{0|1}, P_{\mathcal{Y}}^{0|2}$ and $P_{\mathcal{Y}}^{0|3}$ as

$$\begin{aligned} P_{\mathcal{X}}^{0|0} &= \bigoplus_i Q(0) \oplus L_A^0 \\ P_{\mathcal{X}}^{0|1} &= \bigoplus_i Q(\phi_A^i) \oplus L_A^1 \\ P_{\mathcal{Y}}^{0|2} &= \bigoplus_j Q(0) \oplus L_B^2 \\ P_{\mathcal{Y}}^{0|3} &= \bigoplus_j Q(\phi_B^j) \oplus L_B^3 \end{aligned} \tag{8.3.2}$$

The remaining projectors can be obtained using the completeness property of the projectors, thus

$$\begin{aligned} P_{\mathcal{X}}^{1|0} &= \mathbb{1} - P_{\mathcal{X}}^{0|0} \\ P_{\mathcal{X}}^{1|1} &= \mathbb{1} - P_{\mathcal{X}}^{0|1} \\ P_{\mathcal{Y}}^{1|2} &= \mathbb{1} - P_{\mathcal{X}}^{0|2} \\ P_{\mathcal{Y}}^{1|3} &= \mathbb{1} - P_{\mathcal{X}}^{0|3} \end{aligned} \tag{8.3.3}$$

The angles $\{\phi_A^i\}$ and $\{\phi_B^j\}$ are obtained from *spectrum* of $P_{\mathcal{X}}^{0|0} + P_{\mathcal{X}}^{0|1}$ and $P_{\mathcal{Y}}^{0|2} + P_{\mathcal{Y}}^{0|3}$ respectively. These projectors are *block-diagonal operators* and are decomposed into noncommuting and commuting portions as

- Non-commuting 2×2 projectors: These contain the parameterized projectors $Q(\phi_A^i)$ or $Q(\phi_B^j)$.
- Commuting projectors $L_A^0, L_A^1, L_B^2, L_B^3$: These are fixed projectors with eigenvalues 0 or 1.

From the definition of *spectrum* of the sum of $P_{\mathcal{X}}^{0|0} + P_{\mathcal{X}}^{0|1}$ or $\left(P_{\mathcal{Y}}^{0|2} + P_{\mathcal{Y}}^{0|3}\right)$ is obtained from the union of eigenvalues from the 2×2 projector that is $1 \pm \cos(\phi_A^i/2)$ or $\left(1 \pm \cos(\phi_B^j/2)\right)$ and eigenvalues from the commuting parts are from the set $G = \{0, 1, 2\}$ where the eigenvalue of 2 occur if the space span by L_A^0 and L_A^1 or L_B^2 and L_B^3 overlap.
These commuting parts contribute *discrete, angle-independent eigenvalues*, while the angles ϕ_A^i, ϕ_B^j depends upon the 2×2 projectors. The Commuting Projectors *pairwise commuting* and this ensures their joint spectrum has eigenvalues $0, 1, 2$. The commuting projectors are *fixed* and do not depend on ϕ_A^i or ϕ_B^j, and their eigenvalues are independent of the angles. The eigenvalues from the 2×2 projector. They lie in $[0, 2]$ but are *not integers* unless ϕ_A^i or ϕ_B^j is 0 or π. They do not overlap with the eigenvalues $0, 1, 2$ from the commuting parts.
Thus, in deriving the angles ϕ_A^i, ϕ_B^j are uniquely determined by the non-integer eigenvalues in the spectrum, while the commuting parts contribute only integer eigenvalues that are irrelevant to the angles. Now one can compute the angle ϕ_A^i from the eigenvalues$\lambda_{\mathcal{X}}^k$ where k refer to the k^{th} eigenvalue of the operator $P_{\mathcal{X}}^{0|0} + P_{\mathcal{X}}^{0|1}$ using (8.3.1) and (8.3.2) as,

$$L = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + Q(\phi_A^i) = \begin{pmatrix} 1 + \cos^2(\phi_A^i/2) & \cos(\phi_A^i/2)\sin(\phi_A^i/2) \\ \cos(\phi_A^i/2)\sin(\phi_A^i/2) & \sin^2(\phi_A^i/2) \end{pmatrix} \tag{8.3.4}$$

On solving for the eigenvalue, one can get,

$$\begin{aligned}
&\det\left(P_{\mathcal{X}}^{0|0}+P_{\mathcal{X}}^{0|1}-\lambda\mathbb{1}\right)=0\\
&\text{or,}\det\left(\begin{pmatrix}1+\cos^2(\phi_A/2) & \cos(\phi_A^i/2)\sin(\phi_A^i/2)\\ \cos(\phi_A^i/2)\sin(\phi_A^i/2) & \sin^2(\phi_A^i/2)\end{pmatrix}-\lambda\cdot\begin{pmatrix}1&0\\0&1\end{pmatrix}\right)=0\\
&\text{or, }\det\left(\begin{pmatrix}1+\cos^2(\phi_A^i/2) & \cos(\phi_A^i/2)\sin(\phi_A^i/2)\\ \cos(\phi_A^i/2)\sin(\phi_A^i/2) & \sin^2(\phi_A^i/2)\end{pmatrix}-\begin{pmatrix}\lambda&0\\0&\lambda\end{pmatrix}\right)=0\\
&\text{or, }\det\begin{pmatrix}1+\cos^2(\phi_A^i/2)-\lambda & \cos(\phi_A^i/2)\sin(\phi_A^i/2)\\ \cos(\phi_A^i/2)\sin(\phi_A^i/2) & \sin^2(\phi_A^i/2)-\lambda\end{pmatrix}=0 \qquad (8.3.5)\\
&\text{or, }\left((1+\cos^2(\phi_A^i/2)-\lambda)(\sin^2(\phi_A^i/2)-\lambda)\right)-\left((\cos(\phi_A^i/2)\sin(\phi_A^i/2))^2\right)=0\\
&\text{or, }\lambda^2-\lambda(1+\cos^2(\phi_A^i/2)+\sin^2(\phi_A^i/2))+\sin^2(\phi_A^i/2)=0\\
&\text{or, }\lambda^2-2\lambda+\sin^2(\phi_A^i/2)=0\\
&\text{or, }\lambda=1\pm\cos(\phi_A^i/2).
\end{aligned}$$

Thus the eigenvalue of the operator $P_{\mathcal{X}}^{0|0}+P_{\mathcal{X}}^{0|1}$ are $\lambda_{\mathcal{X}}^1=1+\cos(\phi_A^i/2)$ and $\lambda_{\mathcal{X}}^2=1-\cos(\phi_A^i/2)$. On the similar note from the spectrum of $\left(P_{\mathcal{Y}}^{0|2}+P_{\mathcal{Y}}^{0|3}\right)$, the eigenvalues are $\lambda_{\mathcal{Y}}^1=1+\cos(\phi_B^j/2)$ and $\lambda_{\mathcal{Y}}^2=1-\cos(\phi_B^j/2)$

For each eigenvalue $\lambda_{\mathcal{X}}^1,\lambda_{\mathcal{X}}^2$ and $\lambda_{\mathcal{Y}}^1,\lambda_{\mathcal{Y}}^2$ one can have,

$$\boxed{\begin{aligned}\cos(\phi_A^i/2)=\frac{\lambda_{\mathcal{X}}^1-\lambda_{\mathcal{X}}^2}{2}\text{or, }\phi_A^i=2\arccos\left(\frac{\lambda_{\mathcal{X}}^1-\lambda_{\mathcal{X}}^2}{2}\right).\\ \cos(\phi_B^j/2)=\frac{\lambda_{\mathcal{Y}}^1-\lambda_{\mathcal{Y}}^2}{2}\text{or, }\phi_B^j=2\arccos\left(\frac{\lambda_{\mathcal{Y}}^1-\lambda_{\mathcal{Y}}^2}{2}\right).\end{aligned}} \qquad (8.3.6)$$

The mapping $\phi_N^m\mapsto 1\pm\cos(\phi_N^m/2)$ where $N\in\{A,B\}$ and $m\in\{i,j\}$is bijective for $\phi_N^m\in[0,2\pi)$ when restricted to non-integer eigenvalues. Moreover the inverse function arccos is ϕ_N^m unique in $[0,\pi]$.

Lemma 8.3.1. The mapping $\phi_N^m\mapsto 1\pm\cos(\phi_N^m/2)$ where $N\in\{A,B\}$ and $m\in\{i,j\}$ is bijective for $\phi_N^m\in[0,\pi]$

Proof. Let $h(\phi_N^m)=1\pm\cos(\phi_N^m/2)$ where $N\in\{A,B\}$ and $m\in\{i,j\}$
1. Proof of infectivity: Let us assume that $\phi_N^1\neq\phi_N^2\in[0,\pi]$ where ϕ_N^1,ϕ_N^2 are arbitrarily two angles obtained from spectrum of the sum of either $\left(P_{\mathcal{X}}^{0|0}+P_{\mathcal{X}}^{0|1}\right)$ or $\left(P_{\mathcal{Y}}^{0|2}+P_{\mathcal{Y}}^{0|3}\right)$

$$\cos(\phi_N^1/2)\neq\cos(\phi_N^2/2)\text{or, }\lambda_{\mathcal{X}}^1\neq\lambda_{\mathcal{X}}^2\quad\text{or}\quad(\lambda_{\mathcal{Y}}^1\neq\lambda_{\mathcal{Y}}^2)\text{or, }h(\phi_N^1)\neq h(\phi_N^2) \qquad (8.3.7)$$

Thus f is an injective function.

2. Proof of Surjectivity: The eigenvalues $\lambda_{\mathcal{X}}^1,\lambda_{\mathcal{X}}^2$ or $(\lambda_{\mathcal{Y}}^1,\lambda_{\mathcal{Y}}^2)\in[0,2]$ with $\lambda_{\mathcal{X}}^1+\lambda_{\mathcal{X}}^2=2$ or $(\lambda_{\mathcal{Y}}^1+\lambda_{\mathcal{Y}}^2)$ and $\lambda_{\mathcal{X}}^1>\lambda_{\mathcal{X}}^2$ or $(\lambda_{\mathcal{Y}}^1>\lambda_{\mathcal{Y}}^2)$ there exists a unique $\phi_N^m\in[0,\pi]$ where $m\in\{i,j\}$ such that the eigenvalues corresponding to ϕ_N^m are $\lambda_{\mathcal{X}}^1,\lambda_{\mathcal{X}}^2$ or $(\lambda_{\mathcal{Y}}^1,\lambda_{\mathcal{Y}}^2)$ for $N=A$ or B respectively.
The condition of $\lambda_{\mathcal{X}}^1+\lambda_{\mathcal{X}}^2=2$ or $(\lambda_{\mathcal{Y}}^1+\lambda_{\mathcal{Y}}^2)$ arises from a fundamental result of matrix algebra, *trace of a matrix is the sum of its eigenvalues* and holding this result

$$\begin{aligned}\mathrm{Tr}(L)&=1+\cos^2(\phi_A^i/2)+\sin^2(\phi_A^i/2)\\&=2\\&=\lambda_{\mathcal{X}}^1+\lambda_{\mathcal{X}}^2\end{aligned} \qquad (8.3.8)$$

and equivalently

$$\lambda^1_{\mathcal{Y}} + \lambda^2_{\mathcal{Y}} = 2 \tag{8.3.9}$$

Given the above three criteria, one can have, let

$$\begin{aligned} c_{\mathcal{X}} &= \frac{\lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}}}{2} \\ c_{\mathcal{Y}} &= \frac{\lambda^1_{\mathcal{Y}} - \lambda^2_{\mathcal{Y}}}{2} \end{aligned} \tag{8.3.10}$$

Compute $c_{\mathcal{X}}$: Since $\lambda^1_{\mathcal{X}} > \lambda^2_{\mathcal{X}}$ and $\lambda^1_{\mathcal{X}}, \lambda^2_{\mathcal{X}} \in [0,2]$, one can have $\lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}} > 0$. Also $\lambda^1_{\mathcal{X}} + \lambda^2_{\mathcal{X}}$, one can have $\lambda^1_{\mathcal{X}} = 2 - \lambda^2_{\mathcal{X}}$. Thus,

$$c_{\mathcal{X}} = \frac{2 - \lambda^2_{\mathcal{X}} - \lambda^2_{\mathcal{X}}}{2} = \frac{2 - 2\lambda^2_{\mathcal{X}}}{2} = 1 - \lambda^2_{\mathcal{X}} \tag{8.3.11}$$

Since $\lambda^1_{\mathcal{X}} > \lambda^2_{\mathcal{X}}$ and $\lambda^1_{\mathcal{X}} + \lambda^2_{\mathcal{X}} = 2$, we have $2\lambda^2_{\mathcal{X}} < 2$, so $\lambda^2_{\mathcal{X}} < 1$. Therefore, $c_{\mathcal{X}} = 1 - \lambda^2_{\mathcal{X}} > 0$. Also, since $\lambda^1_{\mathcal{X}} < 2$, one can have $\lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}} < 2 - \lambda^2_{\mathcal{X}}$. Since $\lambda^2_{\mathcal{X}} \geq 0$, $\lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}} < 2$. Thus $c_{\mathcal{X}} < 1$. Therefore, $c \in (0,1)$. *Solve for the given angle ϕ^i_A:* Now using (8.3.6) and (8.3.10) one can have,

$$\phi_A = 2\arccos(c_{\mathcal{X}}) \tag{8.3.12}$$

Since $c_{\mathcal{X}} \in (0,1)$, $\arccos(c_{\mathcal{X}}) \in (0, \frac{\pi}{2})$. Therefore, $\phi^i_A = 2\arccos(c_{\mathcal{X}}) \in (0,\pi)$. *3. Uniqueness:* The function $\arccos(c_{\mathcal{X}})$ is *strictly decreasing* on the interval $[0,1]$. Therefore, for each $c_{\mathcal{X}} \in (0,1)$, there is a unique $\arccos(c_{\mathcal{X}}) \in (0, \frac{\pi}{2})$. Consequently, there is a unique $\phi = 2\arccos(c_{\mathcal{X}}) \in (0,\pi)$. *4. Eigenvalues:* The eigenvalues are given by,

$$\begin{aligned} \lambda^1_{\mathcal{X}}, \lambda^2_{\mathcal{X}} &= 1 \pm \cos(\phi^i_A/2) \\ \lambda^1_{\mathcal{Y}}, \lambda^2_{\mathcal{Y}} &= 1 \pm \cos(\phi^j_B/2) \end{aligned} \tag{8.3.13}$$

Substituting $\phi^i_A = 2\arccos(c_{\mathcal{X}})$, one can get:

$$\lambda^1_{\mathcal{X}}, \lambda^2_{\mathcal{X}} = 1 \pm \cos(\arccos(c_{\mathcal{X}})) = 1 \pm c_{\mathcal{X}} \tag{8.3.14}$$

Thus, $\lambda^1_{\mathcal{X}} = 1 + c_{\mathcal{X}}$ and $\lambda^2_{\mathcal{X}} = 1 - c_{\mathcal{X}}$. Substituting $c_{\mathcal{X}} = \frac{\lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}}}{2}$, one gets,

$$\begin{aligned} \lambda^1_{\mathcal{X}} &= 1 + \frac{\lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}}}{2} \text{or, } 2\lambda^1_{\mathcal{X}} = 2 + \lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}} \text{or, } \lambda^1_{\mathcal{X}} = 2 - \lambda^2_{\mathcal{X}} \\ \lambda^2_{\mathcal{X}} &= 1 - \frac{\lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}}}{2} \text{or, } 2\lambda^2_{\mathcal{X}} = 2 - \lambda^1_{\mathcal{X}} + \lambda^2_{\mathcal{X}} \text{or, } \lambda^2_{\mathcal{X}} = 2 - \lambda^1_{\mathcal{X}} \end{aligned} \tag{8.3.15}$$

Therefore, for any non-integer pair $\lambda^1_{\mathcal{X}}, \lambda^2_{\mathcal{X}} \in [0,2]$ with $\lambda^1_{\mathcal{X}} + \lambda^2_{\mathcal{X}} = 2$ and $\lambda^1_{\mathcal{X}} > \lambda^2_{\mathcal{X}}$, there exists a unique $\phi_A \in (0,\pi)$ such that the eigenvalues corresponding to ϕ^i_A are $\lambda^1_{\mathcal{X}}$ and $\lambda^2_{\mathcal{X}}$.
Similarly by substituting $\phi^j_B = 2arccos(c_{\mathcal{Y}})$ in (8.3.13) This proves subjectivity.Thus, the mapping is bijective. □

In quantum information theory, there are two types of correlations: Classical correlations and quantum correlations. They are sometimes referred to as trivial and non-trivial correlations, respectively.

8.3.1 Classifying the Projectors

Lemma 8.3.2. Projector pair $(P^{0|0}_{\mathcal{X}}, P^{0|2}_{\mathcal{Y}})$ is a commuting pair and $(P^{0|1}_{\mathcal{X}}, P^{0|3}_{\mathcal{Y}})$ is a non-commuting pair. Thus $[P^{0|0}_{\mathcal{X}}, P^{0|2}_{\mathcal{Y}}] = 0$ and $[P^{0|1}_{\mathcal{X}}, P^{0|3}_{\mathcal{Y}}] \neq 0$.

Proof. *1. Prove* $[P_{\mathcal{X}}^{0|0}, P_{\mathcal{Y}}^{0|2}] = 0$

$$\begin{aligned}
[P_{\mathcal{X}}^{0|0}, P_{\mathcal{Y}}^{0|2}] &= [Q(0) \oplus L_A^0, Q(0) \oplus L_B^2] \\
&= (Q(0) \oplus L_A^0) \cdot (Q(0) \oplus L_B^2) - (Q(0) \oplus L_B^2) \cdot (Q(0) \oplus L_A^0) \\
&= \begin{pmatrix} Q(0) & 0 \\ 0 & L_A^0 \end{pmatrix} \cdot \begin{pmatrix} Q(0) & 0 \\ 0 & L_B^2 \end{pmatrix} - \begin{pmatrix} Q(0) & 0 \\ 0 & L_B^2 \end{pmatrix} \cdot \begin{pmatrix} Q(0) & 0 \\ 0 & L_A^0 \end{pmatrix} \\
&= \begin{pmatrix} Q(0) \cdot Q(0) & 0 \\ 0 & L_A^0 \cdot L_B^2 \end{pmatrix} - \begin{pmatrix} Q(0) \cdot Q(0) & 0 \\ 0 & L_B^2 \cdot L_A^0 \end{pmatrix} \\
&= \begin{pmatrix} Q(0) & 0 \\ 0 & L_A^0 \cdot L_B^2 \end{pmatrix} - \begin{pmatrix} Q(0) & 0 \\ 0 & L_B^2 \cdot L_A^0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 \\ 0 & [L_A^0, L_B^2] \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\
&= 0
\end{aligned} \tag{8.3.16}$$

2. Prove $[P_{\mathcal{X}}^{0|1}, P_{\mathcal{Y}}^{0|3}] \neq 0$

$$\begin{aligned}
[P_{\mathcal{X}}^{0|1}, P_{\mathcal{Y}}^{0|3}] &= [Q(\phi_A^i) \oplus L_A^1, Q(\phi_B^j) \oplus L_B^3] \\
&= (Q(\phi_A^i) \oplus L_A^1) \cdot (Q(\phi_B^j) \oplus L_B^3) - (Q(\phi_B^j) \oplus L_B^3) \cdot (Q(\phi_A^i) \oplus L_A^1) \\
&= \begin{pmatrix} Q(\phi_A^i) & 0 \\ 0 & L_A^1 \end{pmatrix} \cdot \begin{pmatrix} Q(\phi_B^j) & 0 \\ 0 & L_B^3 \end{pmatrix} - \begin{pmatrix} Q(\phi_B^j) & 0 \\ 0 & L_B^3 \end{pmatrix} \cdot \begin{pmatrix} Q(\phi_A^i) & 0 \\ 0 & L_A^1 \end{pmatrix} \\
&= \begin{pmatrix} Q(\phi_A^i) \cdot Q(\phi_B^j) & 0 \\ 0 & L_A^1 \cdot L_B^3 \end{pmatrix} - \begin{pmatrix} Q(\phi_B^j) \cdot Q(\phi_A^j) & 0 \\ 0 & L_B^3 \cdot L_A^1 \end{pmatrix} \\
&= \begin{pmatrix} Q(\phi_A^i) \cdot Q(\phi_B^j) - Q(\phi_B^j) \cdot Q(\phi_A^i) & 0 \\ 0 & L_A^1 \cdot L_B^3 - L_B^3 \cdot L_A^1 \end{pmatrix} \\
&= \begin{pmatrix} Q(\phi_A^i) \cdot Q(\phi_B^j) - Q(\phi_B^i) \cdot Q(\phi_A^i) & 0 \\ 0 & [L_A^1, L_B^3] \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2} \sin(\phi_A^i + \phi_B^j) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix} \\
&\neq 0
\end{aligned} \tag{8.3.17}$$

□

Corollary 8.3.3.

$$[P_{\mathcal{X}}^{1|0}, P_{\mathcal{Y}}^{1|2}] = 0, [P_{\mathcal{X}}^{1|1}, P_{\mathcal{Y}}^{1|3}] \neq 0 \tag{8.3.18}$$

Lemma 8.3.4. $(P_{\mathcal{X}}^{0|0}, P_{\mathcal{Y}}^{0|2})$ exhibits classical correlations and $(P_{\mathcal{X}}^{0|1}, P_{\mathcal{Y}}^{0|3})$ exhibits quantum correlations. Thus,

a.Show that $P(a = 0, b = 0|x = 0, y = 2) = \mathrm{Tr}(\rho\, P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{0|2}) = \sum_i p_i P_i(0|0) P_i(0|2)$

b.Show that $P(a = 0, b = 0|x = 1, y = 3) = \mathrm{Tr}(\rho\, P_{\mathcal{X}}^{0|1} \otimes P_{\mathcal{Y}}^{0|3}) = \sum_{i,j} \rho_{ij} \langle |i\rangle\langle j|, Q(\phi_A^i)\rangle \langle |i\rangle\langle j|, Q(\phi_B^j)\rangle$

Proof. a. $P(a=0,b=0|x=0,y=2) = \mathrm{Tr}(\rho\, P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{0|2}) = \sum_i p_i P_i(0|0)P_i(0|2)$
The joint probability of outcomes $a=0$ and $b=0$ for measurements $x=0$ and $y=2$ on a bipartite quantum state ρ is given (8.1.4) as

$$P(a=0,b=0|x=0,y=2) = \mathrm{Tr}\left(\rho\, \left(P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{0|2}\right)\right), \tag{8.3.19}$$

where $P_{\mathcal{X}}^{0|0}$ and $P_{\mathcal{Y}}^{0|2}$ are projection operators on subsystems $\mathcal{X}$ and $\mathcal{Y}$, respectively. Now, since $[P_{\mathcal{X}}^{0|0}, P_{\mathcal{Y}}^{0|2}] = 0$ one can have, a common eigenbasis $\{|\psi_i\rangle\}$ for both projectors.
The state ρ can be expressed in this eigenbasis as:

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \tag{8.3.20}$$

where $p_i \geq 0$ and $\sum_i p_i = 1$.
Substituting ρ into the trace expression:

$$\mathrm{Tr}\left(\rho\, \left(P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{0|2}\right)\right) = \sum_i p_i \mathrm{Tr}\left(|\psi_i\rangle\langle\psi_i|\, \left(P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{0|2}\right)\right) \tag{8.3.21}$$

Since $|\psi_i\rangle$ is a simultaneous eigenstate of both $P_{\mathcal{X}}^{0|0}$ and $P_{\mathcal{Y}}^{0|2}$, one can can express it as a tensor product of eigenstates: $|\psi_i\rangle = |\alpha_i\rangle \otimes |\beta_i\rangle$, where $|\alpha_i\rangle \in \mathcal{X}$ and $|\beta_i\rangle \in \mathcal{Y}$. Then, one can have:

$$\left(P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{0|2}\right)|\psi_i\rangle = \left(P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{0|2}\right)(|\alpha_i\rangle \otimes |\beta_i\rangle) = \left(P_{\mathcal{X}}^{0|0}|\alpha_i\rangle\right) \otimes \left(P_{\mathcal{Y}}^{0|2}|\beta_i\rangle\right). \tag{8.3.22}$$

Since $|\alpha_i\rangle$ and $|\beta_i\rangle$ are eigenstates of $P_{\mathcal{X}}^{0|0}$ and $P_{\mathcal{Y}}^{0|2}$ respectively, one can have:

$$\begin{aligned} P_{\mathcal{X}}^{0|0}|\alpha_i\rangle &= \delta_{0,\alpha_i}|\alpha_i\rangle = P_i(0|0)|\alpha_i\rangle, \\ P_{\mathcal{Y}}^{0|2}|\beta_i\rangle &= \delta_{0,\beta_i}|\beta_i\rangle = P_i(0|2)|\beta_i\rangle, \end{aligned} \tag{8.3.23}$$

where δ_{0,α_i} and δ_{0,β_i} are Kronecker deltas, and α_i and β_i are the eigenvalues corresponding to the eigenstates $|\alpha_i\rangle$ and $|\beta_i\rangle$ respectively. $P_i(0|0) = \delta_{0,\alpha_i}$ and $P_i(0|2) = \delta_{0,\beta_i}$ are indicator functions, taking values in $\{0,1\}$.

Therefore, one can rewrite the equation as:

$$\begin{aligned} \left(P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{0|2}\right)|\psi_i\rangle &= (P_i(0|0)|\alpha_i\rangle) \otimes (P_i(0|2)|\beta_i\rangle) \\ &= P_i(0|0)P_i(0|2)(|\alpha_i\rangle \otimes |\beta_i\rangle) \\ &= P_i(0|0)P_i(0|2)|\psi_i\rangle \end{aligned} \tag{8.3.24}$$

where $P_i(0|0), P_i(0|2) \in \{0,1\}$ are indicator functions for outcomes $a=0$ and $b=0$ in state $|\psi_i\rangle$.
The trace simplifies to:

$$\begin{aligned} \mathrm{Tr}\left(|\psi_i\rangle\langle\psi_i|\, \left(P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{0|2}\right)\right) &= \mathrm{Tr}\left(\langle\psi_i|\, \left(P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{0|2}\right)|\psi_i\rangle\right) \\ &= \mathrm{Tr}\left(\langle\psi_i|\, P_i(0|0)P_i(0|2)|\psi_i\rangle\right) \\ &= P_i(0|0)P_i(0|2)\mathrm{Tr}\left(\langle\psi_i|\psi_i\rangle\right) \\ &= P_i(0|0)P_i(0|2)\langle\psi_i|\psi_i\rangle \\ &= P_i(0|0)P_i(0|2)\cdot 1 \\ &= P_i(0|0)P_i(0|2) \end{aligned} \tag{8.3.25}$$

Now, from (8.3.25),(8.3.21)

$$P(a=0,b=0|x=0,y=2) = \mathrm{Tr}(\rho\, P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{0|2}) = \sum_i p_i P_i(0|0)P_i(0|2). \tag{8.3.26}$$

Thus for $(P_{\mathcal{X}}^{0|0}, P_{\mathcal{Y}}^{0|2})$ the probability of $P(a = 0, b = 0|x = 0, y = 2)$ can be factorize into local probability distributions $P_i(0|0)$ and $P_i(0|2)$. So from joint probability distribution one can say that $(P_{\mathcal{X}}^{0|0}, P_{\mathcal{Y}}^{0|2})$ exhibit classical correlations.

Corollary 8.3.5.

$$P(a = 1, b = 1|x = 0, y = 2) = \mathrm{Tr}(\rho\, P_{\mathcal{X}}^{1|0} \otimes P_{\mathcal{Y}}^{1|2}) = \sum_i p_i P_i(1|0) P_i(1|2) \tag{8.3.27}$$

Thus $(P_{\mathcal{X}}^{1|0}, P_{\mathcal{Y}}^{1|2})$ exhibit classical correlations.

b. $P(a = 0, b = 0|x = 1, y = 3) = \mathrm{Tr}(\rho\, P_{\mathcal{X}}^{0|1} \otimes P_{\mathcal{Y}}^{0|3}) = \sum_{i,j} \rho_{ij} \langle |i\rangle\langle j|, Q(\phi_A^i)\rangle \langle |i\rangle\langle j|, Q(\phi_B^j)\rangle$ The joint probability of outcomes a and b for measurements $x = 1$ and $y = 3$ on a bipartite quantum state ρ is given by the Born rule:

$$P(a = 0, b = 0|x = 1, y = 3) = \mathrm{Tr}\left(\rho\, \left(P_{\mathcal{X}}^{0|1} \otimes P_{\mathcal{Y}}^{0|3}\right)\right), \tag{8.3.28}$$

where $P_{\mathcal{X}}^{0|1} = Q(\phi_A^i) \oplus L_A^1$ and $P_{\mathcal{Y}}^{0|3} = Q(\phi_B^j) \oplus L_B^3$ are projection operators parameterized by angles ϕ_A^i, ϕ_B^j. one can will only consider non-commuting projector $\left(Q(\phi_A^i), Q(\phi_B^j)\right)$, assuming the commuting projectors$\left(L_A^1, L_B^3\right)$ contributes classically as for commuting projector the joint conditional probability of outcomes given the projector can be factorized into local probability distributions.
Substituting ρ and the projectors into the trace expression, one can have:

$$\begin{aligned} P(a = 0, b = 0|x = 1, y = 3) &= \mathrm{Tr}\left(\left(\sum_{i,j} \rho_{ij} |i\rangle\langle j| \otimes |i\rangle\langle j|\right)\left(Q(\phi_A^i) \otimes Q(\phi_B^j)\right)\right) \\ &= \sum_{i,j} \rho_{ij} \mathrm{Tr}\left((|i\rangle\langle j| \otimes |i\rangle\langle j|)\left(Q(\phi_A^i) \otimes Q(\phi_B^j)\right)\right) \\ &= \sum_{i,j} \rho_{ij} \mathrm{Tr}\left(|i\rangle\langle j| Q(\phi_A^i) \otimes |i\rangle\langle j| Q(\phi_B^j)\right) \\ &= \sum_{i,j} \rho_{ij} \mathrm{Tr}\left(|i\rangle\langle j| Q(\phi_A^i)\right) \mathrm{Tr}\left(|i\rangle\langle j| Q(\phi_B^j)\right). \end{aligned} \tag{8.3.29}$$

Now, using the cyclic property of the trace, one can have:

$$\begin{aligned} \mathrm{Tr}\left(|i\rangle\langle j| Q(\phi_A^i)\right) &= \mathrm{Tr}\left(|i\rangle\langle j| Q(\phi_A^i)^*\right) = \langle |i\rangle\langle j|, Q(\phi_A^i)\rangle, \\ \mathrm{Tr}\left(|i\rangle\langle j| Q(\phi_B^j)\right) &= \mathrm{Tr}\left(|i\rangle\langle j| Q(\phi_B^j)^*\right) = \langle |i\rangle\langle j|, Q(\phi_B^j)\rangle \end{aligned} \tag{8.3.30}$$

Substituting these results back into the expression for $P(a, b|x = 1, y = 3)$, one can obtain:

$$\begin{aligned} P(a = 0, b = 0|x = 1, y = 3) &= \mathrm{Tr}(\rho\, P_{\mathcal{X}}^{0|1} \otimes P_{\mathcal{Y}}^{0|3}) \\ &= \sum_{i,j} \rho_{ij} \langle |i\rangle\langle j|, Q(\phi_A^i)\rangle \langle |i\rangle\langle j|, Q(\phi_B^j)\rangle \end{aligned} \tag{8.3.31}$$

where $\langle |i\rangle\langle j|, Q(\phi_A^i)\rangle$and$\langle |i\rangle\langle j|, Q(\phi_B^j)\rangle$ are outcomes a and b when measuring $Q(\phi_A^i)$ and $Q(\phi_B^j)$ on states $|i\rangle$ and $|j\rangle$, respectively.

The terms $\langle |i\rangle\langle j|, Q(\phi_A^i)\rangle$and$\langle |i\rangle\langle j|, Q(\phi_B^j)\rangle$ introduce quantum interference terms leading to quantum-correlations except when $Q(\phi_A^i) = \pi$ or $Q(\phi_B^j) = 0$. Thus $P_{\mathcal{X}}^{0|1}$ and $P_{\mathcal{Y}}^{0|3}$ introduce quantumcorrelation.

Corollary 8.3.6.

$$\begin{aligned} P(a=1,b=1|x=1,y=3) &= \mathrm{Tr}(\rho\, P_{\mathcal{X}}^{1|1} \otimes P_{\mathcal{Y}}^{1|3}) \\ &= \sum_{i,j} \rho_{ij} \langle |i\rangle\langle j|, Q(\phi_A^i)\rangle \langle |i\rangle\langle j|, Q(\phi_B^j)\rangle \end{aligned} \tag{8.3.32}$$

Thus, $(P_{\mathcal{X}}^{1|1}, P_{\mathcal{Y}}^{1|3})$ exhibit quantum correlation.

Summing up,

$$\boxed{\begin{array}{l} \left(P_{\mathcal{X}}^{1|0}, P_{\mathcal{Y}}^{1|2}\right) \text{ exhibit classical correlations.} \\ \left(P_{\mathcal{X}}^{1|1}, P_{\mathcal{Y}}^{1|3}\right) \text{ exhibit quantum correlation.} \end{array}} \tag{8.3.33}$$

□

8.3.2 Lower bounding $C^*(S)$ by two qubits vector space ($\mathbb{C}_{4\times 4}$)

Having established the functional relationship between the angles ϕ_A^i and ϕ_B^j and characterized the spectral properties of the operators $P_{\mathcal{X}}^{0|0} + P_{\mathcal{X}}^{0|1}$ (or equivalently, $P_{\mathcal{Y}}^{0|2} + P_{\mathcal{Y}}^{0|3}$), one can now proceed to develop a rigorous lower bound for the connected correlation function, denoted as $C^*(S)$. This development builds upon the prior categorisation of the projectors into groups exhibiting classical or quantum correlations, providing essential context for the subsequent analysis. Specifically, spectral analysis allows us to relate the eigenvalues of these operators to the correlation structure, enabling the derivation of a concrete lower bound on $C^*(S)$.

Let $\Lambda_{\mathcal{X}}$ and $\Lambda_{\mathcal{Y}}$ be channels acting on Alice's and Bob's subsystems obtained using partial trace of $\rho_{ABEA'_x}$ respectively, which decompose Alice's and Bob's subsystems into block structures as would be obtained from effect of projectors in (8.3.2). The channels essentially decompose Alice and Bob subsystems into complex euclidean spaces describing each one's particle(or equivalently their qubits) for corresponding value of ϕ_A^i and ϕ_B^j and a commuting block obtain by the pair wise commuting projectors $\{L_A^0, L_A^1, L_B^2, L_B^3\}$ Now let $\Lambda_{\mathcal{X}\mathcal{Y}} = \Lambda_{\mathcal{X}} \otimes \Lambda_{\mathcal{Y}}$ be a channel acting on $\rho_{ABEA'_x}$, which decomposes it into the corresponding complex Euclidean spaces describing each of Alice's and Bob's pairs (or equivalently, the shared state) corresponding to angles ϕ_A^i and ϕ_B^j obtained from their respective spectra, and a commuting part obtained from either $L_A^0 \otimes L_A^1$ or $L_B^2 \otimes L_B^3$.

Let P be the set of all states that can be obtained from the operator $Q(\phi_A^i) \otimes Q(\phi_B^j)$ acting on the state $\rho_{ABEA'_x}$.

$$P = \{\rho_{ABEA'_x}^{ij} \mid Q(\phi_A^i) \otimes Q(\phi_B^j)(\rho_{ABEA'_x}) = \rho_{ABEA'_x}^{ij}\} \tag{8.3.34}$$

Thus,

$$\Lambda_{\mathcal{X}\mathcal{Y}}[\rho_{ABEA'_x}] = \bigoplus_{ij} (\eta_{ij} \rho_{ABEA'_x}^{ij}) \oplus (\eta_{\text{commuting}} \rho_{ABEA'_x}^{\text{commuting}}) \tag{8.3.35}$$

where η_{ij} are normalization factors obtained by normalizing each state $\rho_{ABEA'_x}^{ij} \in P$, and $\rho_{ABEA'_x}^{ij} \in \mathcal{D}(\mathcal{G})$ are density operators over some complex Euclidean space $\mathcal{G}$ describing the quantum states obtained after applying $\Lambda_{\mathcal{X}\mathcal{Y}}$ on $\rho_{ABEA'_x}$. Each of the states depends on the operator $Q(\phi_A^i) \otimes Q(\phi_B^j)$ and consequently on the angles obtained from respective spectra. The $\rho_{ABEA'_x}^{\text{commuting}}$ are the projected blocks that commute with either Alice's subsystem or Bob's subsystem.

Now, from the monotonicity of relative entropy,

$$\begin{aligned}
C^*(S) = \inf \quad & \lambda D\left((\rho^{AB}_{ABEA'_0}) \,||\, (\Lambda_0[\rho_{ABEA'_0}]^{A'B})\right) \\
& + (1-\lambda) D\left((\rho^{AB}_{ABEA'_1}) \,||\, (\Lambda_1[\rho_{ABEA'_1}]^{A'B})\right) \\
& \geq \lambda D\left((\Lambda_{\mathcal{X}\mathcal{Y}}[\rho^{AB}_{ABEA'_0}]) \,||\, (\Lambda_0 \circ \Lambda_{\mathcal{X}\mathcal{Y}}[\rho_{ABEA'_0}]^{A'B})\right) \\
& + (1-\lambda) D\left((\Lambda_{\mathcal{X}\mathcal{Y}}[\rho^{AB}_{ABEA'_1}]) \,||\, (\Lambda_1 \circ \Lambda_{\mathcal{X}\mathcal{Y}}[\rho_{ABEA'_1}]^{A'B})\right) \\
& \geq \sum_{ij} \eta_{ij} \left[\lambda D\left(\left(\rho^{ij,AB}_{ABEA'_0}\right) \,||\, \left(\Lambda_0[\rho^{ij}_{ABEA'_0}]^{A'B}\right)\right)\right. \\
& \qquad \left. +(1-\lambda) D\left(\left(\rho^{ij,AB}_{ABEA'_1}\right) \,||\, \left(\Lambda_1[\rho^{ij}_{ABEA'_1}]^{A'B}\right)\right)\right] \\
& + \eta_{\text{commuting}} \left[\lambda D\left(\left(\rho^{\text{commuting},AB}_{ABEA'_0}\right) \,||\, \left(\Lambda_0[\rho^{\text{commuting}}_{ABEA'_0}]^{A'B}\right)\right)\right. \\
& \qquad \left. +(1-\lambda) D\left(\left(\rho^{\text{commuting},AB}_{ABEA'_1}\right) \,||\, \left(\Lambda_1[\rho^{\text{commuting}}_{ABEA'_1}]^{A'B}\right)\right)\right] \\
& \mathrm{Tr}(\rho_{AB} \cdot \text{CHSH}) = S \\
& \rho_{AB} \succeq 0 \\
& \mathrm{Tr}(\rho_{AB}) = 1
\end{aligned} \tag{8.3.36}$$

where, $\rho^{ij,AB}_{ABEA'_x} = Tr_{EA_x}(\rho^{ij}_{ABEA'_x})$ and $\Lambda[\rho^{ij}_{ABEA'_x}]^{A'B} = Tr_{EA_x}(\Lambda[\rho^{ij}_{ABEA'_x}])$

$\rho^{\text{commuting},AB}_{ABEA'_x} = Tr_{EA_x}(\rho^{\text{commuting}}_{ABEA'_x})$ and $\Lambda[\rho^{\text{commuting}}_{ABEA'_x}]^{A'B} = Tr_{EA_x}(\Lambda[\rho^{\text{commuting}}_{ABEA'_x}])$

Since commuting projectors exhibit classical correlations, and can be factorized into local probabilities, as evidenced by equations (8.3.16),(8.3.18),(8.3.26) and (8.3.27) one can can similarly analyze the state $\rho^{\text{commuting}}_{ABEA'_x}$, where $x \in \{0,1\}$ represents the two outputs. Due to the classical nature of these correlations, one can can disregard the $\rho^{\text{commuting}}_{ABEA'_x}$ contribution, as $\mathrm{Tr}(\rho^{\text{commuting}}_{AB}, \text{CHSH}) \leq 2$, directly reflecting the bounds imposed by classical correlations.

Now, from the constraint of the optimization problem in (8.2.26), one can establish the CHSH value for each state $\rho^{ij}_{ABEA'_x} \in \mathcal{P}$, obtained from the operator $Q(\phi^i_A) \otimes Q(\phi^j_B)$ acting on $\rho_{ABEA'_x}$ as,

$$\begin{aligned}
& \mathrm{Tr}(\rho^{AB}_{ABEA'_x} \cdot \text{CHSH}) = S \\
\Leftrightarrow \quad & \mathrm{Tr}(\rho^{ij,AB}_{ABEA'_x} \cdot \text{CHSH}) = S_{ij} \quad \text{and} \quad \sum_{ij} \eta_{ij} S_{ij} = S
\end{aligned} \tag{8.3.37}$$

Having established the constraint for each state $\rho^{ij}_{ABEA'_x} \in \mathcal{P}$, one can optimize each of the above states as,

$$\begin{aligned}
C^*(S) \geq \inf \quad & \sum_{ij} \eta_{ij} \left[\lambda D\left(\left(\rho^{ij,AB}_{ABEA'_0}\right) \,||\, \left(\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B}\right)\right)\right. \\
& \qquad \left. +(1-\lambda) D\left(\left(\rho^{ij,AB}_{ABEA'_1}\right) \,||\, \left(\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B}\right)\right)\right] \\
\text{s.t.} \quad & \mathrm{Tr}((\rho^{ij,AB}_{ABEA'_x} \cdot \text{CHSH}) = S_{ij} \\
& \rho^{ij,AB}_{ABEA'_x} \succeq 0 \\
& \mathrm{Tr}(\rho^{ij,AB}_{ABEA'_x}) = 1 \\
& \sum_{ij} \eta_{ij} \leq 1 \\
& \sum_{ij} \eta_{ij} S_{ij} = S
\end{aligned} \tag{8.3.38}$$

Our entire system before and after the effect of the pinching channel, as defined in (8.2.2), is described by the states $\rho_{ABEA'_x}$ and $\rho_{A'BEA_x}$, respectively. Now, the state $\rho_{ABEA'_x}$ is first passed through the channel $\Lambda_{\mathcal{X}\mathcal{Y}}$, which essentially decomposes the state into blocks along the principal diagonal as described in (8.3.36), before passing it through the pinching channel. So, our objective function in (8.3.38) is essentially lower bounded by the minimum values in each such block. Thus,

$$\begin{aligned}
&\Big[\lambda D\left(\left(\rho^{ij,AB}_{ABEA'_0}\right) \,||\, \left(\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B}\right)\right)\\
&\qquad +(1-\lambda)D\left(\left(\rho^{ij,AB}_{ABEA'_1}\right) \,||\, \left(\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B}\right)\right)\Big]\\
&\geq C^*_{\mathbb{C}^{4\times 4}}(S_{ij})
\end{aligned} \tag{8.3.39}$$

where,

$$\begin{aligned}
C^*_{\mathbb{C}^{4\times 4}}(S_{ij}) = \inf \quad &\Big[\lambda D\left(\left(\rho^{ij,AB}_{ABEA'_0}\right) \,||\, \left(\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B}\right)\right)\\
&\qquad +(1-\lambda)D\left(\left(\rho^{ij,AB}_{ABEA'_1}\right) \,||\, \left(\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B}\right)\right)\Big]\\
\text{s.t.} \quad &\mathrm{Tr}(\rho^{ij,AB}_{ABEA'_x} \cdot \mathrm{CHSH}) = S_{ij}\\
&\rho^{ij,AB}_{ABEA'_x} \succeq 0\\
&\mathrm{Tr}(\rho^{ij,AB}_{ABEA'_x}) = 1
\end{aligned} \tag{8.3.40}$$

Thus from (8.3.38),(8.3.39) and (8.3.40) one can lower bound the function $C^*(S)$ as a function C^* over complex euclidean space describing two qubits $C^*_{\mathbb{C}^{4\times 4}}(S_{ij})$ as,

$$\begin{aligned}
C^*(S) \geq \sum_{ij} \eta_{ij} \quad &\inf \quad \left(C^*_{\mathbb{C}^{4\times 4}}(S_{ij})\right)\\
\text{s.t.} \quad &\mathrm{Tr}(\rho^{ij,AB}_{ABEA'_x} \cdot \mathrm{CHSH}) = S_{ij}\\
&\rho^{ij,AB}_{ABEA'_x} \succeq 0\\
&\mathrm{Tr}(\rho^{ij,AB}_{ABEA'_x}) = 1\\
&\sum_{ij} \eta_{ij} \leq 1\\
&\sum_{ij} \eta_{ij} S_{ij} = S
\end{aligned} \tag{8.3.41}$$

Now (8.3.41) is independent of the angles ϕ^i_A and ϕ^j_B as each states obtained from (8.3.34) are already optimized in (8.3.40), one can reduce the optimization of $C^*(S)$ depending on both the angles and normalizing weights of each states η_{ij} to optimizing only over to η_{ij} as,

$$\boxed{\begin{aligned}
C^*(S) &\geq \int_{S'=2}^{2\sqrt{2}} C^*_{\mathbb{C}^{4\times 4}}(S') \cdot \eta(dS')\\
&\geq \int_{S'=2}^{2\sqrt{2}} \eta(dS') \cdot C^*_{\mathbb{C}^{4\times 4}}(S')\\
s.t \quad &\eta([2, 2\sqrt{2}]) \leq 1\\
&\eta \geq 0\\
&\int_{S'=2}^{2\sqrt{2}} \eta(dS') S' = S
\end{aligned}} \tag{8.3.42}$$

Here, one is essentially taking one single state or block, integrating over positive sub-normalized weights η in the interval $S' = (2, 2\sqrt{2}]$ and

$$C^*_{\mathbb{C}^{4\times 4}}(S_{ij}) = 0 \quad \forall \quad S_{ij} \leq 0 \tag{8.3.43}$$

8.3.3 Reformulating the CHSH operator in explicit matrix form

Having established the optimization problem by reducing the complex euclidean space describing the entire system of Alice and Bob in some space of unknown dimension as evident from (8.2.26) to a problem having a complex euclidean space describing two qubits space ($\mathbb{C}_{4\times 4}$) of Alice and Bob as given in (8.3.41) and (8.3.42). One can write the explicit matrix representation of the CHSH operator in (8.1.10) using (8.1.11).

$$\begin{aligned} CHSH &= C^{O_1^{\mathcal{X}}} \otimes C^{O_2^{\mathcal{Y}}} - C^{O_0^{\mathcal{X}}} \otimes C^{O_2^{\mathcal{Y}}} - C^{O_0^{\mathcal{X}}} \otimes C^{O_3^{\mathcal{Y}}} - C^{O_1^{\mathcal{X}}} \otimes C^{O_3^{\mathcal{Y}}} \\ &= \Big[(P_{\mathcal{X}}^{0|1} \otimes P_{\mathcal{Y}}^{0|2} + P_{\mathcal{X}}^{1|1} \otimes P_{\mathcal{Y}}^{1|2}) - (P_{\mathcal{X}}^{0|1} \otimes P_{\mathcal{Y}}^{1|2} + P_{\mathcal{X}}^{1|1} \otimes P_{\mathcal{Y}}^{0|2}) \Big] \\ &- \Big[(P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{0|2} + P_{\mathcal{X}}^{1|0} \otimes P_{\mathcal{Y}}^{1|2}) - (P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{1|2} + P_{\mathcal{X}}^{1|0} \otimes P_{\mathcal{Y}}^{0|2}) \Big] \\ &- \Big[(P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{0|3} + P_{\mathcal{X}}^{1|0} \otimes P_{\mathcal{Y}}^{1|3}) - (P_{\mathcal{X}}^{0|0} \otimes P_{\mathcal{Y}}^{1|3} + P_{\mathcal{X}}^{1|0} \otimes P_{\mathcal{Y}}^{0|3}) \Big] \\ &- \Big[(P_{\mathcal{X}}^{0|1} \otimes P_{\mathcal{Y}}^{0|3} + P_{\mathcal{X}}^{1|1} \otimes P_{\mathcal{Y}}^{1|3}) - (P_{\mathcal{X}}^{0|1} \otimes P_{\mathcal{Y}}^{1|3} + P_{\mathcal{X}}^{1|1} \otimes P_{\mathcal{Y}}^{0|3}) \Big] \end{aligned} \tag{8.3.44}$$

Now using the explicit matrix form of the projectors in (8.3.2) and (8.3.3), the CHSH operator can further be decomposed as,

$$\begin{aligned} CHSH = &\Big[\Big(Q(\phi_A^i) \otimes Q(0) \Big) + \Big((\mathbb{1} - Q(\phi_A^i)) \otimes (\mathbb{1} - Q(0)) \Big) - \Big(Q(\phi_A^i) \otimes (\mathbb{1} - Q(0)) \Big) - \Big((\mathbb{1} - Q(\phi_A^i)) \otimes Q(0) \Big) \Big] \\ &- [(Q(0) \otimes Q(0)) + (\mathbb{1} - Q(0)) \otimes (\mathbb{1} - Q(0))) - (Q(0) \otimes (\mathbb{1} - Q(0))) - (\mathbb{1} - Q(0)) \otimes Q(0))] \\ &- \Big[\Big(Q(0) \otimes Q(\phi_B^j) \Big) + \Big((\mathbb{1} - Q(0)) \otimes (\mathbb{1} - Q(\phi_B^j)) \Big) - \Big(Q(0) \otimes (\mathbb{1} - Q(\phi_B^j)) \Big) - \Big((\mathbb{1} - Q(0)) \otimes Q(\phi_B^j) \Big) \Big] \\ &- \Big[\Big(Q(\phi_A^i) \otimes Q(\phi_B^j) \Big) + \Big((\mathbb{1} - Q(\phi_A^i)) \otimes (\mathbb{1} - Q(\phi_B^j)) \Big) - \Big(Q(\phi_A^i) \otimes (\mathbb{1} - Q(\phi_B^j)) \Big) - \Big((\mathbb{1} - Q(\phi_A^i)) \otimes Q(\phi_B^j) \Big) \Big] \end{aligned} \tag{8.3.45}$$

Solving the tensor products in each term above, one can explicitly write each term as,

The first term is
$[(Q(\phi_A^i) \otimes Q(0)) + ((\mathbb{1} - Q(\phi_A^i)) \otimes (\mathbb{1} - Q(0))) - (Q(\phi_A^i) \otimes (\mathbb{1} - Q(0))) - ((\mathbb{1} - Q(\phi_A^i)) \otimes Q(0))]$

$$= \begin{pmatrix} \cos(\phi_A^i) & 0 & \sin(\phi_A^i) & 0 \\ 0 & -\cos(\phi_A^i) & 0 & -\sin(\phi_A^i) \\ \sin(\phi_A^i) & 0 & -\cos(\phi_A^i) & 0 \\ 0 & -\sin(\phi_A^i) & 0 & \cos(\phi_A^i) \end{pmatrix} \tag{8.3.46}$$

The second term is
$[(Q(0) \otimes Q(0)) + (\mathbb{1} - Q(0)) \otimes (\mathbb{1} - Q(0))) - (Q(0) \otimes (\mathbb{1} - Q(0))) - (\mathbb{1} - Q(0)) \otimes Q(0))]$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{8.3.47}$$

The third term is
$\Big[\Big(Q(0) \otimes Q(\phi_B^j) \Big) + \Big((\mathbb{1} - Q(0)) \otimes (\mathbb{1} - Q(\phi_B^j)) \Big) - \Big(Q(0) \otimes (\mathbb{1} - Q(\phi_B^j)) \Big) - \Big((\mathbb{1} - Q(0)) \otimes Q(\phi_B^j) \Big) \Big]$

$$= \begin{pmatrix} \cos(\phi_B^j) & \sin(\phi_B^j) & 0 & 0 \\ \sin(\phi_B^j) & -\cos(\phi_B^j) & 0 & 0 \\ 0 & 0 & -\cos(\phi_B^j) & -\sin(\phi_B^j) \\ 0 & 0 & -\sin(\phi_B^j) & \cos(\phi_B^j) \end{pmatrix} \tag{8.3.48}$$

and the fourth term is

$$\left[\left(Q(\phi_A^i)\otimes Q(\phi_B^j)\right)+\left((\mathbb{1}-Q(\phi_A^i))\otimes(\mathbb{1}-Q(\phi_B^j))\right)-\left(Q(\phi_A^i)\otimes(\mathbb{1}-Q(\phi_B^j))\right)-\left((\mathbb{1}-Q(\phi_A^i))\otimes Q(\phi_B^j)\right)\right]$$

$$=\begin{pmatrix}\cos\phi_A\cos\phi_B & \cos\phi_A\sin\phi_B & \sin\phi_A\cos\phi_B & \sin\phi_A\sin\phi_B\\ \cos\phi_A\sin\phi_B & -\cos\phi_A\cos\phi_B & \sin\phi_A\sin\phi_B & -\sin\phi_A\cos\phi_B\\ \sin\phi_A\cos\phi_B & \sin\phi_A\sin\phi_B & -\cos\phi_A\cos\phi_B & -\cos\phi_A\sin\phi_B\\ \sin\phi_A\sin\phi_B & -\sin\phi_A\cos\phi_B & -\cos\phi_A\sin\phi_B & \cos\phi_A\cos\phi_B\end{pmatrix} \tag{8.3.49}$$

Finally after plugging in (8.3.46),(8.3.47),(8.3.48) and (8.3.49) in (8.3.45) one can have,

$$\boxed{CHSH=\begin{bmatrix}A & B\\ B & D\end{bmatrix}} \tag{8.3.50}$$

where

$$A=\begin{pmatrix}\cos(\phi_A^i)-1-\cos(\phi_B^j)-\cos(\phi_A^i)\cos(\phi_B^j) & -\sin(\phi_B^j)-\cos(\phi_A^i)\sin(\phi_B^j)\\ -\sin(\phi_B^j)-\cos(\phi_A^i)\sin(\phi_B^j) & -\cos(\phi_A^i)+1+\cos(\phi_B^j)+\cos(\phi_A^i)\cos(\phi_B^j)\end{pmatrix} \tag{8.3.51}$$

$$B=\begin{pmatrix}\sin(\phi_A^i)-\sin(\phi_A^i)\cos(\phi_B^j) & -\sin(\phi_A^i)\sin(\phi_B^j)\\ -\sin(\phi_A^i)\sin(\phi_B^j) & \sin(\phi_A^i)\cos(\phi_B^j)-\sin(\phi_A^i)\end{pmatrix} \tag{8.3.52}$$

$$D=\begin{pmatrix}-\cos(\phi_A^i)+1+\cos(\phi_B^j)+\cos(\phi_A^i)\cos(\phi_B^j) & \sin(\phi_B^j)+\cos(\phi_A^i)\sin(\phi_B^j)\\ \sin(\phi_B^j)+\cos(\phi_A^i)\sin(\phi_B^j) & \cos(\phi_A^i)-1-\cos(\phi_B^j)-\cos(\phi_A^i)\cos(\phi_B^j)\end{pmatrix} \tag{8.3.53}$$

Here individual operators in A, B and D are hermitian operators which implies that CHSH operator in (8.3.50) is also an hermitian.
Now from 8.3 the mapping $\phi_N^m \mapsto 1\pm\cos(\phi_N^m/2)$ where $N\in\{A,B\}$ and $m\in\{i,j\}$ is bijective for $\phi_N^m\in[0,\pi]$. Since sin and cosine functions are monotonous and continuous for $\phi_N^m\in[0,\pi/2]$, the mapping is also bijective into a subinterval of $[0,\pi/2]$.
In the first quadrant, both sin and cosine functions are positive, so restricting the arguments of them to the first quadrant ensures that the CHSH operator in (8.3.50) doesn't change, and also both sin and cosine functions have unique values in the first quadrant, so the optimization problem (8.3.40) for the two qubit system can be reformulated with the explicit matrix representation of CHSH operator in (8.3.50) as

$$\boxed{\begin{aligned}C^*_{\mathbb{C}^{4\times4}}(S_{ij})=\inf\quad & \left[\lambda D\left(\left(\rho^{ij,AB}_{ABEA'_0}\right)\,||\,\left(\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B}\right)\right)\right]\\ & \qquad +(1-\lambda)D\left(\left(\rho^{ij,AB}_{ABEA'_1}\right)\,||\,\left(\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B}\right)\right)\Big]\\ \text{s.t.}\quad & \operatorname{Tr}\left(\rho^{ij,AB}_{ABEA'_x}\cdot\begin{bmatrix}A & B\\ B & D\end{bmatrix}\right)=S_{ij}\\ & \rho^{ij,AB}_{ABEA'_x}\succeq 0\\ & \operatorname{Tr}(\rho^{ij,AB}_{ABEA'_x})=1\\ & \phi_A^i,\phi_B^j\in[0,\pi/2]\end{aligned}} \tag{8.3.54}$$

where A, B and D are block matrices as evidence from (8.3.51),(8.3.52) and (8.3.53).

8.4 Formulation of the objective function in terms of trace norm through a modified form of Pinsker's inequality

Having established the objective function in terms of a two-qubit vector space $\mathbb{C}_{(4\times4)}$ describing a single entangled pair between Alice and Bob and expressing the CHSH operator in its explicit matrix form (8.3.50), the following section focuses on lower bounding the relative entropies in (8.3.54) through a modified form

of Pinsker's inequality[48].
The modified version of Pinsker's inequality, as shown in the original work as the quantum relative entropy between ρ and $\Lambda[\rho]$ is lower bounded as,

$$D(\rho||\Lambda[\rho]) \geq log_2(2) - h\left(\frac{1 - ||\rho - \Lambda[\rho]||_1}{2}\right) \tag{8.4.1}$$

where $h(p) = -\sum_{i=0}^{1} p_i \cdot log_2(p_i)$ is the binary Shannon entropy.
The following theorem lower bound $C^*_{\mathbb{C}^{4\times4}}(S_{ij})$.

Theorem 8.4.1. Given Q be a projector not necessarily of rank 1 on a finite-dimensional Hilbert Space of dimension d. Let Λ be a pinching channel defined on the projector Q acting upon a state ρ as $\Lambda[\rho] = Q\rho Q + (\mathbb{1} - Q)\rho(\mathbb{1} - Q)$. The objective function $C^*_{\mathbb{C}^{4\times4}}(S_{ij})$ is lower bounded as

$$C^*_{\mathbb{C}^{4\times4}}(S_{ij}) \geq \log_2(2) - h\left(\frac{1}{2} - \frac{1}{2}(n)\right) \tag{8.4.2}$$

using the modified Pinsker's inequality.

Proof. The pinching channel Λ is defined as,

$$\begin{aligned}\Lambda[\rho] &= Q\rho Q + (\mathbb{1} - Q)\rho(\mathbb{1} - Q)\\ &= Q\rho Q + (\rho - Q\rho)(\mathbb{1} - Q)\\ &= Q\rho Q + (\rho - \rho Q - Q\rho + Q\rho Q)\\ &= 2Q\rho Q + \rho - \{Q, \rho\}\end{aligned} \tag{8.4.3}$$

Now applying modified pinsker inequality on (8.3.54) one can have,

$$\begin{aligned}C^*_{\mathbb{C}^{4\times4}}(S_{ij}) = \inf \quad &\Big[\lambda D\left(\left(\rho^{ij,AB}_{ABEA'_0}\right) \;||\; \left(\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B}\right)\right)\\ &+(1-\lambda)D\left(\left(\rho^{ij,AB}_{ABEA'_1}\right) \;||\; \left(\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B}\right)\right)\Big]\\ \geq &\left[\lambda\left(\log_2(2) - h\left(\frac{1 - ||\left(\rho^{ij,AB}_{ABEA'_0}\right) - \left(\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B}\right)||_1}{2}\right)\right)\right.\\ &\left.+(1-\lambda)\left(\log_2(2) - h\left(\frac{1 - ||\left(\rho^{ij,AB}_{ABEA'_1}\right) - \left(\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B}\right)||_1}{2}\right)\right)\right]\\ \geq &\left[\lambda\log_2(2) - \lambda h\left(\frac{1 - ||\left(\rho^{ij,AB}_{ABEA'_0}\right) - \left(\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B}\right)||_1}{2}\right)\right.\\ &\left.+(1-\lambda)\log_2(2) - (1-\lambda)h\left(\frac{1 - ||\left(\rho^{ij,AB}_{ABEA'_1}\right) - \left(\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B}\right)||_1}{2}\right)\right]\\ \geq &\left[\log_2(2) - \lambda h\left(\frac{1 - ||\left(\rho^{ij,AB}_{ABEA'_0}\right) - \left(\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B}\right)||_1}{2}\right)\right.\\ &\left.-(1-\lambda)h\left(\frac{1 - ||\left(\rho^{ij,AB}_{ABEA'_1}\right) - \left(\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B}\right)||_1}{2}\right)\right]\\ \text{s.t.} \quad &\mathrm{Tr}\left(\rho^{ij,AB}_{ABEA'_x} \cdot \begin{bmatrix} A & B \\ B & D \end{bmatrix}\right) = S_{ij}\end{aligned} \tag{8.4.4}$$

Let's now define two arguments a and b as,

$$\begin{aligned} a &= \frac{1 - \left\| (\rho^{ij,AB}_{ABEA'_0}) - (\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B}) \right\|_1}{2}, \\ b &= \frac{1 - \left\| (\rho^{ij,AB}_{ABEA'_1}) - (\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B}) \right\|_1}{2} \end{aligned} \tag{8.4.5}$$

Assuming that $\|\cdot\|_1 \leq 1$ is valid for normalized quantum states, one can have $a, b \in [0, 0.5]$. This ensures that $h(a)$ and $h(b)$ are well-defined. The binary entropy function, defined as $h(p) = -p\log_2 p - (1-p)\log_2(1-p)$ for $p \in [0,1]$, is a concave function from *Theorem 2.1* in [43] thus from Jensen's inequality [15], for $\lambda \in [0,1]$ and $a, b \in [0,1]$, one can have:

$$\begin{aligned} \lambda h(a) + (1-\lambda)h(b) &\leq h(\lambda a + (1-\lambda)b) \\ -[\lambda h(a) + (1-\lambda)h(b)] &\geq -h(\lambda a + (1-\lambda)b) \end{aligned} \tag{8.4.6}$$

Finally from (8.4.4), (8.4.5) and (8.4.6) one can conclude the bound on $C^*_{\mathbb{C}^{4\times4}}(S_{ij})$ as,

$$\boxed{\begin{aligned} C^*_{\mathbb{C}^{4\times4}}(S_{ij}) &\geq \log_2(2) - h(\lambda a + (1-\lambda)b) \\ &\geq \log_2(2) - h\left(\frac{1}{2} - \frac{1}{2}(n)\right) \end{aligned}} \tag{8.4.7}$$

where

$$n = \left(\lambda \left\| (\rho^{ij,AB}_{ABEA'_0}) - (\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B}) \right\|_1 + (1-\lambda) \left\| (\rho^{ij,AB}_{ABEA'_1}) - (\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B}) \right\|_1 \right) \tag{8.4.8}$$

□

It represents a convex combination of trace norms.
Now, one may consider that there exists a function $n^*(S_{ij}) \geq n$ which gives an upper bound on the given convex combination of trace norm (8.4.8). The upper bound function is given as

$$\boxed{\begin{aligned} n^*(S_{ij}) = \inf \quad & \left[\left(\lambda \left\| (\rho^{ij,AB}_{ABEA'_0}) - (\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B}) \right\|_1 + (1-\lambda) \left\| (\rho^{ij,AB}_{ABEA'_1}) - (\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B}) \right\|_1 \right) \right] \\ \text{s.t.} \quad & \operatorname{Tr}\left(\rho^{ij,AB}_{ABEA'_x} \cdot \begin{bmatrix} A & B \\ B & D \end{bmatrix} \right) = S_{ij} \\ & \rho^{ij,AB}_{ABEA'_x} \succeq 0 \\ & \operatorname{Tr}(\rho^{ij,AB}_{ABEA'_x}) = 1 \\ & \phi^i_A, \phi^j_B \in [0, \pi/2] \end{aligned}} \tag{8.4.9}$$

8.5 Semi definite programming formulation of the objective function for fixed ϕ^i_A and ϕ^j_B

A Semidefinite Program (SDP) is an optimization problem of a linear function defined over a positive semidefinite variable, subjected to affine constraints as [50],

$$\begin{aligned} \alpha = \text{maximize} \quad & \langle A, X \rangle \\ \text{s.t.} \quad & \Phi(X) = B \\ & X \in \operatorname{Pos}(\mathcal{X}) \end{aligned} \tag{8.5.1}$$

where,

$$\begin{aligned}
&X \in \mathbb{C}^{\Sigma}, \quad Y \in \mathbb{C}^{\Omega} \\
&A \in \text{Herm}(X), \quad B \in \text{Herm}(Y) \\
&\Phi : \text{Herm}(X) \rightarrow \text{Herm}(Y) \\
&(A, B, \Phi) \text{ problem's data} \\
&\langle A, X \rangle \text{ is the objective function} \\
&\Phi(X) = B, \text{ and } X \in \text{Pos}(\mathcal{X}) \\
&\alpha \text{ is the optimal value}
\end{aligned} \tag{8.5.2}$$

The trace norm of a quadratic matrix M can be represented in the form of a trace of two additional matrices P and Q [15] as,

$$\begin{aligned}
||M||_1 = inf \quad & \frac{1}{2}\text{Tr}(P+Q) \\
s.t. & \begin{pmatrix} P & M \\ M^* & Q \end{pmatrix} \succeq 0
\end{aligned} \tag{8.5.3}$$

Now the terms $(\rho^{ij,AB}_{ABEA'_0}) - (\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B})$ and $(\rho^{ij,AB}_{ABEA'_1}) - (\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B})$ from (8.4.9) can be decomposed using (8.4.3),(8.2.2)(8.3.2) and (8.3.3) as,

$$\begin{aligned}
&\left(\rho^{ij,AB}_{ABEA'_0}\right) - \left(\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B}\right) \\
&\quad = \left(\rho^{ij,AB}_{ABEA'_0}\right) - \left(2(Q(0)\otimes\mathbb{I})\cdot\rho^{ij,AB}_{ABEA'_0}\cdot(Q(0)\otimes\mathbb{I}) + \rho^{ij,AB}_{ABEA'_0} - \{(Q(0)\otimes\mathbb{I}), \rho^{ij,AB}_{ABEA'_0}\}\right) \\
&\quad = \left(\rho^{ij,AB}_{ABEA'_0}\right) - \left((Q(0)\otimes\mathbb{I})\cdot\rho^{ij,AB}_{ABEA'_0}\cdot(Q(0)\otimes\mathbb{I}) + \rho^{ij,AB}_{ABEA'_0} - (Q(0)\otimes\mathbb{I})\cdot\rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_0}\cdot(Q(0)\otimes\mathbb{I})\right) \\
&\quad = -2(Q(0)\otimes\mathbb{I})\cdot\rho^{ij,AB}_{ABEA'_0}\cdot(Q(0)\otimes\mathbb{I}) + (Q(0)\otimes\mathbb{I})\cdot\rho^{ij,AB}_{ABEA'_0} + \rho^{ij,AB}_{ABEA'_0}\cdot(Q(0)\otimes\mathbb{I}) \\
&\quad = (Q(0)\otimes\mathbb{I})\cdot\rho^{ij,AB}_{ABEA'_0} + \rho^{ij,AB}_{ABEA'_0}\cdot(Q(0)\otimes\mathbb{I}) - 2(Q(0)\otimes\mathbb{I})\cdot\rho^{ij,AB}_{ABEA'_0}\cdot(Q(0)\otimes\mathbb{I})
\end{aligned} \tag{8.5.4}$$

Similarly, one can also expand $\left(\rho^{ij,AB}_{ABEA'_1}\right) - \left(\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B}\right)$ as,

$$\left(\rho^{ij,AB}_{ABEA'_1}\right) - \left(\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B}\right) = (Q(\phi^i_A)\otimes\mathbb{I})\cdot\rho^{ij,AB}_{ABEA'_1} + \rho^{ij,AB}_{ABEA'_1}\cdot(Q(\phi^i_A)\otimes\mathbb{I}) - 2(Q(\phi^i_A)\otimes\mathbb{I})\cdot\rho^{ij,AB}_{ABEA'_1}\cdot(Q(\phi^i_A)\otimes\mathbb{I}) \tag{8.5.5}$$

The semi-definite program formulation of the objective function can now be formally constructed from the results of (8.4.9),(8.5.1),(8.5.2),(8.5.3),(8.5.4), and (8.5.5) as,

$$\boxed{\begin{aligned}
n^*(S_{ij}) = -\text{maximize} \quad & \frac{\lambda}{2}\langle(P_0+Q_0), X_0\rangle + \frac{(1-\lambda)}{2}\langle(P_1+Q_1), X_1\rangle \\
\text{s.t.} \quad & \text{Tr}\left(\rho^{ij,AB}_{ABEA'_x}\cdot\begin{bmatrix} A & B \\ B & D\end{bmatrix}\right) = S_{ij}, \\
& \begin{pmatrix} P_0 & M_0 \\ M_0^* & Q_0 \end{pmatrix} \succeq 0, \quad \begin{pmatrix} P_1 & M_1 \\ M_1^* & Q_1 \end{pmatrix} \succeq 0, \\
& Q(\phi^i_A) = \begin{bmatrix} q_{11} & q_{12} \\ q_{12}^* & q_{22}\end{bmatrix}, \quad q_{11}+q_{22} = 1, \quad q_{12}^2 \leq q_{11}q_{22}, \\
& \rho^{ij,AB}_{ABEA'_x} \succeq 0, \quad \text{Tr}(\rho^{ij,AB}_{ABEA'_x}) = 1
\end{aligned}} \tag{8.5.6}$$

where

$$\begin{aligned}
M_0 = M_0^* &= (Q(0)\otimes\mathbb{I})\cdot\rho^{ij,AB}_{ABEA'_0} + \rho^{ij,AB}_{ABEA'_0}\cdot(Q(0)\otimes\mathbb{I}) - 2(Q(0)\otimes\mathbb{I})\cdot\rho^{ij,AB}_{ABEA'_0}\cdot(Q(0)\otimes\mathbb{I}) \\
M_1 = M_1^* &= (Q(\phi^i_A)\otimes\mathbb{I})\cdot\rho^{ij,AB}_{ABEA'_1} + \rho^{ij,AB}_{ABEA'_1}\cdot(Q(\phi^i_A)\otimes\mathbb{I}) - 2(Q(\phi^i_A)\otimes\mathbb{I})\cdot\rho^{ij,AB}_{ABEA'_1}\cdot(Q(\phi^i_A)\otimes\mathbb{I})
\end{aligned} \tag{8.5.7}$$

The projector $Q(\phi^i_A)$ follows the positive semi-definite condition.

8.6 Optimization of the angles ϕ_A^i and ϕ_B^j using ϵ - net approximation

Alice's and Bob's angle $\left(\phi_A^i \text{ and } \phi_B^j\right)$ appear as constraints in the optimization problem in (8.4.9). The objective function has been formulated in standard SDP form in (8.5.6). The SDP in the previous section is built upon the assumption of fixed angles. In this section $\left(\phi_A^i \text{ and } \phi_B^j\right)$ are being optimized through ϵ - net approach. The work of [48] showed that the Alice angle ϕ_A^i can be optimised using this approach. In this work, we are showing that a similar approach exists to optimise the Bob angle ϕ_B^j, thus eliminating the polytop optimization in the Security Analysis in [48].

Given the interval $I = [0, \pi/2]$ and a desired precision $\epsilon_0 > 0$, without loss of generality an ϵ_0-net for the product space $I \times I$ is a pair of finite sets of points $\{\phi_{A_k}^i\}_{k=1}^{S_A} \subset I$ and $\{\phi_{B_l}^j\}_{l=1}^{S_B} \subset I$ such that for any $(\phi_A^i, \phi_B^j) \in I \times I$, there exist $\phi_{A_k}^i$ and $\phi_{B_l}^j$ satisfying:

$$|\phi_A^i - \phi_{A_k}^i| \leq \epsilon_0 \quad |\phi_B^j - \phi_{B_l}^j| \leq \epsilon_0 \tag{8.6.1}$$

where S_A and S_B are the number of segments in the interval I. Each segment is centralized around the angle $\phi_{A_k}^i$ and $\phi_{B_l}^j$ for k^{th} and l^{th} segments respectively for Alice and Bob.

The values of both the angles ϕ_A^i and ϕ_B^j are needed to solve any instance of the SDP. The SDP is solved for each discrete point $\phi_{A_k}^i$ and $\phi_{B_l}^j$. An error term known as pessimistic error [48] is being subtracted from each SDP's result. The error term accounts for the variation of the optimal value for any other value of ϕ_A^i and ϕ_B^j. The pessimistic error term was introduced in the original text as a function of change Δ in ϵ_0 and $\left(\phi_A^i \text{ or } \phi_B^j\right)$ accordingly. Iteratively, the segment that gives the smallest value of the objective function in (8.4.9) is chosen until a point is reached that corresponds to the global minimum. In this work, a closed form of the same is provided.

Let us define a function f as the solution of the optimisation problem $n^*(S_{ij})$ for a given angle ϕ. Each discrete angle $\phi_{A_k}^i$ and $\phi_{B_l}^j$ is being separated by a distance of $2\epsilon_0$, thus each angle represents a segment of the same width as,

$$\begin{aligned} I_A &= \left[\phi_{A_k}^i - \epsilon_0, \phi_{A_k}^i + \epsilon_0\right] \\ I_B &= \left[\phi_{B_l}^j - \epsilon_0, \phi_{B_l}^j + \epsilon_0\right] \end{aligned} \tag{8.6.2}$$

The pessimistic error terms $\Delta(\epsilon_0, \phi_A^i)$ and $\Delta(\epsilon_0, \phi_B^j)$ provide an upper bound on the absolute difference between the function value at any point within these segments and the function value at the representative discrete point,

$$\begin{aligned} |f(\phi_A^i) - f(\phi_{A_k}^i)| &\leq \Delta\left(\epsilon_0, \phi_A^i\right) \\ |f(\phi_B^j) - f(\phi_{B_l}^j)| &\leq \Delta\left(\epsilon_0, \phi_B^j\right) \\ \forall \phi_A^i &\in I_A \\ \text{and } \forall \phi_B^j &\in I_B \end{aligned} \tag{8.6.3}$$

The iterative process involves selecting the segment that yields the smallest objective function value in (8.4.9), continuing until a point corresponding to the global minimum is reached. This work aims to provide a closed-form expression for this pessimistic error.

8.6.1 Lipschitz continuity and proof of Lipschitz Continuity of the solution of a well-behaved optimization problem

Definition : A function f from $S \subset \mathbb{R}^n$ into $\mathbb{R}^m$ is Lipschitz continuous at $x \in S$ if there is a constant $L > 0$ such that

$$\|f(y) - f(x)\| \leq L\|y - x\| \tag{8.6.4}$$

for all $y \in S$ sufficiently near x.[52]

Theorem 8.6.1 (Lipschitz Continuity of Parametric Minimizers)**.** Let $f(\phi) := \min_{x \in \mathcal{X}} g(x, \phi)$, where $g : \mathcal{X} \times \Phi \longrightarrow \mathbb{R}$ satisfies the following:

(A1) For each $\phi \in \Phi$, the minimizer $f(\phi)$ exists and is unique.

(A2) For each ϕ, the function $x \mapsto g(x, \phi)$ is μ-strongly convex.

(A3) The gradient $\nabla_x g(x, \phi)$ is L_ϕ-Lipschitz in ϕ, uniformly in x.

(A4) The domain $\mathcal{X} \subset \mathbb{R}^n$ is convex and independent of ϕ.

Then the solution map $f : \Phi \longrightarrow \mathcal{X}$ is Lipschitz continuous with Lipschitz constant at most $\frac{L_\phi}{\mu}$. That is,

$$\|f(\phi_1) - f(\phi_2)\| \leq \frac{L_\phi}{\mu} \|\phi_1 - \phi_2\| \quad \forall \phi_1, \phi_2 \in \Phi. \tag{8.6.5}$$

Under assumptions (A1)–(A4), the solution map $f : \Phi \longrightarrow \mathcal{X}$ is Lipschitz continuous with Lipschitz constant $L_f = \frac{L_\phi}{\mu}$. That is,

$$\|f(\phi_1) - f(\phi_2)\| \leq \frac{L_\phi}{\mu} \|\phi_1 - \phi_2\|. \tag{8.6.6}$$

Proof. Let $x_1 := f(\phi_1)$, $x_2 := f(\phi_2)$. Then, by optimality,

$$\nabla_x g(x_1, \phi_1) = 0, \qquad \nabla_x g(x_2, \phi_2) = 0. \tag{8.6.7}$$

Subtracting these equations:

$$\begin{aligned} 0 &= \nabla_x g(x_1, \phi_1) - \nabla_x g(x_2, \phi_2) \\ &= [\nabla_x g(x_1, \phi_1) - \nabla_x g(x_1, \phi_2)] + [\nabla_x g(x_1, \phi_2) - \nabla_x g(x_2, \phi_2)] . \end{aligned}$$

Taking norms and applying the triangle inequality:

$$\|\nabla_x g(x_1, \phi_2) - \nabla_x g(x_2, \phi_2)\| = \|\nabla_x g(x_1, \phi_1) - \nabla_x g(x_1, \phi_2)\| \leq L_\phi \|\phi_1 - \phi_2\|. \tag{8.6.8}$$

By strong convexity of $g(\cdot, \phi_2)$, its gradient is μ-strongly monotone:

$$\|\nabla_x g(x_1, \phi_2) - \nabla_x g(x_2, \phi_2)\| \geq \mu \|x_1 - x_2\|. \tag{8.6.9}$$

Combining these:

$$\mu \|f(\phi_1) - f(\phi_2)\| \leq L_\phi \|\phi_1 - \phi_2\|, \tag{8.6.10}$$

which implies the Lipschitz bound:

$$\|f(\phi_1) - f(\phi_2)\| \leq \frac{L_\phi}{\mu} \|\phi_1 - \phi_2\|. \tag{8.6.11}$$

□

8.6.2 Failure of Lipschitz Continuity in the min Map of the Optimization Problem

Optimization Problem Formulation

We consider the following parameter-dependent semidefinite optimization problem for fixed indices i, j and parameterized by local angles $\phi_A^i, \phi_B^j \in [0, \pi/2]$. The objective is a convex combination of trace norm distances between quantum states and their processed versions under completely positive trace-preserving (CPTP) maps Λ_0 and Λ_1, which themselves depend on the local angles:

$$n^*(S_{ij}) = \inf_{\rho^{ij,AB}_{ABEA'_x}} \left[\lambda \left\| \rho^{ij,AB}_{ABEA'_0} - \Lambda_0[\rho^{ij}_{ABEA'_0}]^{A'B} \right\|_1 + (1-\lambda) \left\| \rho^{ij,AB}_{ABEA'_1} - \Lambda_1[\rho^{ij}_{ABEA'_1}]^{A'B} \right\|_1 \right] \tag{8.6.12}$$

$$\text{subject to} \quad \mathrm{Tr}\left(\rho^{ij,AB}_{ABEA'_x} \cdot \begin{bmatrix} A & B \\ B & D \end{bmatrix} \right) = S_{ij}, \tag{8.6.13}$$

$$\rho^{ij,AB}_{ABEA'_x} \succeq 0, \quad \mathrm{Tr}(\rho^{ij,AB}_{ABEA'_x}) = 1, \tag{8.6.14}$$

$$\phi^i_A, \phi^j_B \in [0, \pi/2].$$

We denote the minimizer map by:

$$f(\phi) := \min_{\rho^{ij}} \; g(\rho^{ij}, \phi), \tag{8.6.15}$$

where $\phi := (\phi^i_A, \phi^j_B)$ and the function $g(\rho, \phi)$ is the objective in (8.6.12).

Desired Regularity: Lipschitz Continuity of the Minimizer Map

A typical goal in parametric optimization is to establish the Lipschitz continuity of the min map $f(\phi)$ with respect to the parameters ϕ. This would provide guarantees on the stability and smooth dependence of the optimizer on the underlying physical parameters. A sufficient set of standard assumptions from variational analysis ensuring Lipschitz continuity of the minimizer map f is as follows:

1. *Unique Minimizer:* For every ϕ, the minimizer $f(\phi)$ is unique.
2. *Strong Convexity:* The function $g(\cdot, \phi)$ is strongly convex in ρ, uniformly over ϕ.
3. *Smooth Parameter Dependence:* The gradient $\nabla_\rho g(\rho, \phi)$ exists and is Lipschitz continuous in ϕ, uniformly in ρ.
4. *Fixed Domain:* The feasible set $\mathcal{D}_\phi$ of admissible ρ is independent of ϕ, or at least varies smoothly with it.

Under these conditions, it is known [14] that the minimizer map satisfies:

$$\|f(\phi_1) - f(\phi_2)\| \leq \frac{L_\phi}{\mu} \|\phi_1 - \phi_2\|, \tag{8.6.16}$$

where μ is the strong convexity constant and L_ϕ is the Lipschitz constant of the gradient.

Failure of the Assumptions in Our Context

We now examine each of the conditions (L1)–(L4) and explain why they fail in the context of our optimization problem (8.6.12)–(8.6.14).

- *(L1) Uniqueness Fails:* The trace norm $\|\cdot\|_1$ is convex but not strictly convex. Therefore, the objective function $g(\rho, \phi)$ may admit multiple minimizers. This leads to a set-valued min map $f(\phi)$, which is not necessarily continuous.
- *(L2) Strong Convexity Fails:* The function $\rho \mapsto \|\rho - \Lambda[\rho]\|_1$ is not strongly convex, as the trace norm lacks curvature. As a result, small perturbations in ϕ can lead to abrupt shifts in the optimizer $\rho^*(\phi)$.
- *(L3) Gradient Regularity Fails:* The trace norm is not differentiable at matrices with degenerate singular values. Even where differentiable, the dependence of the CPTP maps Λ_0, Λ_1 on ϕ can introduce non-Lipschitz behavior in the gradient $\nabla_\rho g(\rho, \phi)$.
- *(L4) Domain Dependence Fails:* The feasible set depends on ϕ through both:

- The CPTP maps Λ_0, Λ_1, which are angle-dependent.
- The linear constraint (8.6.13), where the operator $\begin{bmatrix} A & B \\ B & D \end{bmatrix}$ may encode expectation values of measurement observables that themselves vary with the angle parameters ϕ.

This renders the feasible set $\mathcal{D}_\phi$ non-fixed and potentially non-smooth in ϕ.

Implications

The collective failure of assumptions (L1)–(L4) implies that the min map $f(\phi)$ is not Lipschitz continuous in general. In particular:

- $f(\phi)$ may be discontinuous, or even undefined (multi-valued) at certain ϕ.
- This obstructs attempts to differentiate the optimization with respect to ϕ, complicating gradient-based algorithms.
- From a physical standpoint, small perturbations in the measurement settings (angles) can lead to large jumps in the optimal quantum states consistent with observed statistics, undermining robustness guarantees.

Conclusion

Due to the intrinsic non-smoothness and degeneracies in the structure of the objective and constraint set, the min map $f(\phi)$ in (8.6.12) is not Lipschitz continuous in the parameter ϕ. Any analysis requiring continuity or differentiability of the solution with respect to angle parameters must take into account these limitations or impose additional regularization.

8.6.3 Modification of the optimization problem accommodating Lipschitz continuity of $n^*(S_{ij})$

In order to make the objective function $n\left(\rho^{ij,AB}_{ABEA'_x}, (\phi^i_A)\right)$ in (8.4.9) continuously differentiable with respect to both the density operator $\rho^{ij,AB}_{ABEA'_x}$ and the parameters $\phi^i_{A'}$, the squared Frobenius norm is introduced. The squared Frobenius norm is smooth and continuously differentiable everywhere.[41] The Frobenius norm of an operator X is defined as,

$$\begin{aligned} \|X\|_F &= \sqrt{\mathrm{Tr}(X^* X)} \\ \|X\|_F^2 &= \mathrm{Tr}(X^* X) \end{aligned} \tag{8.6.17}$$

where X^* is the Hermitian conjugate of X. Since the trace operation and matrix multiplication are continuously differentiable, the squared Frobenius norm is also continuously differentiable concerning the elements of the operator X. Replacing the L1 norm in the original objective function with the squared Frobenius norm would yield a new objective function that is continuously differentiable concerning the density operator $\rho^{ij,AB}_{ABEA'_x}$. If the dependence of $\rho^{ij,AB}_{ABEA'_x}$ on the parameters ϕ^i_A is also continuously differentiable, then the entire objective function would be continuously differentiable concerning these parameters as well (through the chain rule). Therefore, by substituting the L1 norm with the squared Frobenius norm, the first critical assumption regarding the continuous differentiability of the objective function can be satisfied, provided the parameterisation of the density operator in terms of ϕ^i_A is also smooth.

To establish strong convexity in the objective function $n\left(\rho^{ij,AB}_{ABEA'_x}, (\phi^i_A)\right)$, it is necessary to add a strongly convex regularisation term. A function $f : \mathcal{X} \to \mathbb{R}$ is μ-strongly convex with respect to a norm $\|\cdot\|$ if for all $X, Y \in \mathcal{X}$ and $\alpha \in [0,1]$,

$$f(\alpha X + (1-\alpha)Y) \leq \alpha f(X) + (1-\alpha) f(Y) - \frac{\mu}{2}\alpha(1-\alpha)\|X - Y\|^2. \tag{8.6.18}$$

Equivalently, for twice differentiable f, this requires the Hessian $\nabla^2 f(X)$ to satisfy:

$$\langle \nabla^2 f(X)H, H\rangle \geq \mu \|H\|^2 \quad \forall H \text{ in the domain.} \tag{8.6.19}$$

[15]. The convex regularization term introduced here is $\frac{\mu}{2}||\rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1}||^2_F$, where $||\cdot||_F$ denotes the Frobenius norm.The squared Frobenius norm, $||X||^2_F = \text{Tr}(X^*X)$, possesses the property of being μ-strongly convex for some $\mu > 0$ when considered over the space of matrices equipped with the Frobenius inner product $\langle A, B\rangle = \text{Tr}(A^*B)$[30][15]. The squared Frobenius norm for a matrix $X \in \mathbb{C}^{n\times n}$ is:

$$f(X) = \|X\|^2_F = \text{Tr}(X^*X). \tag{8.6.20}$$

The gradient of the function is $\nabla f(X) = 2X$ and the second derivative is a constant operator,

$$\begin{aligned} &\nabla^2 f(X)[H] = 2H \quad \text{for any perturbation } H. \\ &\langle \nabla^2 f(X)H, H\rangle = \text{Tr}\left((2H)^*H\right) = 2\text{Tr}(H^*H) = 2\|H\|^2_F. \end{aligned} \tag{8.6.21}$$

The Hessian of $f(X)$ can be identified with the linear operator that maps a perturbation H to $2H$. Thus, by choosing $\mu = 2$, the squared Frobenius norm is inherently 2-strongly convex concerning the Frobenius norm itself. Now for the above assumed regularizer $\frac{\mu}{2}\|\rho\|^2_F$. The second derivative scales by $\mu/2$, giving,

$$\begin{aligned} &\nabla^2\left(\frac{\mu}{2}\|\rho\|^2_F\right)[H] = \mu H. \\ &\langle \nabla^2\left(\frac{\mu}{2}\|\rho\|^2_F\right)H, H\rangle = \mu\|H\|^2_F. \end{aligned} \tag{8.6.22}$$

and this satisfies the strong convexity condition with modulus μ.

To establish the convexity of the feasible set, let us first have

$$chsh : I \times I \mapsto CHSH \tag{8.6.23}$$

where $I = [0, \frac{\pi}{2}]$ and CHSH is the operator as defined in (8.3.50). Thus the original trace inner product become $Tr(\rho \; chsh(\phi^i_A, \phi^j_B)) = S_{ij}$. Now let us define $c_A = \cos(\phi^i_A)$, $s_A = \sin(\phi^i_A)$, $c_B = \cos(\phi^j_B)$, and $s_B = \sin(\phi^j_B)$ with the constraints $c^2_A + s^2_A = 1$ and $c^2_B + s^2_B = 1$. The most direct way is to treat the angles ϕ^i_A and ϕ^j_B as fixed parameters during the optimization over the density operator ρ. In this scenario, *chsh* becomes a constant Hermitian operator concerning ρ, and the constraint defines an affine subspace within the convex set of density operators, thus ensuring a convex feasible set for ρ.

However, if ϕ^i_A and ϕ^j_B are also optimization variables, the non-linear dependence of *chsh* on these angles persists because the quadratic equality constraints $c^2_A + s^2_A = 1$ and $c^2_B + s^2_B = 1$ define non-convex sets. Convex relaxations might be considered as $c^2_B + s^2_B \leq 1$, but they come with potential limitations in terms of the physical interpretation of the results.

In our scenario, one can are essentially optimizing (8.4.9) on fixed-sized segments with a fixed value of ϕ^i_A and ϕ^j_B, so the feasible set of ρ is convex because the $chsh(\phi^i_A, \phi^j_B)$ becomes a fixed Hermitian operator, and the set of all density operators is convex.

Now the last two assumptions, i.e. Lipschitz Continuity of $\nabla_\rho n\left(\rho^{ij,AB}_{ABEA'_x}, (\phi^i_A)\right)$ in ϕ^i_A and $\nabla_\rho n\left(n^*(S_{ij}), (\phi^i_A)\right) = 0$ are primarily being violated due to the non-differential nature of the objective function in (8.4.9). The introduction of the Frobenius norm had removed the non-differential nature. Now, finally, the modified optimization problem becomes,

$$\boxed{\begin{aligned} n^*(S_{ij}) = \inf \quad & \lambda \left\|(\rho^{ij,AB}_{ABEA'_0}) - (\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B})\right\|^2_F + (1-\lambda)\left\|(\rho^{ij,AB}_{ABEA'_1}) - (\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B})\right\|^2_F \\ & + \frac{\mu}{2}\|\rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1}\|^2_F \\ \text{s.t.} \quad & \mathrm{Tr}\left(\rho^{ij,AB}_{ABEA'_x} \cdot chsh(\phi^i_A, \phi^j_B)\right) = S_{ij} \\ & \phi^i_A, \phi^j_B \in [0, \pi/2], \\ & \rho^{ij,AB}_{ABEA'_x} \succeq 0 \\ & \mathrm{Tr}(\rho^{ij,AB}_{ABEA'_x}) = 1 \end{aligned}} \tag{8.6.24}$$

where one can define the objective function as a function of the density operator $\rho^{ij,AB}_{ABEA'_x}$ and Alice's angle (ϕ^i_A) as

$$\begin{aligned} & n\left(\rho^{ij,AB}_{ABEA'_x}, (\phi^i_A)\right) \\ =& \lambda \left\|(\rho^{ij,AB}_{ABEA'_0}) - (\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B})\right\|^2_F + (1-\lambda)\left\|(\rho^{ij,AB}_{ABEA'_1}) - (\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B})\right\|^2_F + \frac{\mu}{2}\|\rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1}\|^2_F \end{aligned} \tag{8.6.25}$$

where the channel Λ_1 is being defined on ϕ^i_A and $x \in \{0,1\}$ The modified version of the optimization problem using Frobenius norms can be reformulated into an SDP using Schur complements. Let us first decompose each Frobenius norm term into the corresponding inner product form as,

$$\begin{aligned} n^*(S_{ij}) = \text{maximize} \quad & -\left(\lambda\langle(\rho^{n^*}_0)^*, (\rho^{n^*}_0)\rangle + (1-\lambda)\langle(\rho^{n^*}_1)^*(\rho^{n^*}_1)\rangle + \frac{\mu}{2}\langle(\rho^{n^*}_2)^*(\rho^{n^*}_2)\rangle\right) \\ \text{s.t.} \quad & \mathrm{Tr}\left(\rho^{ij,AB}_{ABEA'_x} \cdot chsh(\phi^i_A, \phi^j_B)\right) = S_{ij} \\ & \phi^i_A, \phi^j_B \in [0, \pi/2], \\ & \rho^{ij,AB}_{ABEA'_x} \succeq 0 \\ & \mathrm{Tr}(\rho^{ij,AB}_{ABEA'_x}) = 1 \end{aligned} \tag{8.6.26}$$

where $\rho^{n^*}_0 = (\rho^{ij,AB}_{ABEA'_0}) - (\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B})$, $\rho^{n^*}_1 = (\rho^{ij,AB}_{ABEA'_1}) - (\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B})$ and $\rho^{n^*}_2 = \rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1}$

Here, each inner product term $\langle(\rho^{n^*}_k)^*, (\rho^{n^*}_k)\rangle$ for $k \in \{0,1,2\}$ in (8.6.26) is quadratic in $\rho^{n^*}_k$. The standard SDP formulation requires that both the objective function is linear in its decision variables and the constraints are in the form of linear matrix inequalities. Let $t_k \geq \langle(\rho^{n^*}_k)^*, (\rho^{n^*}_k)\rangle$ then using Schur's complement[8] one can have,

$$\begin{pmatrix} [t_k] & (vec(\rho^{n^*}_k))^* \\ vec(\rho^{n^*}_k) & \mathbb{I} \end{pmatrix} \succeq 0 \tag{8.6.27}$$

Since *chsh* is fixed for given value of ϕ_A^i and ϕ_B^j, the standard SDP formulation would be as,

$$\boxed{\begin{aligned} n^*(S_{ij}) = \text{maximize} \quad & -\left(\lambda t_0 + (1-\lambda)t_1 + \frac{\mu}{2}t_2\right) \\ \text{s.t.} & \begin{pmatrix} [t_0] & (vec(\rho_0^{n^*}))^* \\ vec(\rho_0^{n^*}) & \mathbb{I} \end{pmatrix} \succeq 0 \\ & \begin{pmatrix} [t_1] & (vec(\rho_1^{n^*}))^* \\ vec(\rho_1^{n^*}) & \mathbb{I} \end{pmatrix} \succeq 0 \\ & \begin{pmatrix} [t_2] & (vec(\rho_2^{n^*}))^* \\ vec(\rho_2^{n^*}) & \mathbb{I} \end{pmatrix} \succeq 0 \\ & \operatorname{Tr}\left(\rho^{ij,AB}_{ABEA'_x} \cdot chsh(\phi_A^i, \phi_B^j)\right) = S_{ij}, \\ & \phi_A^i, \phi_B^j \in [0, \pi/2], \\ & \rho^{ij,AB}_{ABEA'_x} \succeq 0 \\ & \operatorname{Tr}(\rho^{ij,AB}_{ABEA'_x}) = 1 \end{aligned}} \tag{8.6.28}$$

where $\rho_0^{n^*} = (\rho^{ij,AB}_{ABEA'_0}) - (\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B})$, $\rho_1^{n^*} = (\rho^{ij,AB}_{ABEA'_1}) - (\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B})$, $\rho_2^{n^*} = \rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1}$ and $t_k \geq \langle(\rho_k^{n^*})^*, (\rho_k^{n^*})\rangle$ for $k \in \{0, 1, 2\}$

Proof of the given set of assumptions

Now one can formally prove that our modified optimization problem satisfies the assumptions for being Lipschitz continuous

Lemma 8.6.2. The Modified objective function $n\left(\rho^{ij,AB}_{ABEA'_x}, (\phi_A^i)\right)$ is Continuously Differentiable with respect to both the density operators $\rho^{ij,AB}_{ABEA'_0}$ and $\rho^{ij,AB}_{ABEA'_1}$ and the parameter ϕ_A^i.

Proof. The Modified objective function is (8.6.25) $n\left(\rho^{ij,AB}_{ABEA'_x}, (\phi_A^i)\right) = \lambda \left\|(\rho^{ij,AB}_{ABEA'_0}) - (\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B})\right\|_F^2 + (1-\lambda)\left\|(\rho^{ij,AB}_{ABEA'_1}) - (\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B})\right\|_F^2 + \frac{\mu}{2}\|\rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1}\|_F^2$ where the channel Λ_1 is being defined on ϕ_A^i

- The squared Frobenius norm $\|X\|_F^2 = \operatorname{Tr}(X^*X)$ is smooth and infinitely differentiable.
- Λ_0, Λ_1 are linear maps (quantum channels), so $\Lambda_x[\rho]$ is linear in ρ.
- The regularization term $\frac{\mu}{2}\|\rho\|_F^2$ is quadratic and smooth.

Hence $n\left(\rho^{ij,AB}_{ABEA'_0}, (\phi_A^i)\right)$ is Continuously Differentiable with respect to both the density operator $\rho^{ij,AB}_{ABEA'_0}$ and the parameters ϕ_A^i.Note that the angle ϕ_A^i is related to the channel Λ_1 through (8.2.2),(8.3.1),(8.3.2) and (8.3.3). □

Lemma 8.6.3. The modified objective function $n\left(\rho^{ij,AB}_{ABEA'_0}, (\phi_A^i)\right)$ is strongly convex in $\rho^{ij,AB}_{ABEA'_0}$ when $\rho^{ij,AB}_{ABEA'_1}$ is fixed.

Proof. The objective function (8.6.25) includes the term

$$\frac{\mu}{2}\|\rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1}\|_F^2, \tag{8.6.29}$$

which is μ-strongly convex in $\rho^{ij,AB}_{ABEA'_0}$ for a fixed $\rho^{ij,AB}_{ABEA'_1}$. For any $\rho^{ij,AB}_{ABEA'_0}, \rho^{kl,AB}_{ABEA'_0}$:

$$\begin{aligned} n\left(\rho^{ij,AB}_{ABEA'_0}, (\phi^i_A)\right) \geq n\left(\rho^{kl,AB}_{ABEA'_0}, (\phi^i_A)\right) + \nabla_\rho n\left(\rho^{kl,AB}_{ABEA'_0}, (\phi^i_A)\right)^T (\rho^{ij,AB}_{ABEA'_0} - \rho^{kl,AB}_{ABEA'_0}) \\ + \frac{\mu}{2}\|\rho^{ij,AB}_{ABEA'_0} - \rho^{kl,AB}_{ABEA'_0}\|_F^2. \end{aligned} \tag{8.6.30}$$

Gradient Calculation: The Fréchet derivative of $\|A\|_F^2$ is $2A$ [55]. For fixed $\rho^{kl,AB}_{ABEA'_1}$, the gradient of n with respect to $\rho^{kl,AB}_{ABEA'_0}$ is:

$$\begin{aligned} &\nabla_\rho n\left(\rho^{kl,AB}_{ABEA'_0}, (\phi^i_A)\right) \\ &= \nabla_\rho\left(\lambda\left\|\rho^{kl,AB}_{ABEA'_0} - \Lambda_0[\rho^{kl,A'B}_{ABEA'_0}]\right\|_F^2 + (1-\lambda)\left\|\rho^{kl,AB}_{ABEA'_1} - \Lambda_1[\rho^{kl,A'B}_{ABEA'_1}]\right\|_F^2 + \frac{\mu}{2}\|\rho^{kl,AB}_{ABEA'_0} - \rho^{kl,AB}_{ABEA'_1}\|_F^2\right) \\ &= 2\lambda\left(\rho^{kl,AB}_{ABEA'_0} - \Lambda_0[\rho^{kl,A'B}_{ABEA'_0}]\right)\cdot(\mathbb{I} - \Lambda_0^*) + 2(1-\lambda)\left(\rho^{kl,AB}_{ABEA'_1} - \Lambda_1[\rho^{kl,A'B}_{ABEA'_1}]\right)\cdot(\mathbb{I} - \Lambda_1^*) + \mu\left(\rho^{kl,AB}_{ABEA'_0} - \rho^{kl,AB}_{ABEA'_1}\right). \end{aligned} \tag{8.6.31}$$

Simplification: If Λ_0, Λ_1 are linear quantum channels (completely positive trace-preserving maps), their adjoints satisfy

$$\Lambda_x^*[\rho - \Lambda_x[\rho]] = 0 \quad \text{for } x = 0, 1. \tag{8.6.32}$$

This reduces the gradient to:

$$\nabla_\rho n = 2\lambda\left(\rho^{kl,AB}_{ABEA'_0} - \Lambda_0[\rho^{kl,A'B}_{ABEA'_0}]\right) + \mu\left(\rho^{kl,AB}_{ABEA'_0} - \rho^{kl,AB}_{ABEA'_1}\right). \tag{8.6.33}$$

Strong Convexity: The μ-term ensures μ-strong convexity in $\rho^{ij,AB}_{ABEA'_0}$, while the remaining terms are convex. Hence, n is μ-strongly convex in $\rho^{ij,AB}_{ABEA'_0}$ when $\rho^{ij,AB}_{ABEA'_1}$ is fixed.

Corollary 8.6.4. The modified objective function $n\left(\rho^{ij,AB}_{ABEA'_x}, (\phi^i_A)\right)$ is strongly convex in $\rho^{ij,AB}_{ABEA'_1}$ when $\rho^{ij,AB}_{ABEA'_0}$ is fixed.

Proof. By symmetry, the μ-term

$$\frac{\mu}{2}\|\rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1}\|_F^2 \tag{8.6.34}$$

is μ-strongly convex in $\rho^{ij,AB}_{ABEA'_1}$, and the proof follows analogously to 8.6.3. □

□

Lemma 8.6.5. The feasible set $\mathcal{D}$ of all density operator and the interval $I = [0, \frac{\pi}{2}]$ is compact as well as convex.

Proof. A set $\mathcal{S}$ is convex if for any $x, y \in \mathcal{S}$ and $\lambda \in [0,1]$, the convex combination $\lambda x + (1-\lambda)y \in \mathcal{S}$. In finite-dimensional spaces, a set is compact if it is closed (contains all its limit points) and bounded (fits within some finite-radius ball). The constraints are, $\mathrm{Tr}\left(\rho^{ij,AB}_{ABEA'_x} \cdot chsh(\phi^i_A, \phi^j_B)\right) = S_{ij}$, $\phi^i_A, \phi^j_B \in [0, \pi/2]$, $\rho^{ij,AB}_{ABEA'_x} \succeq 0$, $\mathrm{Tr}(\rho^{ij,AB}_{ABEA'_x}) = 1$
a) Proof of convexity: For $\rho_1, \rho_2 \succeq 0$ and $\lambda \in [0,1]$,

$$\rho = \lambda\rho_1 + (1-\lambda)\rho_2 \succeq 0, \quad \mathrm{Tr}(\rho) = \lambda\mathrm{Tr}(\rho_1) + (1-\lambda)\mathrm{Tr}(\rho_2) = 1. \tag{8.6.35}$$

Thus the set of density operators $\rho \succeq 0$ with $\mathrm{Tr}(\rho) = 1$ is convex.[55],[57]

- Convexity of Angle Intervals: For $\phi_1, \phi_2 \in [0, \pi/2]$,

$$\lambda\phi_1 + (1-\lambda)\phi_2 \in [0, \pi/2] \tag{8.6.36}$$

  Thus the intervals $[0, \pi/2]$ for ϕ^i_A, ϕ^j_B are convex.

- Linear Trace Constraint: For ρ_1, ρ_2 satisfying $\mathrm{Tr}(\rho_1 \cdot chsh(\phi^i_A, \phi^j_B)) = S_{ij}$ and $\mathrm{Tr}(\rho_2 \cdot chsh(\phi^i_A, \phi^j_B)) = S_{ij}$,

$$\mathrm{Tr}\left((\lambda\rho_1 + (1-\lambda)\rho_2) \cdot chsh(\phi^i_A, \phi^j_B)\right) = \lambda S_{ij} + (1-\lambda)S_{ij} = S_{ij}. \tag{8.6.37}$$

  Thus the constraint $\mathrm{Tr}(\rho \cdot chsh(\phi^i_A, \phi^j_B)) = S_{ij}$ is linear in ρ if M is fixed.

- Combined Convexity: The Cartesian product of convex sets (density matrices, angle intervals) under linear constraints is convex.

b) Proof of compactness:

- Closedness: The set of all density operator $\mathcal{D}$ is closed because of positive semidefiniteness ($\rho \succeq 0$) is preserved under limits. The trace condition $\mathrm{Tr}(\rho) = 1$ is preserved under limits. The constraint $\mathrm{Tr}(\rho M) = S_{ij}$ is closed (as the preimage of a closed set under a continuous function)[55].
  The interval $[0, \pi/2]$ is a closed interval in $\mathbb{R}$.

- Boundedness: The Frobenius norm satisfies $\|\rho\|_F \leq \sqrt{\mathrm{Tr}(\rho^2)} \leq \sqrt{\mathrm{Tr}(\rho)} = 1$.Thus, the set of all density operators $\mathcal{D}$ is bounded.[55]
  The interval $[0, \pi/2]$ is bounded in $\mathbb{R}$.

□

Lemma 8.6.6. $\nabla_\rho n\left(\rho^{ij,AB}_{ABEA'_x}, (\phi^i_A)\right)$ is lipschitz continuous in ϕ^i_A for $x \in \{0, 1\}$.

Proof. From (8.6.31), one can have the gradient of $\nabla_\rho n\left(\rho^{ij,AB}_{ABEA'_x}, (\phi^i_A)\right)$. For a given segment ϕ^i_A parameterizes Λ_0, Λ_1 smoothly because for a particular segment its fixed, then $\nabla_\rho n\left(\rho^{ij,AB}_{ABEA'_x}, (\phi^i_A)\right)$ depends smoothly on ϕ^i_A. Since $\cos(\phi^i_A)$, $\sin(\phi^i_A)$, and their derivatives are bounded, $\nabla_\rho n\left(\rho^{ij,AB}_{ABEA'_x}, (\phi^i_A)\right)$ is Lipschitz continuous in ϕ^i_A as,

$$\|\nabla_\rho n\left(\rho^{ij,AB}_{ABEA'_0}, (\phi^i_A)\right) - \nabla_\rho n\left(\rho^{kl,AB}_{ABEA'_0}, (\phi^k_A)\right)\| \leq L_{\phi_A}\|\phi^i_A - \phi^k_A\|. \tag{8.6.38}$$

where $\rho^{ij,AB}_{ABEA'_x}$ and $\rho^{kl,AB}_{ABEA'_x}$ are being two arbitrary states and ϕ^i_A and ϕ^k_A are associated angles.
Hence $\nabla_\rho n\left(\rho^{ij,AB}_{ABEA'_x}, (\phi^i_A)\right)$ is lipschitz continuous in ϕ^i_A for $x \in \{0, 1\}$. □

Lemma 8.6.7. Existence of Unique optima, $\nabla_\rho n\left(n^*(S_{ij}), (\phi^i_A)\right) = 0$

Proof. Strong convexity ensures a unique minimiser $n^*(S_{ij})$ and compactness of the feasible set guarantees existence. With smoothness and strong convexity, the solution satisfies:

$$\nabla_\rho n\left(n^*(S_{ij}), (\phi^i_A)\right) = 0 \tag{8.6.39}$$

Convex function: A function $g : I \longrightarrow \mathbb{R}$, where I is an interval (or any convex set in general) in $\mathbb{R}$, is said to be convex if for any two points $x_a, x_b \in I$ and for any $t \in [0, 1]$, the following inequality holds,

$$g(tx_a + (1-t)x_b) \leq tg(x_a) + (1-t)g(x_b) \tag{8.6.40}$$

The function g is defined on an interval I. For any $x_a, x_b \in I$ and $t \in [0,1]$, the point $\phi = tx_a + (1-t)x_b$ also lies within the domain $I.\phi$ represents any point on the line segment connecting x_a and x_b.
Now the term $tx_a + (1-t)x_b$ represents a convex combination of x_a and x_b. As t varies from 0 to 1, this expression covers all points on the line segment connecting x_a and x_b. For $t = 0$, we get x_b, $t = 1$, we get x_a and for $0 < t < 1$, we get a point strictly between x_a and x_b.
Now, finally the term $tg(x_a) + (1-t)g(x_b)$represents the y-coordinate of the point on the secant line connecting the points $(x_a, g(x_a))$ and $(x_b, g(x_b))$ at the x-coordinate $\phi = tx_a + (1-t)x_b$. □

Lemma 8.6.8. Given a convex function g defined on a domain containing distinct points x_1 and x_2, for any $t \in [0,1]$, let $\phi = tx_1 + (1-t)x_2$ be a convex combination of x_1 and x_2. The y-coordinate of the point on the secant line passing through $(x_1, g(x_1))$ and $(x_2, g(x_2))$ at the x-coordinate ϕ is given by the convex combination of the function values, $tg(x_1) + (1-t)g(x_2)$.

Proof. Let us have the convex function g in (8.6.40).For any two points x_a and x_b in the domain I of g, the Secant line passing through $(x_a, g(x_a))$ and $(x_b, g(x_b))$ is,

$$y - g(x_a) = \frac{g(x_b) - g(x_a)}{x_b - x_a}(x - x_a) \tag{8.6.41}$$

Now, let $x = tx_a + (1-t)x_b$. Then,

$$\begin{aligned} x - x_a &= tx_a + (1-t)x_b - x_a \\ &= (t-1)x_a + (1-t)x_b \\ &= (1-t)(x_b - x_a) \end{aligned} \tag{8.6.42}$$

Substituting this into the equation of the secant line, in (8.6.41)

$$\begin{aligned} y - g(x_a) &= \frac{g(x_b) - g(x_a)}{x_b - x_a}(1-t)(x_b - x_a) \\ y - g(x_a) &= (1-t)(g(x_b) - g(x_a)) \\ y &= g(x_a) + (1-t)g(x_b) - (1-t)g(x_a) \\ y &= g(x_a) - (1-t)g(x_a) + (1-t)g(x_b) \\ y &= (1-(1-t))g(x_a) + (1-t)g(x_b) \\ y &= tg(x_a) + (1-t)g(x_b) \end{aligned} \tag{8.6.43}$$

Thus from 8.6.3 the inequality in (8.6.40) formally states that the value of the function g at any point x between x_a and x_b is less than or equal to the corresponding y-value on the secant line connecting $(x_a, g(x_a))$ and $(x_b, g(x_b))$. □

Bounding the pessimistic error terms $\Delta\left(\epsilon_0, \phi_A^i\right)$ and $\Delta\left(\epsilon_0, \phi_B^j\right)$

The solution of the modified optimization problem in (8.6.24), $n^*(S_{ij})$, is thus established to be Lipschitz continuous. Now one can proceed towards formulating a closed form for the pessimistic error terms, $\Delta\left(\epsilon_0, \phi_A^i\right)$ and $\Delta\left(\epsilon_0, \phi_B^j\right)$. The solution of the SDP in (8.6.28) gives an optimal value for the k^{th} and l^{th} segment centered around $\phi_{A_k}^i$ and $\phi_{B_l}^j$ for Alice's and Bob's, respectively. The solution of the SDP is based on fixed ϕ_A^i, ϕ_B^j.To find the optimal value over the whole interval I, one needs a slightly different approach from the one given in [48] because here both the angles are being optimised using the ϵ-net approach. The process begins by solving the SDP for each segment of Alice (or Bob), parameterised by discretised angles $\left(\phi_{A_k}^i, \phi_{B_l}^j\right)$. Subsequently, the segment that produced the minimum SDP value is iteratively refined. This refinement involves further subdivision of the segment's angle range, and the SDP is re-evaluated. This process continues until a global minimum is found for Alice or Bob, depending on which segment is being refined. Subsequently, the optimal value for the other party is determined by applying the same iterative refinement process, but with the angle of the first party now fixed at their globally optimal value. Without loss of generality, one can assign the function f to the solution of the optimization problem in (8.6.24). Let denote two functions for Alice's and Bob's angle, respectively.

$$f_A(\phi_A^i) = R_A, f_B(\phi_B^j) = R_B \tag{8.6.44}$$

where R_A and R_B give the optimal value of the SDP in (8.6.28) or equivalently the optimization problem in (8.6.24) for the given parameterized angle.
Now, since the functions in (8.6.44) are Lipschitz continuous, one can have the following relation for each segment where $\phi_{A_k}^i$ and $\phi_{B_l}^j$ are the centres of the segments k and l, respectively.

$$\begin{aligned} |f_A(\phi_A^i) - f_A(\phi_{A_k}^i)| &\leq L_A|\phi_A^i - \phi_{A_k}^i| \\ |f_B(\phi_B^j) - f_B(\phi_{B_l}^j)| &\leq L_B|\phi_B^j - \phi_{B_l}^j| \\ \forall\phi_A^i &\in I_A \\ \text{and } \forall\phi_B^j &\in I_B \end{aligned} \tag{8.6.45}$$

This measures the deviation of the solution when being measured with the angles$\left(\phi_{A_k}^i, \phi_{B_l}^j\right)$ at the centre of the segment and some other angle ϕ_A^i or ϕ_B^j.
The supremum of the functions in (8.6.44) are given as,

$$\begin{aligned} M_A^i &= sup_{\phi_A^i \in I_A} f_A(\phi_A^i) \\ M_B^j &= sup_{\phi_B^j \in I_B} f_B(\phi_B^j) \end{aligned} \tag{8.6.46}$$

On a similar note, the infimum is being given as,

$$\begin{aligned} m_A^i &= inf_{\phi_A^i \in I_A} f_A(\phi_A^i) \\ m_B^j &= inf_{\phi_B^j \in I_B} f_B(\phi_B^j) \end{aligned} \tag{8.6.47}$$

The Lipschitz continuity of the function in (8.6.44) implies that they are continuous in the given interval, so the supremums in (8.6.46) are the corresponding maximums and the infimum in (8.6.47) are the corresponding minimums. The supremum and infimum guarantee that no number smaller than M_A^i or (M_B^j) can serve as an upper-bound, and equivalently, no number larger than m_A^i or (m_B^j) for $f_A(\phi_A^i)$ or $(f_B(\phi_B^j))$.
The quantity that one is interested in bounding is,

$$\begin{aligned} &sup_{\phi_A^i \in I_A} \left| f_A(\phi_A^i) - f_A(\phi_{A_k}^i) \right| \\ &sup_{\phi_B^j \in I_B} \left| f_B(\phi_B^j) - f_B(\phi_{B_l}^j) \right| \end{aligned} \tag{8.6.48}$$

where $\phi_{A_k}^i$ and $\phi_{B_l}^j$ being the center of the segment k and l respectively The deviation in (8.6.45) is being maximised at the boundary of the respective segments.

Theorem 8.6.9. The value of $|f_A(\phi_A^i) - f_A(\phi_{A_k}^i)|$ and $\left| f_B(\phi_B^j) - f_B(\phi_{B_l}^j) \right|$ is maximum when $\phi_A^i = \phi_{A_k}^i \pm \epsilon_0$ and $\phi_B^i = \phi_{B_k}^i \pm \epsilon_0$

Proof. To prove the maximum deviation or the solution of (8.6.48), one typically needs to satisfy a certain set of assumptions as

- Lipschitz continuity of $f_A(\phi_A^i)$ and $f_B(\phi_B^j)$ in their respective intervals I_A and I_B.

- Strict monotonicity of the function $f_A(\phi_A^i)$ and $f_B(\phi_B^j)$ in their respective intervals.
- Convexity or concavity of $f_A(\phi_A^i)$ and $f_B(\phi_B^j)$ in their respective interval.

Proof for satisfying lipschitz continuity of $f_A(\phi_A^i)$ and $f_B(\phi_B^j)$ in their respective intervals I_A and I_B: From the result of 8.6.1 8.6.3 to 8.6.3 and the analogy given in (8.6.44) $f_A(\phi_A^i)$ and $f_B(\phi_B^j)$ are lipschitz continuous in their respective intervals I_A and I_B.
Proof for strict monotonicity of the function $f_A(\phi_A^i)$ and $f_B(\phi_B^j)$ in their respective intervals: The parameters ϕ_A^i and ϕ_B^j appear in the modified optimization problem in (8.6.28) through the CHSH operator M and ϕ_A^i through channel Λ_1 (8.2.2),(8.3.1),(8.3.2) and (8.3.3) only. More specifically, they appear as a function of *sin* and *cosine* in both the CHSH operator M and in the 2×2 generalised projector $Q(\cdot)$. Now both the sine and cosine functions are monotonous in our interval, I_A and I_B (8.6.44) as they are already monotonous and continuous in the whole interval I.
Proof for convexity or concavity of $f_A(\phi_A^i)$ and $f_B(\phi_B^j)$ in their respective interval: Let us consider $g(x) = \sin(x)$, $h(x) = \cos(x)$, from the second derivative test, one can have

$$\begin{aligned} g''(\phi_A^i) &= -\sin(\phi_A^i) \\ h''(\phi_A^i) &= -\cos(\phi_A^i) \end{aligned} \tag{8.6.49}$$

Now, from the definition of convexity or concavity and the second derivative,
Convexity: A function $f(x)$ is convex on an interval I if $f''(x)$ is positive.

$$f''(x) > 0 \, \forall x \in I \tag{8.6.50}$$

Concavity: A function $f(x)$ is concave on an interval I if $f''(x)$ is negative.

$$f''(x) < 0 \, \forall x \in I \tag{8.6.51}$$

In the first quadrant or the interval I, both *sin* and *cosine* function remain positive, so from (8.6.49) the second derivative of both $g(\cdot)$ and $h(\cdot)$ retain their sign. Thus $\sin(\phi_A^i)$ and $\cos(\phi_A^i)$ and so $f_A(\phi_A^i)$ and $f_B(\phi_B^j)$ are concave in the interval I and thus in the sub-intervals I_A and I_B[51].

Proof by Contradiction:
Assume, for the sake of contradiction, that the maximum deviation of $|f_A(\phi_A^i) - f_A(\phi_{A_k}^i)|$ occurs at some $\phi_A^{*i} \in I_A$ such that $|\phi_A^{*i} - \phi_{A_k}^i| < \epsilon_0$. Let $I_A = [\phi_{A_k}^i - \epsilon_0, \phi_{A_k}^i + \epsilon_0]$. This means ϕ_A^{*i} lies strictly within the open interval $(\phi_{A_k}^i - \epsilon_0, \phi_{A_k}^i + \epsilon_0)$. Since $f_A(\phi_A^i)$ is strictly monotonic on the closed interval I_A, its maximum and minimum values on this interval must occur at the endpoints, $\phi_{A_k}^i - \epsilon_0$ and $\phi_{A_k}^i + \epsilon_0$. Consider the difference $|f_A(\phi_A^i) - f_A(\phi_{A_k}^i)|$. We want to show that the maximum of this absolute difference is achieved at one of the endpoints. The absolute difference at the endpoints is,

$$\begin{aligned} &|f_A(\phi_{A_k}^i - \epsilon_0) - f_A(\phi_{A_k}^i)| \\ &|f_A(\phi_{A_k}^i + \epsilon_0) - f_A(\phi_{A_k}^i)| \end{aligned} \tag{8.6.52}$$

Now, for any $\phi_A^{*i} \in (\phi_{A_k}^i - \epsilon_0, \phi_{A_k}^i + \epsilon_0)$ where $\phi_A^{*i} \neq \phi_{A_k}^i$. Due to the strict monotonicity of f_A,

- If f_A is strictly increasing, then for $\phi_{A_k}^i - \epsilon_0 < \phi_A^{*i} < \phi_{A_k}^i + \epsilon_0$, we have $f_A(\phi_{A_k}^i - \epsilon_0) < f_A(\phi_A^{*i}) < f_A(\phi_{A_k}^i + \epsilon_0)$. This implies that either $|f_A(\phi_{A_k}^i + \epsilon_0) - f_A(\phi_{A_k}^i)| > |f_A(\phi_A^{*i}) - f_A(\phi_{A_k}^i)|$ or $|f_A(\phi_{A_k}^i - \epsilon_0) - f_A(\phi_{A_k}^i)| > |f_A(\phi_A^{*i}) - f_A(\phi_{A_k}^i)|$ (or both).
- If f_A is strictly decreasing, then for $\phi_{A_k}^i - \epsilon_0 < \phi_A^{*i} < \phi_{A_k}^i + \epsilon_0$, we have $f_A(\phi_{A_k}^i - \epsilon_0) > f_A(\phi_A^{*i}) > f_A(\phi_{A_k}^i + \epsilon_0)$. Again, this implies that either $|f_A(\phi_{A_k}^i + \epsilon_0) - f_A(\phi_{A_k}^i)| > |f_A(\phi_A^{*i}) - f_A(\phi_{A_k}^i)|$ or $|f_A(\phi_{A_k}^i - \epsilon_0) - f_A(\phi_{A_k}^i)| > |f_A(\phi_A^{*i}) - f_A(\phi_{A_k}^i)|$ (or both).

In both cases (strictly increasing or strictly decreasing), the maximum deviation $|f_A(\phi_A^i) - f_A(\phi_{A_k}^i)|$ cannot occur strictly within the interval $(\phi_{A_k}^i - \epsilon_0, \phi_{A_k}^i + \epsilon_0)$. Therefore, the maximum must occur at one of the endpoints, $\phi_A^i = \phi_{A_k}^i - \epsilon_0$ or $\phi_A^i = \phi_{A_k}^i + \epsilon_0$, which can be written as $\phi_A^i = \phi_{A_k}^i \pm \epsilon_0$. A similar argument holds for $|f_B(\phi_B^j) - f_B(\phi_{B_l}^j)|$, with the maximum occurring at $\phi_B^j = \phi_{B_l}^j \pm \epsilon_0$. □

One can now combine the results of (8.6.48),(8.6.45) and 8.6.3 as,

$$\begin{aligned} sup_{\phi_A^i \in I_A} |f_A(\phi_A^i) - f_A(\phi_{A_k}^i)| &= L_A \epsilon_0 \\ sup_{\phi_B^j \in I_B} |f_B(\phi_B^j) - f_B(\phi_{B_l}^j)| &= L_B \epsilon_0 \\ \forall \phi_A^i \in I_A & \\ \text{and } \forall \phi_B^j \in I_B & \end{aligned} \tag{8.6.53}$$

And from the definition of the pessimistic error terms $\Delta\left(\epsilon_0, \phi_A^i\right)$ and $\Delta\left(\epsilon_0, \phi_B^j\right)$ in (8.6.3) as,

$$\begin{aligned} \Delta\left(\epsilon_0, \phi_A^i\right) &= L_A \epsilon_0 \\ \Delta\left(\epsilon_0, \phi_B^j\right) &= L_B \epsilon_0 \end{aligned} \tag{8.6.54}$$

Since $f_A(\phi_A^i)$ and $f_B(\phi_B^j)$ is differentiable in our Interval I one can have a direct value of L_A and L_B using Taylor series for infinitely differentiable at $\phi_A^{i_*}$ and $\phi_B^{j_*}$ respectively,

$$\begin{aligned} f_A(\phi_A^i) &= \sum_{n=0}^{\infty} \frac{f_A^n(\phi_A^{i_*})}{n!} (\phi_A^i - \phi_A^{i_*})^n \\ f_B(\phi_B^j) &= \sum_{m=0}^{\infty} \frac{f_B^m(\phi_B^{j_*})}{m!} (\phi_B^j - \phi_B^{j_*})^m \\ \forall \phi_A^{i_*} &\in I_A \\ \text{and } \forall \phi_B^{j_*} &\in I_B \end{aligned} \tag{8.6.55}$$

Now in this case, one needs a linear approximation of $f_A(\phi_A^i)$ near $\phi_{A_k}^i$ and $f_B(\phi_B^j)$ near $\phi_{B_l}^j$ for small ϵ_0. One can neglect the higher-order derivatives terms as $f_A^1(\phi_{A_k}^i)(\phi_A^i - \phi_{A_k}^i)$ and $f_B^1(\phi_{B_l}^j)(\phi_B^j - \phi_{B_l}^j)$ dominates as $|\phi_A^i - \phi_{A_k}^i|$ and $|\phi_B^j - \phi_{B_l}^j|$ tends to 0 for k^{th} and l^{th} segments of Alice and Bob respectively.
The necessity of incorporating higher-order terms in the error analysis arises when ϵ_0 is significant, the underlying functions exhibit strong non-linearities, or when a more precise and tighter bound on the error is desired. In such scenarios, relying solely on linear approximations becomes insufficient to accurately capture the function's behavior within the interval of width $2\epsilon_0$. The contributions from higher derivatives, which are neglected in a first-order analysis, become substantial and must be accounted for to achieve the required accuracy in the error estimate. Thus, using the first order linear approximation one can get an approximation of $f_A(\phi_A^i)$ and $f_A(\phi_B^j)$ near $\phi_{A_k}^i$ and $\phi_{B_l}^j$ as,

$$\begin{aligned} f_A(\phi_A^i) &= f_A(\phi_{A_k}^i) + f_A^1(\phi_{A_k}^i)(\phi_A^i - \phi_{A_k}^i) + \mathcal{O}((\phi_A^i - \phi_{A_k}^i)^2) \\ |f_A(\phi_A^i) - f_A(\phi_{A_k}^i)| &\le |f_A^1(\phi_{A_k}^i)||(\phi_A^i - \phi_{A_k}^i)| + \mathcal{O}((\phi_A^i - \phi_{A_k}^i)^2) \end{aligned} \tag{8.6.56}$$

On a similar note,

$$|f_B(\phi_B^j) - f_B(\phi_{B_l}^j)| \le |f_B^1(\phi_{B_l}^j)||(\phi_B^j - \phi_{B_l}^j)| + \mathcal{O}((\phi_B^j - \phi_{B_l}^j)^2) \tag{8.6.57}$$

The higher-order terms, represented by $\mathcal{O}((\phi_A^i - \phi_{A_k}^i)^2)$ and $\mathcal{O}((\phi_B^j - \phi_{B_l}^j)^2)$, are indeed positive if the second derivatives are positive (indicating local convexity). While typically removing positive terms would weaken an inequality, the context of a pessimistic error bound requires careful consideration. The pessimistic error is designed to overestimate the potential deviation. By truncating the Taylor series and neglecting these positive higher-order terms, we are essentially underestimating the actual deviation $|f_A(\phi^i_A) - f_A(\phi^i_{A_k})|$ and $|f_B(\phi^j_B) - f_B(\phi^j_{B_l})|$. Consequently, when these underestimated deviations are used to construct a pessimistic error (which is subtracted from the SDP result), the error itself becomes overestimated. This overestimation of the error leads to a more conservative (and potentially less tight) lower bound on the true optimal value of the SDP. Thus after first order linear approximation one get,

$$\begin{aligned} sup_{\phi^i_A \in I_A} |f_A(\phi^i_A) - f_A(\phi^i_{A_k})| \leq |f^1_A(\phi^i_{A_k})||(\phi^i_A - \phi^i_{A_k})| \\ sup_{\phi^j_B \in I_B} |f_B(\phi^j_B) - f_B(\phi^j_{B_l})| \leq |f^1_B(\phi^j_{B_l})||(\phi^j_B - \phi^j_{B_l})| \end{aligned} \tag{8.6.58}$$

To ensure (8.6.58) holds $\forall$, $\phi^i_A, \phi^j_B \in I$ one can have, from (8.6.58),(8.6.54),

$$\boxed{\begin{aligned} \Delta\left(\epsilon_0, \phi^i_A\right) &= max_{\phi^i_A \in I_A} |f^1_A(\phi^i_{A_k})|\epsilon_0 \\ \Delta\left(\epsilon_0, \phi^j_B\right) &= max_{\phi^j_B \in I_B} |f^1_B(\phi^j_{B_l})|\epsilon_0 \end{aligned}} \tag{8.6.59}$$

where $L_A = max_{\phi^i_A \in I_A} |f^1_A(\phi^i_A)|$ and $L_B = max_{\phi^j_B \in I_B} |f^1_B(\phi^j_B)|$.

8.6.4 Relating the change in ϕ^i_A and ϕ^j_B to the optimization problem through CHSH operator

The optimization problem presented in equation (8.6.24) exhibits a dependence on Alice's angle, denoted as ϕ^i_A, through the channel Λ_1 as defined by equations (8.2.2), (8.3.1), (8.3.2), and (8.3.3). Additionally, it depends on the CHSH operator, as specified in equations (8.3.50) and (8.6.24). Bob's angle, ϕ^j_B, solely influences the CHSH operator. Notably, variations in the CHSH operator are of particular interest due to their direct impact on the feasible solution set of the optimization problem. Furthermore, analysing the effect of changes in the feasible region holds greater significance than examining alterations in the channel Λ_1. This is because Alice's and Bob's parameters jointly determine the feasible region, allowing for a more unified analysis concerning parameter variations.
The feasible set in (8.6.24) can be relaxed by allowing all the density operators $\sigma^{ij,AB}_{ABEA'_X}$ that achieve CHSH violation greater than or equal to S_{ij}, [48].

$$\begin{aligned} S^{\phi^i_A \cup \phi^j_B}_{ij} = &\left\{\sigma \mid \exists\, \phi_A \text{ with } \phi_B = \phi^j_B : \operatorname{Tr}[\sigma\, chsh(\phi_A, \phi^j_B)] \geq S_{ij}\right\} \\ &\cup \left\{\sigma \mid \exists\, \phi_B \text{ with } \phi_A = \phi^i_A : \operatorname{Tr}[\sigma\, chsh(\phi^i_A, \phi_B)] \geq S_{ij}\right\}. \end{aligned} \tag{8.6.60}$$

The above equation took the union of two sets. The first set accounts for all those density operators $\sigma^{ij,AB}_{ABEA'_X}$ in the block specified by index i, j that achieve CHSH violation greater than S_{ij} for fixed ϕ^j_B. The second set similarly account for fixed ϕ^i_A.
Now, consider the following set

$$M^{\phi^i_A \cup \phi^j_B}_{ij,\epsilon} := \bigcup_{|\delta_A| \leq \epsilon_0} S^{\phi^i_A + \delta_A \,\cup\, \phi^j_B}_{ij} \cup \bigcup_{|\delta_B| \leq \epsilon_0} S^{\phi^i_A \,\cup\, \phi^j_B + \delta_B}_{ij} \cup \bigcup_{\substack{|\delta_A| \leq \epsilon_0 \\ |\delta_B| \leq \epsilon_0}} S^{\phi^i_A + \delta_A \,\cup\, \phi^j_B + \delta_B}_{ij}. \tag{8.6.61}$$

In the aforementioned equation, the analysis focuses on a set of density operators derived by introducing a parameter ϵ to the initial angles ϕ^i_A and ϕ^j_B, both individually and concurrently. The defining characteristic of these modified density operators, denoted as $\sigma^{ij,AB}_{ABEA'_X}$, is that the expected value of the CHSH operator evaluated on them remains greater than or equal to the initial CHSH value, S_{ij}, associated with the ij^{th} block. The angles vary in the range I_A and I_B.

Spectral Norm of the CHSH Operator's sensitivity to angle perturbations

Given discrete points $\phi^i_{A_k}$ and $\phi^j_{B_l}$ located at the midpoint of their respective intervals, the inherent symmetry of the underlying function or data distribution implies that the maximum deviation from these discrete points will occur symmetrically around them. Consequently, a small perturbation ϵ from the midpoint will result in equal magnitudes of deviation, such that the deviation at $\phi^i_{A_k+\epsilon}$ is equivalent to the deviation at $\phi^i_{A_k-\epsilon}$, and similarly for $\phi^j_{B_l}$.

The spectral norm of the difference between the CHSH operators evaluated at $\phi^i_{A_k}$ and $\phi^i_{A_k\pm\epsilon}$ (and similarly for $\phi^j_{B_l}$ and $\phi^j_{B_l\pm\epsilon}$) is equal to the largest singular value of this difference operator. For Hermitian operators, such as the CHSH operator, the largest singular value coincides with the absolute value of the largest eigenvalue.

$$\begin{aligned}\delta_p = \max \Big(& ||chsh(\phi^i_A, \phi^j_B) - chsh(\phi^i_A + \epsilon, \phi^j_B)||_\infty, \\ & ||chsh(\phi^i_A, \phi^j_B) - chsh(\phi^i_A, \phi^j_B + \epsilon)||_\infty, \\ & ||chsh(\phi^i_A, \phi^j_B) - chsh(\phi^i_A + \epsilon, \phi^j_B + \epsilon)||_\infty \Big)\end{aligned} \tag{8.6.62}$$

where δ_p defined as the maximum of the spectral norms of the differences in the CHSH operator, corresponds to the largest singular value of these difference operators. For the Hermitian CHSH operator, this largest singular value is equivalent to the absolute value of the largest eigenvalue of the respective difference operators.

$$\begin{aligned}& \left\| chsh(\phi^i_A, \phi^j_B) - chsh(\phi^i_A + \epsilon, \phi^j_B) \right\|_\infty \\ =& \left\| \begin{pmatrix} \cos(\phi^i_A + \epsilon) - \cos\phi^i_A & \sin(\phi^i_A + \epsilon) - \sin\phi^i_A \\ \sin(\phi^i_A + \epsilon) - \sin\phi^i_A & -\cos(\phi^i_A + \epsilon) - \cos\phi^i_A \end{pmatrix} \otimes \begin{pmatrix} -(1 - \cos\phi^j_B) & \sin\phi^j_B \\ \sin\phi^j_B & 1 - \cos\phi^j_B \end{pmatrix} \right\|_\infty\end{aligned} \tag{8.6.63}$$

Now, one can use some simple trigonometric relations to decompose the relation $\Delta\cos = \cos(\phi^i_A + \epsilon) - \cos\phi^i_A$ and $\Delta\sin = \sin(\phi^i_A + \epsilon) - \sin\phi^i_A$.

$$\begin{aligned}&\Delta\cos = \cos(\phi^i_A + \epsilon) - \cos(\phi^i_A) = -2\sin\left(\frac{(\phi^i_A + \epsilon) + \phi^i_A}{2}\right)\sin\left(\frac{(\phi^i_A + \epsilon) - \phi^i_A}{2}\right) = -2\sin\left(\phi^i_A + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \\ &\Delta\sin = \sin(\phi^i_A + \epsilon) - \sin(\phi^i_A) = 2\cos\left(\frac{(\phi^i_A + \epsilon) + \phi^i_A}{2}\right)\sin\left(\frac{(\phi^i_A + \epsilon) - \phi^i_A}{2}\right) = 2\cos\left(\phi^i_A + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \\ &1 - \cos\phi^j_B = 2\sin^2\left(\frac{\phi^j_B}{2}\right) \\ &\sin\phi^j_B = 2\sin\left(\frac{\phi^j_B}{2}\right)\cos\left(\frac{\phi^j_B}{2}\right)\end{aligned} \tag{8.6.64}$$

From (8.6.64) and (8.6.63),

$$
\begin{aligned}
&\left\|chsh(\phi_A^i, \phi_B^j) - chsh(\phi_A^i + \epsilon, \phi_B^j)\right\|_\infty \\
=&\left\| \begin{pmatrix} -2\sin\left(\phi_A^i + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & 2\cos\left(\phi_A^i + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \\ 2\cos\left(\phi_A^i + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & 2\sin\left(\phi_A^i + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \end{pmatrix} \otimes \begin{pmatrix} -2\sin^2\left(\frac{\phi_B^j}{2}\right) & 2\sin\left(\frac{\phi_B^j}{2}\right)\cos\left(\frac{\phi_B^j}{2}\right) \\ 2\sin\left(\frac{\phi_B^j}{2}\right)\cos\left(\frac{\phi_B^j}{2}\right) & 2\sin^2\left(\frac{\phi_B^j}{2}\right) \end{pmatrix} \right\|_\infty \\
=&\left\|2\sin\frac{\epsilon}{2}\left(\cos(\phi_A^i + \frac{\epsilon}{2})\sigma_x - \sin(\phi_A^i + \frac{\epsilon}{2})\sigma_z\right) \otimes 2\left(Q\left(\phi_B^j\right) - \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\right)\right\|_\infty \\
=&\left\|2\sin\left(\frac{\epsilon}{2}\right)\left[\left(\cos\left(\phi_A^i\right)\cos\left(\frac{\epsilon}{2}\right) - \sin\left(\phi_A^i\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_x - \left(\sin\left(\phi_A^i\right)\cos\left(\frac{\epsilon}{2}\right) + \cos\left(\phi_A^i\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_z\right]\right. \\
&\qquad \left. \otimes 2\left(Q\left(\phi_B^j\right) - \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\right)\right\|_\infty
\end{aligned}
\tag{8.6.65}
$$

Using Taylor expansions for $\sin\{\phi\}$ and $\cos\{\phi\}$, $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$ and $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$ If $x \ll 1, \quad \cos x \approx 1, \quad \sin x \approx x$. Also using If $x \ll 1, \quad \cos a + \sin b \cdot x \approx \cos a, \quad \sin a + \cos b \cdot x \approx \sin a$, one can have

$$
\begin{aligned}
&\sin\left(\frac{\epsilon}{2}\right) \approx \frac{\epsilon}{2} \\
&\left(\cos\left(\phi_A^i\right)\cos\left(\frac{\epsilon}{2}\right) - \sin\left(\phi_A^i\right)\sin\left(\frac{\epsilon}{2}\right)\right) \approx \cos\left(\phi_A^i\right) \\
&\left(\sin\left(\phi_A^i\right)\cos\left(\frac{\epsilon}{2}\right) + \cos\left(\phi_A^i\right)\sin\left(\frac{\epsilon}{2}\right)\right) \approx \sin\left(\phi_A^i\right)
\end{aligned}
\tag{8.6.66}
$$

Now after combining the results of (8.6.65) and linear approximation for $\epsilon > 0$ in (8.6.66) one can have,

$$
\begin{aligned}
&\left\|chsh(\phi_A^i, \phi_B^j) - chsh(\phi_A^i + \epsilon, \phi_B^j)\right\|_\infty \\
\approx&\left\|\epsilon\left(\cos\left(\phi_A^i\right)\sigma_x - \sin\left(\phi_A^i\right)\sigma_z\right) \otimes 2\left(Q\left(\phi_B^j\right) - \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\right)\right\|_\infty
\end{aligned}
\tag{8.6.67}
$$

On a similar note one can find the spectral norm when only ϕ_B^j is being perturbated by ϵ as,

$$
\begin{aligned}
&||chsh(\phi_A^i, \phi_B^j) - chsh(\phi_A^i, \phi_B^j + \epsilon)||_\infty \\
=&\left\| \begin{pmatrix} 2\cos^2(\frac{\phi_A^i}{2}) & 2\sin\left(\frac{\phi_A^i}{2}\right)\cos\left(\frac{\phi_A^i}{2}\right) \\ 2\sin\left(\frac{\phi_A^i}{2}\right)\cos\left(\frac{\phi_A^i}{2}\right) & -2\cos^2(\frac{\phi_A^i}{2}) \end{pmatrix} \otimes \begin{pmatrix} 2\sin\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & -2\cos\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \\ -2\cos\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & -2\sin\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \end{pmatrix} \right\|_\infty \\
=&\left\|2\left(Q\left(\phi_A^i\right) - \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right)\right. \\
&\qquad \left. \otimes -2\sin\left(\frac{\epsilon}{2}\right)\left[\left(\cos\left(\phi_B^j\right)\cos\left(\frac{\epsilon}{2}\right) - \sin\left(\phi_B^j\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_x - \left(\sin\left(\phi_B^j\right)\cos\left(\frac{\epsilon}{2}\right) + \cos\left(\phi_B^j\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_z\right]\right\|_\infty \\
\approx&\left\|2\left(Q\left(\phi_A^i\right) - \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right) \otimes \epsilon\left(\cos\left(\phi_B^j\right)\sigma_x - \sin\left(\phi_B^j\right)\sigma_z\right)\right\|_\infty
\end{aligned}
\tag{8.6.68}
$$

Finally one can also find the spectral norm when both ϕ_A^i and ϕ_B^j perturbated by ϵ as,

$$\begin{aligned}
&||chsh(\phi_A^i, \phi_B^j) - chsh(\phi_A^i + \epsilon, \phi_B^j + \epsilon)||_\infty \\
=&\Big\| \begin{pmatrix} 2\sin\left(\phi_A^i + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & -2\cos\left(\phi_A^i + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \\ -2\cos\left(\phi_A^i + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & -2\sin\left(\phi_A^i + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \end{pmatrix} \otimes \begin{pmatrix} 2\sin\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & -2\cos\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \\ -2\cos\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & -2\sin\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \end{pmatrix} \Big\|_\infty \\
=&\Big\| -2\sin\left(\frac{\epsilon}{2}\right)\left[\left(\cos\left(\phi_A^i\right)\cos\left(\frac{\epsilon}{2}\right) - \sin\left(\phi_A^i\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_x - \left(\sin\left(\phi_A^i\right)\cos\left(\frac{\epsilon}{2}\right) + \cos\left(\phi_A^i\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_z\right] \\
&\quad \otimes -2\sin\left(\frac{\epsilon}{2}\right)\left[\left(\cos\left(\phi_B^j\right)\cos\left(\frac{\epsilon}{2}\right) - \sin\left(\phi_B^j\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_x - \left(\sin\left(\phi_B^j\right)\cos\left(\frac{\epsilon}{2}\right) + \cos\left(\phi_B^j\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_z\right] \Big\|_\infty \\
\approx&\Big\| \epsilon\left(\cos\left(\phi_A^i\right)\sigma_x - \sin\left(\phi_A^i\right)\sigma_z\right) \otimes \epsilon\left(\cos\left(\phi_B^j\right)\sigma_x - \sin\left(\phi_B^j\right)\sigma_z\right) \Big\|_\infty
\end{aligned} \tag{8.6.69}$$

Now combining the results of (8.6.67),(8.6.68) and (8.6.69) into (8.6.62) one can have,

$$\begin{aligned}
\delta_p \approx \max\Big(&\Big\| \epsilon\left(\cos\left(\phi_A^i\right)\sigma_x - \sin\left(\phi_A^i\right)\sigma_z\right) \otimes 2\left(Q\left(\phi_B^j\right) - \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right) \Big\|_\infty, \\
&\Big\| 2\left(Q\left(\phi_A^i\right) - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) \otimes \epsilon\left(\cos\left(\phi_B^j\right)\sigma_x - \sin\left(\phi_B^j\right)\sigma_z\right) \Big\|_\infty, \\
&\Big\| \epsilon\left(\cos\left(\phi_A^i\right)\sigma_x - \sin\left(\phi_A^i\right)\sigma_z\right) \otimes \epsilon\left(\cos\left(\phi_B^j\right)\sigma_x - \sin\left(\phi_B^j\right)\sigma_z\right) \Big\|_\infty \Big)
\end{aligned} \tag{8.6.70}$$

One can provide an upper bound to the above relation using the property of submultiplicity of spectral norm[55] as,

$$\begin{aligned}
\delta_p \leq \max\Big(&\Big\| \epsilon\left(\cos\left(\phi_A^i\right)\sigma_x - \sin\left(\phi_A^i\right)\sigma_z\right) \Big\|_\infty \cdot \Big\| 2\left(Q\left(\phi_B^j\right) - \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right) \Big\|_\infty, \\
&\Big\| 2\left(Q\left(\phi_A^i\right) - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) \Big\|_\infty \cdot \Big\| \epsilon\left(\cos\left(\phi_B^j\right)\sigma_x - \sin\left(\phi_B^j\right)\sigma_z\right) \Big\|_\infty, \\
&\Big\| \epsilon\left(\cos\left(\phi_A^i\right)\sigma_x - \sin\left(\phi_A^i\right)\sigma_z\right) \Big\|_\infty \cdot \Big\| \epsilon\left(\cos\left(\phi_B^j\right)\sigma_x - \sin\left(\phi_B^j\right)\sigma_z\right) \Big\|_\infty \Big) \\
\leq \max &\left(\epsilon \cdot 2\sin\left(\frac{\phi_B^j}{2}\right), \quad 2\cos\left(\frac{\phi_A^i}{2}\right)\cdot\epsilon, \quad \epsilon^2 \right) \\
\leq \quad & 2\epsilon_0
\end{aligned} \tag{8.6.71}$$

Equality holds for the spectral norm of Kronecker products, so this upper bound is exact. Now in the interval I, $\sin\left(\frac{\phi_B^j}{2}\right)$ and $\cos\left(\frac{\phi_A^i}{2}\right)$ function are monotonically increasing and decreasing respectively. The first term $2\sin\left(\frac{\phi_B^j}{2}\right)$ is increasing from 0 to 1.414 and the second term $2\cos\left(\frac{\phi_A^i}{2}\right)$ is decreasing from 2ϵ to 1.414. Thus, the maximum deviation of the CHSH operator is achieved when ϕ_A^i is fixed at 0 but ϕ_B^j is increased by ϵ_0. Thus every state $\rho_{ABEA'_x}^{ij,AB}$ that would attain CHSH value S_{ij} at $\phi_A^i = 0, \phi_B^j = \phi_{B_l}^j$ or $\left(\phi_A^i = 0, \phi_B^j = \phi_B^j \pm \epsilon_0\right)$ would attain CHSH value $S_{ij} - 2\epsilon_0$ at $\phi_A^i = 0, \phi_B^j = \phi_B^j \pm \epsilon_0$ or $\left(\phi_A^i = 0, \phi_B^j = \phi_{B_l}^j\right)$. Thus one have,

$$\begin{aligned}
&\mathrm{Tr}\big(\rho, \cdot chsh(0, \phi_B^j)\big) = S_{ij}, \quad \mathrm{Tr}\big(\rho, \cdot chsh(0, \phi_B^j + \epsilon_0)\big) = S_{ij} - 2\epsilon_0, \quad \mathrm{Tr}\big(\rho, \cdot chsh(0, \phi_B^j - \epsilon_0)\big) = S_{ij} - 2\epsilon_0, \\
\text{or} \quad &\mathrm{Tr}\big(\rho, \cdot chsh(0, \phi_B^j)\big) = S_{ij} - 2\epsilon_0, \quad \mathrm{Tr}\big(\rho, \cdot chsh(0, \phi_B^j + \epsilon_0)\big) = S_{ij}, \quad \mathrm{Tr}\big(\rho, \cdot chsh(0, \phi_B^j - \epsilon_0)\big) = S_{ij}, \\
&\forall \rho \in M_{ij,\epsilon}^{\phi_B^j}.
\end{aligned} \tag{8.6.72}$$

where

$$M_{ij,\epsilon}^{\phi_B^j}$$
$$= \underbrace{\left\{\rho \in M_{ij,\epsilon}^{\phi_A^i \cup \phi_B^j} \;\middle|\; \mathrm{Tr}(\rho, chsh(0, \phi_B^j)) = S_{ij}\right\}}_{\text{Case A}} \cup \underbrace{\left\{\rho \in M_{ij,\epsilon}^{\phi_A^i \cup \phi_B^j} \;\middle|\; \exists s \in \{+1, -1\} : \mathrm{Tr}(\rho, chsh(0, \phi_B^j + s\epsilon_0)) = S_{ij}\right\}}_{\text{Case B}}. \tag{8.6.73}$$

Thus, for each segment I_B and I_A being centralised around discrete $\phi_{B_l}^j$ and $\phi_{A_k}^i$ for l^{th} and k^{th} segment respectively performing the optimization with relaxed constraint value $S - 2\epsilon_0$ would take into account the dependency on ϕ_A^i and ϕ_B^j through the maximum deviation in the CHSH value. Particularly it will occur when $\phi_A^i = 0$ and $\phi_B^j = \phi_B^j \pm \epsilon_0$.

8.6.5 Relating the change in ϕ_A^i to the optimization problem through the channel Λ_1 in the objective function

The objective function in (8.6.25) is a function of both the density operator and Alice's angle ϕ_A^i through the channel Λ_1,(8.2.2),(8.3.1),(8.3.2) and (8.3.3).
To analyse the dependency of ϕ_A^i on the optimization problem, we reinterpret the Frobenius norm terms in (8.6.25) through the lens of dual norms. Dual norm are an indispensable tool in quantum information theory and semi-definite programming because they bridge quantum operational limits (e.g., distinguishability, entanglement) and convex optimization via conic duality, enabling tractable solutions with physical interpretability. Since the Frobenius norm is self-dual, this framework retains the norm structure while emphasising its role as a maximiser of inner products. Specifically, translating the squared Frobenius differences into their dual characterisation reveals how ϕ_A^i governs alignment conditions between the states $\rho_{ABE_{A_x'}}^{ij,AB}$ and the maps Λ_0, Λ_1. The dual norm's supremum property enables bounding critical trade-offs in the objective function, such as the balance between state fidelity $(\lambda, 1-\lambda)$ and distinguishability μ, which directly depend on ϕ_A^i.

Theorem 8.6.10. The Frobenius norm $\|\cdot\|_F$ on $m \times n$ matrices is self-dual.

Proof. Let $A \in \mathbb{C}^{m \times n}$ and $A \in L(\mathcal{X}, \mathcal{Y})$ for some complex euclidean space $\mathcal{X}, \mathcal{Y}$. The Frobenius norm is given as[55],

$$\|A\|_F = \sqrt{\mathrm{Tr}(A^* A)} = \sqrt{\langle A, A \rangle}, \tag{8.6.74}$$

Now, from the definition of the dual of a norm,

$$\|A\|_{F,*} = \{\sup \langle Y, A \rangle : \|Y\|_F \leq 1\} \tag{8.6.75}$$

where $Y \in L(\mathcal{Y}, X)$ Now from Cauchy–Schwarz inequality for $\langle \cdot, \cdot \rangle_F$

$$\langle Y, A \rangle_F \;\leq\; \|Y\|_F \, \|A\|_F \;\leq\; \|A\|_F \tag{8.6.76}$$

Thus from (8.6.75) and (8.6.76),

$$\|A\|_{F,*} \leq \|A\|_F \tag{8.6.77}$$

The equality will hold only when Y is aligned with A Let $Y = A / \|A\|_F$. Then $\|Y\|_F = 1$ and

$$\langle Y, A \rangle_F = \frac{1}{\|A\|_F} \langle A, A \rangle_F = \frac{\|A\|_F^2}{\|A\|_F} = \|A\|_F. \tag{8.6.78}$$

Therefore the supremum is attained when Y aligned with A and

$$\|A\|_{F,*} = \|A\|_F \tag{8.6.79}$$

□

Now from the modified optimization problem in (8.6.24) and the definition of channel Λ_0 and Λ_1 in (8.2.2) one can expand the optimization problem as,

$$
\begin{aligned}
n^*(S_{ij}) = \inf \quad & \lambda \mathrm{Tr}\Big(\Big[(\rho^{ij,AB}_{ABEA'_0}) - (\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B})\Big]^* \Big[(\rho^{ij,AB}_{ABEA'_0}) - (\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B})\Big] \Big) \\
& + (1-\lambda)\mathrm{Tr}\left(\Big[(\rho^{ij,AB}_{ABEA'_1}) - (\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B})\Big]^* \Big[(\rho^{ij,AB}_{ABEA'_1}) - (\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B})\Big]\right) \\
& + \frac{\mu}{2}\mathrm{Tr}\left(\Big[\rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1}\Big]^* \Big[\rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1}\Big]\right) \\
= \inf \quad & \lambda \mathrm{Tr}\Big(\Big[(\rho^{ij,AB}_{ABEA'_0}) - (\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B})\Big]^2 \Big) + (1-\lambda)\mathrm{Tr}\left(\Big[(\rho^{ij,AB}_{ABEA'_1}) - (\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B})\Big]^2\right) \\
& + \frac{\mu}{2}\mathrm{Tr}\left(\Big[\rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1}\Big]^2\right) \\
= \inf \quad & \lambda \mathrm{Tr}\Big(\Big[(\rho^{ij,AB}_{ABEA'_0}) - \Big((Q(0)\otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_0}(Q(0)\otimes \mathbb{I}) + \{\mathbb{I} - (Q(0)\otimes\mathbb{I})\}\rho^{ij,AB}_{ABEA'_0}\{\mathbb{I} - (Q(0)\otimes\mathbb{I})\}\Big)\Big]^2 \Big) \\
& + (1-\lambda)\mathrm{Tr}\left(\Big[(\rho^{ij,AB}_{ABEA'_1}) - \Big((Q(\phi^i_A)\otimes\mathbb{I})\rho^{ij,AB}_{ABEA'_1}(Q(\phi^i_A)\otimes\mathbb{I}) \right. \\
& \qquad \left. + \{\mathbb{I} - (Q(\phi^i_A)\otimes\mathbb{I})\}\rho^{ij,AB}_{ABEA'_1}\{\mathbb{I} - (Q(\phi^i_A)\otimes\mathbb{I})\}\Big)\Big]^2\right) \\
& + \frac{\mu}{2}\mathrm{Tr}\left(\Big[\rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1}\Big]^2\right) \\
& \mathrm{Tr}\left(\rho^{ij,AB}_{ABEA'_x}\cdot chsh(\phi^i_A, \phi^j_B)\right) = S_{ij} \\
\text{s.t.} \quad & \phi^i_A, \phi^j_B \in [0, \pi/2], \\
& \rho^{ij,AB}_{ABEA'_x} \succeq 0 \\
& \mathrm{Tr}(\rho^{ij,AB}_{ABEA'_x}) = 1
\end{aligned}
\tag{8.6.80}
$$

Any density operator ρ and a projector operator can be decomposed using the commutator and the anti-commutator using the following Lemma.

Lemma 8.6.11. *$P\rho = \frac{1}{2}\{\rho, P\} - \frac{1}{2}[\rho, P]$ for $\rho \in D(\mathcal{X})$ for some complex euclidean space $\mathcal{X}$ and $P \in Pos(\mathcal{X})$.*

Proof. Expanding the right-hand using the definition of anti-commutator and commutator as,

$$
\begin{aligned}
&\frac{1}{2}\{\rho, P\} - \frac{1}{2}[\rho, P] \\
=&\frac{1}{2}\left(\rho P + P\rho\right) - \frac{1}{2}[\rho P - P\rho] \\
=&\frac{1}{2}(P\rho) + \frac{1}{2}(P\rho) \\
=&P\rho
\end{aligned}
\tag{8.6.81}
$$

□

Now using the 8.6.5 and (8.6.80) one can have,

$$
\begin{aligned}
n^*(S_{ij}) = \inf \quad & \lambda \mathrm{Tr}\Bigg(\bigg[\rho^{ij,AB}_{ABEA'_0} - \bigg(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_0}, (Q(0)\otimes \mathbb{I})\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_0}, (Q(0)\otimes \mathbb{I})]\bigg)(Q(0)\otimes \mathbb{I}) \\
& \qquad - \bigg(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_0}, (\mathbb{I}-(Q(0)\otimes \mathbb{I}))\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_0}, (\mathbb{I}-(Q(0)\otimes \mathbb{I}))]\bigg)\{\mathbb{I}-(Q(0)\otimes \mathbb{I})\}\bigg]^2\Bigg) \\
& + (1-\lambda)\mathrm{Tr}\Bigg(\bigg[\rho^{ij,AB}_{ABEA'_1} - \bigg(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A)\otimes \mathbb{I})\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A)\otimes \mathbb{I})]\bigg)(Q(\phi^i_A)\otimes \mathbb{I}) \\
& \qquad - \bigg(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1}, (\mathbb{I}-(Q(\phi^i_A)\otimes \mathbb{I}))\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1}, (\mathbb{I}-(Q(\phi^i_A)\otimes \mathbb{I}))]\bigg)\{\mathbb{I}-(Q(\phi^i_A)\otimes \mathbb{I})\}\bigg]^2\Bigg) \\
& + \frac{\mu}{2}\mathrm{Tr}\bigg(\Big[\rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1}\Big]^2\bigg) \\
\text{s.t.} \quad & \mathrm{Tr}\left(\rho^{ij,AB}_{ABEA'_x} \cdot chsh(\phi^i_A, \phi^j_B)\right) = S_{ij} \\
& \phi^i_A, \phi^j_B \in [0, \pi/2], \\
& \rho^{ij,AB}_{ABEA'_x} \succeq 0 \\
& \mathrm{Tr}(\rho^{ij,AB}_{ABEA'_x}) = 1
\end{aligned}
\tag{8.6.82}
$$

Perturbing the parameter ϕ^i_A in $Q(\phi^i_A)$

The projector $Q(\cdot)$ is of dimension 2×2 and is being defined in (8.3.1) is an orthogonal projector onto the one-dimensional subspace spanned by the unit vector $v(\theta) = \begin{pmatrix}\cos\frac{\theta}{2} \\ \sin\frac{\theta}{2}\end{pmatrix}$.

$$
v(\theta) = \begin{pmatrix}\cos(\theta/2) \\ \sin(\theta/2)\end{pmatrix}, \quad Q(\theta) = v(\theta)\, v(\theta)^T. \tag{8.6.83}
$$

One can verify $Q(\theta)^2 = Q(\theta)$ (idempotent property) and $\mathrm{tr}(Q(\theta)) = 1$. In this form $Q(\theta)$ is the outer product of $v(\theta)$ with itself. For a smooth matrix-valued function $Q(\theta)$, the first-order Taylor expansion about θ under a small increment $\epsilon \ll 1$ is obtained entrywise from the usual Taylor theorem.

$$
Q(\theta+\epsilon) \;=\; Q(\theta) + \epsilon\, Q'(\theta) + \mathcal{O}(\epsilon^2), \tag{8.6.84}
$$

where $Q'(\theta)$ is the first order derivative of the projector Q Now, as $\epsilon \longrightarrow 0$, norm of the higher order terms in $\mathcal{O}(\epsilon^2)$ are bounded by $C\,\epsilon^2$. Thus, $Q(\phi+\epsilon)$ is linearly approximated up to first order derivative of $Q(\phi)$ with respect to θ in ϵ.

$$
Q(\theta+\epsilon) \approx Q(\theta) + \epsilon Q'(\theta) \tag{8.6.85}
$$

The first order derivative of the vector $v(\theta)$ is given as,

$$
v'(\theta) = \frac{d}{d\theta}\begin{pmatrix}\cos(\theta/2) \\ \sin(\theta/2)\end{pmatrix} = \begin{pmatrix}-\frac{1}{2}\sin(\theta/2) \\ \frac{1}{2}\cos(\theta/2)\end{pmatrix}. \tag{8.6.86}
$$

Now from the definition of $Q(\theta)$ in (8.6.83) and matrix product rule one have,

$$
Q'(\theta) \;=\; \frac{d}{d\theta}\big(v(\theta)v(\theta)^T\big) = v'(\theta)\, v(\theta)^T + v(\theta)\,\big(v'(\theta)\big)^T \tag{8.6.87}
$$

Now, from the definition of $v(\theta)$ and $v(\theta)^T$,

$$
\begin{aligned}
v(\theta)'v(\theta)^T \;&=\; \begin{pmatrix}-\frac{1}{2}\sin(\theta/2)\,\cos(\theta/2) & -\frac{1}{2}\sin(\theta/2)\,\sin(\theta/2) \\ \frac{1}{2}\cos(\theta/2)\,\cos(\theta/2) & \frac{1}{2}\cos(\theta/2)\,\sin(\theta/2)\end{pmatrix} \\
v(\theta)(v(\theta)')^T \;&=\; \begin{pmatrix}\cos(\theta/2)\,(-\frac{1}{2}\sin(\theta/2)) & \cos(\theta/2)\,\frac{1}{2}\cos(\theta/2) \\ \sin(\theta/2)\,(-\frac{1}{2}\sin(\theta/2)) & \sin(\theta/2)\,\frac{1}{2}\cos(\theta/2)\end{pmatrix}
\end{aligned}
\tag{8.6.88}
$$

Thus finally,

$$Q'(\theta) \;=\; v(\theta)'v(\theta)^T + v(\theta)(v(\theta)')^T = \begin{pmatrix} -\cos(\theta/2)\sin(\theta/2) & \frac{1}{2}(\cos^2(\theta/2)-\sin^2(\theta/2)) \\ \frac{1}{2}(\cos^2(\theta/2)-\sin^2(\theta/2)) & \cos(\theta/2)\sin(\theta/2) \end{pmatrix}. \tag{8.6.89}$$

Using double-angle identities $\sin\theta = 2\sin(\theta/2)\cos(\theta/2)$ and $\cos\theta = \cos^2(\theta/2) - \sin^2(\theta/2)$, this simplifies to the following compact form

$$Q'(\theta) \;=\; \frac{1}{2}\begin{pmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{pmatrix} \tag{8.6.90}$$

Putting it all together, the first-order Taylor approximation of $Q(\theta)$ is

$$\begin{aligned} Q(\theta+\epsilon) &\approx Q(\theta) + \epsilon\, Q'(\theta) \\ &\approx \begin{pmatrix} \cos^2\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) \\ \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) & \sin^2\left(\frac{\theta}{2}\right) \end{pmatrix} + \frac{\epsilon}{2}\begin{pmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{pmatrix} \\ &\approx \begin{pmatrix} \cos^2\left(\frac{\theta}{2}\right) - \frac{\epsilon\sin(\theta)}{2} & \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) + \frac{\epsilon\cos(\theta)}{2} \\ \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) + \frac{\epsilon\cos(\theta)}{2} & \sin^2\left(\frac{\theta}{2}\right) + \frac{\epsilon\sin(\theta)}{2} \end{pmatrix} \end{aligned} \tag{8.6.91}$$

After establishing an approximate value of $Q(\phi+\epsilon)$ one can proceed by expressing $(Q(\phi_A^i+\epsilon)\otimes\mathbb{I})\rho^{ij,AB}_{ABEA_0'}$ in first order linear approximation as,

$$(Q(\phi_A^i+\epsilon)\otimes\mathbb{I})\rho^{ij,AB}_{ABEA_0'} = (Q(\phi_A^i)\otimes\mathbb{I})\rho^{ij,AB}_{ABEA_0'} + \epsilon(Q'(\phi_A^i)\otimes\mathbb{I})\rho^{ij,AB}_{ABEA_0'} + \mathcal{O}(\epsilon^2) \tag{8.6.92}$$

and in terms of anti-commuting and commuting operators from the result of 8.6.5 and (8.6.84) as,

$$(Q(\phi_A^i+\epsilon)\otimes\mathbb{I})\rho^{ij,AB}_{ABEA_0'} = \frac{1}{2}\{\rho^{ij,AB}_{ABEA_0'}, (Q(\phi_A^i+\epsilon)\otimes\mathbb{I})\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA_0'}, (Q(\phi_A^i+\epsilon)\otimes\mathbb{I})] \tag{8.6.93}$$

Now expanding the commuting and the anti-commuting term individually as,

$$\begin{aligned} [\rho^{ij,AB}_{ABEA_0'}, (Q(\phi_A^i+\epsilon)\otimes\mathbb{I})] &= \rho^{ij,AB}_{ABEA_0'}(Q(\phi_A^i+\epsilon)\otimes\mathbb{I}) - (Q(\phi_A^i+\epsilon)\otimes\mathbb{I})\rho^{ij,AB}_{ABEA_0'} \\ &= \rho^{ij,AB}_{ABEA_0'}\Big((Q(\phi_A^i)\otimes\mathbb{I}) + \epsilon(Q'(\phi_A^i)\otimes\mathbb{I}) + \mathcal{O}(\epsilon^2)\Big) \\ &\qquad - \Big((Q(\phi_A^i)\otimes\mathbb{I}) + \epsilon(Q'(\phi_A^i)\otimes\mathbb{I}) + \mathcal{O}(\epsilon^2)\Big)\rho^{ij,AB}_{ABEA_0'} \\ &= \Big(\rho^{ij,AB}_{ABEA_0'}(Q(\phi_A^i)\otimes\mathbb{I}) - (Q(\phi_A^i)\otimes\mathbb{I})\rho^{ij,AB}_{ABEA_0'}\Big) \\ &\qquad + \epsilon\left(\rho^{ij,AB}_{ABEA_0'}(Q'(\phi_A^i)\otimes\mathbb{I}) - (Q'(\phi_A^i)\otimes\mathbb{I})\rho^{ij,AB}_{ABEA_0'}\right) + \mathcal{O}\left(\epsilon^2\right) \\ &= [\rho^{ij,AB}_{ABEA_0'}, (Q(\phi_A^i)\otimes\mathbb{I})] + \epsilon[\rho^{ij,AB}_{ABEA_0'}, (Q'(\phi_A^i)\otimes\mathbb{I})] + \mathcal{O}(\epsilon^2) \end{aligned} \tag{8.6.94}$$

Thus finally,

$$\begin{aligned} &[\rho^{ij,AB}_{ABEA_0'}, (Q(\phi_A^i+\epsilon)\otimes\mathbb{I})] = [\rho^{ij,AB}_{ABEA_0'}, (Q(\phi_A^i)\otimes\mathbb{I})] + \epsilon[\rho^{ij,AB}_{ABEA_0'}, (Q'(\phi_A^i)\otimes\mathbb{I})] + \mathcal{O}(\epsilon^2) \\ &\text{Similarly, } \{\rho^{ij,AB}_{ABEA_0'}, (Q(\phi_A^i+\epsilon)\otimes\mathbb{I})\} = \{\rho^{ij,AB}_{ABEA_0'}, (Q(\phi_A^i)\otimes\mathbb{I})\} + \epsilon\{\rho^{ij,AB}_{ABEA_0'}, (Q'(\phi_A^i)\otimes\mathbb{I})\} + \mathcal{O}(\epsilon^2) \\ &\text{Also, } [\rho^{ij,AB}_{ABEA_1'}, (\mathbb{I}-(Q(\phi_A^i)\otimes\mathbb{I}))] = [\rho^{ij,AB}_{ABEA_1'}, (\mathbb{I}-(Q(\phi_A^i)\otimes\mathbb{I}))] + \epsilon[\rho^{ij,AB}_{ABEA_1'}, (\mathbb{I}-(Q'(\phi_A^i)\otimes\mathbb{I}))] + \mathcal{O}(\epsilon^2) \\ &\{\rho^{ij,AB}_{ABEA_1'}, (\mathbb{I}-(Q(\phi_A^i)\otimes\mathbb{I}))\} = \{\rho^{ij,AB}_{ABEA_1'}, (\mathbb{I}-(Q(\phi_A^i)\otimes\mathbb{I}))\} + \epsilon\{\rho^{ij,AB}_{ABEA_1'}, (\mathbb{I}-(Q'(\phi_A^i)\otimes\mathbb{I}))\} + \mathcal{O}(\epsilon^2) \end{aligned} \tag{8.6.95}$$

Before proceeding further, let us have the notion behind the hermiticity of the commutator and anti-commutator of two hermitian operators.

Theorem 8.6.12. Commutator and anti-commutator of two Hermitian operators are anti-hermitian and hermitian respectively.

Proof. Let A and B be two Hermitian operators, $A = A^*$ and $B = B^*$.
The commutator of two operators A and B is defined as:

$$[A, B] = AB - BA. \tag{8.6.96}$$

Taking the adjoint,

$$\begin{aligned} [A, B]^* &= (AB - BA)^* \\ &= (AB)^* - (BA)^* \\ &= B^*A^* - A^*B^* \\ &= BA - AB \\ &= -[A, B]. \end{aligned} \tag{8.6.97}$$

Since $[A, B]^* = -[A, B]$, the commutator of two Hermitian operators is anti-Hermitian.
The anti-commutator of two operators A and B is defined as:

$$\{A, B\} = AB + BA. \tag{8.6.98}$$

Taking the adjoint, we get:

$$\begin{aligned} \{A, B\}^* &= (AB + BA)^* \\ &= (AB)^* + (BA)^* \\ &= B^*A^* + A^*B^* \\ &= BA + AB \\ &= \{A, B\}. \end{aligned} \tag{8.6.99}$$

□

Now, let us consider the following function

$$h : (\lambda, \phi^i_A, \rho^{ij,AB}_{ABEA'_0}) \mapsto \mathbb{R} \tag{8.6.100}$$

being defined as follows

$$\begin{aligned} h(\lambda, \phi^i_A, \rho^{ij,AB}_{ABEA'_0}) = (1-\lambda)\mathrm{Tr}\Bigg(\Bigg[\rho^{ij,AB}_{ABEA'_1} - \left(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A) \otimes \mathbb{I})\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A) \otimes \mathbb{I})]\right)(Q(\phi^i_A) \otimes \mathbb{I}) \\ - \left(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1}, (\mathbb{I} - (Q(\phi^i_A) \otimes \mathbb{I}))\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1}, (\mathbb{I} - (Q(\phi^i_A) \otimes \mathbb{I}))]\right)\{\mathbb{I} - (Q(\phi^i_A) \otimes \mathbb{I})\}\Bigg]^2\Bigg) \end{aligned} \tag{8.6.101}$$

Now, one can bound the perturbation in ϕ^i_A by small ϵ using triangular inequality as,

$$\begin{aligned}
&\left|h(\lambda,\phi^i_A,\rho^{ij,AB}_{ABEA'_0}) - h(\lambda,\phi^i_A+\epsilon,\rho^{ij,AB}_{ABEA'_0})\right|\\
\leq&\left|h(\lambda,\phi^i_A,\rho^{ij,AB}_{ABEA'_0})\right| + \left|h(\lambda,\phi^i_A+\epsilon,\rho^{ij,AB}_{ABEA'_0})\right|\\
\leq&(1-\lambda)\mathrm{Tr}\left(\left[\rho^{ij,AB}_{ABEA'_1} - \left(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1},(Q(\phi^i_A)\otimes\mathbb{I})\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1},(Q(\phi^i_A)\otimes\mathbb{I})]\right)(Q(\phi^i_A)\otimes\mathbb{I})\right.\right.\\
&\qquad\left.\left. - \left(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1},(\mathbb{I}-(Q(\phi^i_A)\otimes\mathbb{I}))\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1},(\mathbb{I}-(Q(\phi^i_A)\otimes\mathbb{I}))]\right)\{\mathbb{I}-(Q(\phi^i_A)\otimes\mathbb{I})\}\right]^2\right)\\
&+(1-\lambda)\mathrm{Tr}\left(\left[\rho^{ij,AB}_{ABEA'_1} - \left(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1},(Q(\phi^i_A+\epsilon)\otimes\mathbb{I})\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1},(Q(\phi^i_A+\epsilon)\otimes\mathbb{I})]\right)(Q(\phi^i_A+\epsilon)\otimes\mathbb{I})\right.\right.\\
&\qquad\left.\left. - \left(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1},(\mathbb{I}-(Q(\phi^i_A+\epsilon)\otimes\mathbb{I}))\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1},(\mathbb{I}-(Q(\phi^i_A+\epsilon)\otimes\mathbb{I}))]\right)\{\mathbb{I}-(Q(\phi^i_A+\epsilon)\otimes\mathbb{I})\}\right]^2\right)\\
\leq&(1-\lambda)\Bigg[\mathrm{Tr}\left(\left[\rho^{ij,AB}_{ABEA'_1} - \left(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1},(Q(\phi^i_A)\otimes\mathbb{I})\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1},(Q(\phi^i_A)\otimes\mathbb{I})]\right)(Q(\phi^i_A)\otimes\mathbb{I})\right.\right.\\
&\qquad\left. - \left(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1},(\mathbb{I}-(Q(\phi^i_A)\otimes\mathbb{I}))\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1},(\mathbb{I}-(Q(\phi^i_A)\otimes\mathbb{I}))]\right)(\mathbb{I}-(Q(\phi^i_A)\otimes\mathbb{I}))\right]^2\\
&\qquad+\left[\rho^{ij,AB}_{ABEA'_1} - \left(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1},(Q(\phi^i_A+\epsilon)\otimes\mathbb{I})\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1},(Q(\phi^i_A+\epsilon)\otimes\mathbb{I})]\right)(Q(\phi^i_A+\epsilon)\otimes\mathbb{I})\right.\\
&\qquad\left.\left. - \left(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1},(\mathbb{I}-(Q(\phi^i_A+\epsilon)\otimes\mathbb{I}))\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1},(\mathbb{I}-(Q(\phi^i_A+\epsilon)\otimes\mathbb{I}))]\right)(\mathbb{I}-(Q(\phi^i_A+\epsilon)\otimes\mathbb{I}))\right]^2\right)\Bigg]
\end{aligned}\tag{8.6.102}$$

To simplify the calculation, we are simplifying the perturbing and non-perturbing steps separately.

$$\begin{aligned}
\text{Let A} =& \left[\rho^{ij,AB}_{ABEA'_1} - \left(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1},(Q(\phi^i_A)\otimes\mathbb{I})\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1},(Q(\phi^i_A)\otimes\mathbb{I})]\right)(Q(\phi^i_A)\otimes\mathbb{I})\right.\\
&\left. - \left(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1},(\mathbb{I}-(Q(\phi^i_A)\otimes\mathbb{I}))\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1},(\mathbb{I}-(Q(\phi^i_A)\otimes\mathbb{I}))]\right)(\mathbb{I}-(Q(\phi^i_A)\otimes\mathbb{I}))\right]^2\\
\text{B} =& \left[\rho^{ij,AB}_{ABEA'_1} - \left(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1},(Q(\phi^i_A+\epsilon)\otimes\mathbb{I})\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1},(Q(\phi^i_A+\epsilon)\otimes\mathbb{I})]\right)(Q(\phi^i_A+\epsilon)\otimes\mathbb{I})\right.\\
&\left. - \left(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1},(\mathbb{I}-(Q(\phi^i_A+\epsilon)\otimes\mathbb{I}))\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1},(\mathbb{I}-(Q(\phi^i_A+\epsilon)\otimes\mathbb{I}))]\right)(\mathbb{I}-(Q(\phi^i_A+\epsilon)\otimes\mathbb{I}))\right]^2
\end{aligned}\tag{8.6.103}$$

Now solving the A and B separately,

$$
\begin{aligned}
\mathrm{A} =& \Big[\rho^{ij,AB}_{ABEA'_1} - \Big(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A) \otimes \mathbb{I})\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A) \otimes \mathbb{I})]\Big)(Q(\phi^i_A) \otimes \mathbb{I}) \\
&\quad - \Big(\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1}, (\mathbb{I} - (Q(\phi^i_A) \otimes \mathbb{I}))\} - \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1}, (\mathbb{I} - (Q(\phi^i_A) \otimes \mathbb{I}))]\Big)(\mathbb{I} - (Q(\phi^i_A) \otimes \mathbb{I}))\Big]^2 \\
=& \Big[\rho^{ij,AB}_{ABEA'_1} - \frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A) \otimes \mathbb{I})\}(Q(\phi^i_A) \otimes \mathbb{I}) + \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A) \otimes \mathbb{I})](Q(\phi^i_A) \otimes \mathbb{I}) \\
&\quad -\frac{1}{2}\{\rho^{ij,AB}_{ABEA'_1}, (\mathbb{I} - (Q(\phi^i_A) \otimes \mathbb{I}))\}(\mathbb{I} - (Q(\phi^i_A) \otimes \mathbb{I})) + \frac{1}{2}[\rho^{ij,AB}_{ABEA'_1}, (\mathbb{I} - (Q(\phi^i_A) \otimes \mathbb{I}))](\mathbb{I} - (Q(\phi^i_A) \otimes \mathbb{I}))\Big]^2 \\
=& \Big[\rho^{ij,AB}_{ABEA'_1} - (Q(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}(Q(\phi^i_A) \otimes \mathbb{I}) - (\mathbb{I} - (Q(\phi^i_A) \otimes \mathbb{I}))\rho^{ij,AB}_{ABEA'_1}(\mathbb{I} - (Q(\phi^i_A) \otimes \mathbb{I}))\Big]^2 \\
=& \Big[\rho^{ij,AB}_{ABEA'_1}(Q(\phi^i_A) \otimes \mathbb{I}) + (Q(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1} - 2(Q(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}(Q(\phi^i_A) \otimes \mathbb{I})\Big]^2 \\
=& \Big[-(\rho^{ij,AB}_{ABEA'_1}(Q(\phi^i_A) \otimes \mathbb{I}))^2 - ((Q(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1})^2 + \rho^{ij,AB}_{ABEA'_1}(Q(\phi^i_A) \otimes \mathbb{I})(Q(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1} \\
&\quad +(Q(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1} \cdot \rho^{ij,AB}_{ABEA'_1}(Q(\phi^i_A) \otimes \mathbb{I})\Big] \\
=& -\Big[\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A) \otimes \mathbb{I})\Big]^2
\end{aligned}
\tag{8.6.104}
$$

Similarly, one can find the value for B by using the above result

$$
\begin{aligned}
\mathrm{B} =& \Big[\rho^{ij,AB}_{ABEA'_1} - \Big(\frac{1}{2}\Big\{\rho^{ij,AB}_{ABEA'_1}, Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\Big\} - \frac{1}{2}\Big[\rho^{ij,AB}_{ABEA'_1}, Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\Big]\Big)(Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}) \\
&\quad - \Big(\frac{1}{2}\Big\{\rho^{ij,AB}_{ABEA'_1}, \mathbb{I} - Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\Big\} - \frac{1}{2}\Big[\rho^{ij,AB}_{ABEA'_1}, \mathbb{I} - Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\Big]\Big)(\mathbb{I} - Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I})\Big]^2 \\
=& -\Big[\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I})\Big]^2 \\
=& -\Big[[\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A) \otimes \mathbb{I})] + \epsilon_0[\rho^{ij,AB}_{ABEA'_1}, (Q'(\phi^i_A) \otimes \mathbb{I})] + \mathcal{O}(\epsilon_0^2)\Big]^2 \\
=& -\Big[[\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A) \otimes \mathbb{I})] + \epsilon_0[\rho^{ij,AB}_{ABEA'_1}, (Q'(\phi^i_A) \otimes \mathbb{I})]\Big]^2 + \mathcal{O}(\epsilon_0^2) \\
=& -\Big([\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A) \otimes \mathbb{I})]^2 + \epsilon_0\{[\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A) \otimes \mathbb{I})], [\rho^{ij,AB}_{ABEA'_1}, (Q'(\phi^i_A) \otimes \mathbb{I})]\} \\
&\quad +\epsilon_0^2 \cdot [\rho^{ij,AB}_{ABEA'_1}, (Q'(\phi^i_A) \otimes \mathbb{I})]^2\Big) + \mathcal{O}(\epsilon_0^2)
\end{aligned}
\tag{8.6.105}
$$

With explicit forms of A and B, one can proceed to (8.6.102) as,

$$\begin{aligned}
&\left|h(\lambda,\phi^i_A,\rho^{ij,AB}_{ABEA_1'}) - h(\lambda,\phi^i_A+\epsilon_0,\rho^{ij,AB}_{ABEA_1'})\right| \\
&\leq (1-\lambda)\left[\mathrm{Tr}\left(-[\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbf{I})]^2-\epsilon_0\{[\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbf{I})],[\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbf{I})]\}\right.\right. \\
&\qquad\left.\left.-\epsilon_0^2\cdot[\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbf{I})]^2-\left[\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbf{I})\right]^2\right)\right] \\
&\leq (1-\lambda)\left[\mathrm{Tr}\left(-\epsilon_0\{[\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbf{I})],[\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbf{I})]\}-\epsilon_0^2\cdot[\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbf{I})]^2\right.\right. \\
&\qquad\left.\left.-2[\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbf{I})]^2\right)\right] \\
&\leq (1-\lambda)\left[-\epsilon_0\,\mathrm{Tr}\left(\{[\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbf{I})],[\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbf{I})]\}\right)-\epsilon_0^2\,\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbf{I})]^2\right)\right. \\
&\qquad\left.-\,\mathrm{Tr}\left(2[\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbf{I})]^2\right)\right] \\
&\leq (1-\lambda)\left[-2\epsilon_0\,\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbf{I})]\cdot[\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbf{I})]\right)-\epsilon_0^2\,\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbf{I})]^2\right)\right. \\
&\qquad\left.-\,\mathrm{Tr}\left(2[\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbf{I})]^2\right)\right]
\end{aligned}\tag{8.6.106}$$

Now, one can apply Hilbert Schmidt inner product[55] on $\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbb{I})]\cdot[\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})]\right)$ as,

$$\begin{aligned}
&\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbb{I})]\cdot[\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})]\right) \\
=&\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})]\cdot[\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbb{I})]\right) \\
=&\quad\left\langle[\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbb{I})],[\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})]\right\rangle
\end{aligned}\tag{8.6.107}$$

The above inner product holds because cyclic property of trace and the product of $[\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbb{I})]$ and $[\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})]$ being self adjoint because product of two anti hermitian is hermitian. Now one can use the holder inequality [55] to bound the inner product as,

$$\begin{aligned}
&\left\langle[\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbb{I})],[\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})]\right\rangle \\
\leq&\left\|[\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbb{I})]\right\|_2\cdot\left\|[\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})]\right\|_2 \\
\leq&\sqrt{\mathrm{Tr}([\rho^{ij,AB}_{ABEA_1'},(Q(\phi^i_A)\otimes\mathbb{I})]^2)\cdot\mathrm{Tr}([\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})]^2)}
\end{aligned}\tag{8.6.108}$$

On a similar note on applying Hilbert Schmidt followed by Holder inequality on$\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})]^2\right)$ as,

$$\begin{aligned}
&\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})]^2\right) \\
&=\left\langle[\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})],[\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})]\right\rangle \\
&\leq\sqrt{\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})]^2\cdot\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})]^2\right)\right)} \\
&\leq\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA_1'},(Q'(\phi^i_A)\otimes\mathbb{I})]^2\right)
\end{aligned}\tag{8.6.109}$$

Thus,

$$\begin{aligned}
&\left|h(\lambda, \phi_A^i, \rho^{ij,AB}_{ABEA'_1}) - h(\lambda, \phi_A^i + \epsilon_0, \rho^{ij,AB}_{ABEA'_1})\right| \\
\leq &(1-\lambda)[-2.|\epsilon_0|.\sqrt{\mathrm{Tr}([\rho^{ij,AB}_{ABEA'_1}, (Q(\phi_A^i)\otimes \mathbb{I})]^2)\cdot \mathrm{Tr}([\rho^{ij,AB}_{ABEA'_1}, (Q'(\phi_A^i)\otimes \mathbb{I})]^2)} - |\epsilon_0^2|.\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA'_1}, (Q'(\phi_A^i)\otimes \mathbb{I})]^2\right. \\
&\qquad \left. -2\mathrm{Tr}([\rho^{ij,AB}_{ABEA'_1}, (Q'(\phi_A^i)\otimes \mathbb{I})]^2)\right)]
\end{aligned} \tag{8.6.110}$$

Now one can bound the term $\mathrm{Tr}([\rho^{ij,AB}_{ABEA'_1}, (Q(\phi_A^i)\otimes \mathbb{I})]^2)$ and $\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA'_1}, (Q'(\phi_A^i)\otimes \mathbb{I})]^2\right)$ from the above equation as from the definition of Frobenius norm and hermiticity of product of two anti-hermitian one can have,

$$\begin{aligned}
&\mathrm{Tr}([\rho^{ij,AB}_{ABEA'_1}, (Q(\phi_A^i)\otimes \mathbb{I})]) \\
&= \|[\rho^{ij,AB}_{ABEA'_1}, (Q(\phi_A^i)\otimes \mathbb{I})]\|_2 \\
&= \|(\rho^{ij,AB}_{ABEA'_1}\cdot (Q(\phi_A^i)\otimes \mathbb{I}) + (-(Q(\phi_A^i)\otimes \mathbb{I})\cdot \rho^{ij,AB}_{ABEA'_1}))\|_2 \\
&\leq \|(\rho^{ij,AB}_{ABEA'_1}\cdot (Q(\phi_A^i)\otimes \mathbb{I})\|_2 + \|(Q(\phi_A^i)\otimes \mathbb{I})\cdot \rho^{ij,AB}_{ABEA'_1})\|_2 \\
&\leq \|(\rho^{ij,AB}_{ABEA'_1})\|_2\cdot \|(Q(\phi_A^i)\otimes \mathbb{I})\|_\infty + \|(Q(\phi_A^i)\otimes \mathbb{I})\|_\infty\cdot \|(\rho^{ij,AB}_{ABEA'_1})\|_2 \\
&\leq 2\|(\rho^{ij,AB}_{ABEA'_1})\|_2\cdot \|Q(\phi_A^i)\|_\infty \\
&\leq 2.\mathrm{Tr}(\rho^{ij,AB}_{ABEA'_1})\cdot \|Q(\phi_A^i)\|_\infty \\
&\leq 2
\end{aligned} \tag{8.6.111}$$

The above relation holds because $\rho^{ij,AB}_{ABEA'_1}$, being a density operator, its trace is 1 and $Q(\phi_A^i)$, being a rank one projector, its spectral norm is 1. Further, here two other important inequalities have been used, viz., the triangular inequality[55] and mixed norm submultiplicativity [30]. Thus

$$\mathrm{Tr}([\rho^{ij,AB}_{ABEA'_1}, (Q(\phi_A^i)\otimes \mathbb{I})]^2) \leq 4 \tag{8.6.112}$$

Now, for $\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA'_1}, (Q'(\phi_A^i)\otimes \mathbb{I})]^2\right)$ one can show its being bounded by 1 due to $Q'(\phi_A^i)$ being not a projector but a Hermitian operator that can be expressed as a sum of two rank-1 operators.

Lemma 8.6.13. $Q'(\phi_A^i)$ is a hermitian but not a projector and its spectral norm is bounded by 1.

Proof. a. Hermiticity: For $Q'(\phi_A^i)$, the Hermitian conjugate is given as,

$$\begin{aligned}
(Q'(\phi_A^i))^* &= \left(|\psi'(\phi_A^i)\rangle\langle\psi(\phi_A^i)|\right)^* + \left(|\psi(\phi_A^i)\rangle\langle\psi'(\phi_A^i)|\right)^* \\
&= |\psi(\phi_A^i)\rangle\langle\psi'(\phi_A^i)| + |\psi'(\phi_A^i)\rangle\langle\psi(\phi_A^i)| \\
&= Q'(\phi_A^i)
\end{aligned} \tag{8.6.113}$$

Thus, $Q'(\phi_A^i)$ is a Hermitian operator. *b. $Q'(\phi_A^i)$ is not a projector:* Expanding $Q'(\phi_A^i)^2$:

$$\begin{aligned}
Q'(\phi_A^i)^2 &= \left(|\psi'(\phi_A^i)\rangle\langle\psi(\phi_A^i)| + |\psi(\phi_A^i)\rangle\langle\psi'(\phi_A^i)|\right)^2 \\
&= (|\psi'(\phi_A^i)\rangle\langle\psi(\phi_A^i)||\psi'(\phi_A^i)\rangle\langle\psi(\phi_A^i)| + |\psi'(\phi_A^i)\rangle\langle\psi(\phi_A^i)||\psi(\phi_A^i)\rangle\langle\psi'(\phi_A^i)| \\
&\quad + |\psi(\phi_A^i)\rangle\langle\psi'(\phi_A^i)||\psi'(\phi_A^i)\rangle\langle\psi(\phi_A^i)| + |\psi(\phi_A^i)\rangle\langle\psi'(\phi_A^i)||\psi(\phi_A^i)\rangle\langle\psi'(\phi_A^i)|) \\
&= \left(0 + |\psi'(\phi_A^i)\rangle\langle\psi'(\phi_A^i)| + |\psi(\phi_A^i)\rangle\langle\psi(\phi_A^i)| + 0\right) \\
&\neq (|\psi'(\phi_A^i)\rangle\langle\psi(\phi_A^i)| + |\psi(\phi_A^i)\rangle\langle\psi'(\phi_A^i)|) \\
&\neq Q'(\phi_A^i)
\end{aligned} \tag{8.6.114}$$

Thus, $Q'(\phi_A^i)$ is not a projector.

b. Spectral Norm Analysis: The spectral norm $\|A\|_\infty$ of an operator A is defined as the largest singular value or the largest absolute value of the eigenvalues. For $Q'(\phi_A^i)$, observe that it is a sum of two rank-1 operators.

$$Q'(\phi_A^i) = A + A^* \tag{8.6.115}$$

where $A = |\psi'(\phi_A^i)\rangle\langle\psi(\phi_A^i)|$ The spectral norm of A is given by:

$$\|A\|_\infty = \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\|=1} |\langle x|A|x\rangle| \tag{8.6.116}$$

Since A is a rank-1 operator, its spectral norm is at most 1 (assuming normalized states). Since $Q'(\phi_A^i)$ is a sum of two such rank-1 operators and each has a norm of at most 1, the spectral norm is given as,

$$\|Q'(\phi_A^i)\|_\infty \leq \|A\|_\infty + \|A^*\|_\infty \leq 1 + 1 = 2 \tag{8.6.117}$$

However, due to the symmetry and specific structure of $Q'(\phi_A^i)$, the norm is actually tighter and can be bounded as,

$$\|Q'(\phi_A^i)\|_\infty \leq 1 \tag{8.6.118}$$

The above result holds because the two components A and A^* share the same singular value structure, and the combined operator's action is constrained. Moreover, by inspecting the structure of $Q'(\phi_A^i)$ from (8.6.87), one can easily verify that its eigenvalues are coming as $\sin(\phi_A^i)$, whose largest value is 1. Hence,

$$\|Q'(\phi_A^i)\|_\infty \leq 1 \tag{8.6.119}$$

This tighter bound is a consequence of the Hermitian structure of $Q'(\phi_A^i)$ and the fact that its constituent rank-one operators act on a shared, low-dimensional subspace, leading to eigenvalues bounded by ±1. □

Singular Value Analysis of $Q'(\phi_A^i)$

The singular values of the components A and A^* are equal because they are related by Hermitian conjugation. Each component has at most one singular value equal to 1, as they are rank-1 operators. However, when forming the sum $A + A^*$, the resulting matrix can no longer be treated as two independent rank-1 components. Instead, they act on the same subspace, effectively sharing the same eigenvector structure. Since both A and A^* project onto the same 1-dimensional subspace spanned by $|\psi(\phi_A^i)\rangle$ and $|\psi'(\phi_A^i)\rangle$, the maximum eigenvalue of the combined operator $Q'(\phi_A^i)$ cannot exceed 1.
Geometrically, the action of $Q'(\phi_A^i)$ can be visualised as a combination of two reflections or projections onto overlapping subspaces. Since these subspaces are not orthogonal, the combined action is not a simple sum but a weighted interaction within the same span, resulting in a norm bound of 1 rather than 2.

Lemma 8.6.14. $\mathrm{Tr}([\rho^{ij,AB}_{ABEA_1'}, (Q(\phi_A^i) \otimes \mathbb{I})]^2) \leq 4$ and $\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA_1'}, (Q'(\phi_A^i) \otimes \mathbb{I})]^2\right) \leq 1$.

Proof. **1. Structure and Properties of $Q(\phi_A^i)$ and $Q'(\phi_A^i)$**

1. Projector $Q(\phi_A^i)$ The projector can can be decomposed as,

$$Q(\phi_A^i) = |\psi(\phi_A^i)\rangle\langle\psi(\phi_A^i)| \tag{8.6.120}$$

where $Q(\phi_A^i)$ is a rank-1 Hermitian projector satisfying $Q(\phi_A^i)^2 = Q(\phi_A^i)$, and $\|Q(\phi_A^i)\|_\infty = 1$.

2. Derivative $Q'(\phi_A^i)$ The derivative $Q'(\phi_A^i)$ is given as,

$$Q'(\phi_A^i) = |\psi'(\phi_A^i)\rangle\langle\psi(\phi_A^i)| + |\psi(\phi_A^i)\rangle\langle\psi'(\phi_A^i)| \tag{8.6.121}$$

Now, from 8.6.5 $Q'(\phi_A^i)$ is not a projector but a Hermitian operator that can be expressed as a sum of two rank-1 operators and the spectral norm is $\|Q'(\phi_A^i)\|_\infty \leq 1$.

2. Commutator Analysis for $[\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A) \otimes \mathbb{I})]$

a. Commutator Structure

The commutator is being defined as

$$[\rho^{ij,AB}_{ABEA'_1}, Q(\phi^i_A) \otimes \mathbb{I}] = \rho^{ij,AB}_{ABEA'_1}(Q(\phi^i_A) \otimes \mathbb{I}) - (Q(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1} \tag{8.6.122}$$

Now upon squaring and expanding the commutator,

$$\begin{aligned}[\rho^{ij,AB}_{ABEA'_1}, Q(\phi^i_A) \otimes \mathbb{I}]^2 = {} & \rho^{ij,AB}_{ABEA'_1}(Q(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}(Q(\phi^i_A) \otimes \mathbb{I}) \\ & - \rho^{ij,AB}_{ABEA'_1}(Q(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1} \\ & - (Q(\phi^i_A) \otimes \mathbb{I})(\rho^{ij,AB}_{ABEA'_1})^2(Q(\phi^i_A) \otimes \mathbb{I}) \\ & + (Q(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}(Q(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}\end{aligned} \tag{8.6.123}$$

b. Bounding the Trace

Now, finding the Frobenius norm of (8.6.122) and applying the triangle inequality as,

$$\|[\rho^{ij,AB}_{ABEA'_1}, Q(\phi^i_A) \otimes \mathbb{I}]\|_2 \leq \|\rho^{ij,AB}_{ABEA'_1}(Q(\phi^i_A) \otimes \mathbb{I})\|_2 + \|(Q(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}\|_2 \tag{8.6.124}$$

Now using mixed norm submultiplicativity, $\|AB\|_2 \leq \|A\|_2\|B\|_\infty$ on the two terms in RHS of (8.6.124) one has,

$$\|\rho^{ij,AB}_{ABEA'_1}(Q(\phi^i_A) \otimes \mathbb{I})\|_2 \leq \|\rho^{ij,AB}_{ABEA'_1}\|_2\|Q(\phi^i_A) \otimes \mathbb{I}\|_\infty = \|\rho^{ij,AB}_{ABEA'_1}\|_2\|Q(\phi^i_A)\|_\infty\|\mathbb{I}\|_\infty = \|\rho^{ij,AB}_{ABEA'_1}\|_2 \cdot 1 \cdot 1 = \|\rho^{ij,AB}_{ABEA'_1}\|_2$$

$$\|(Q(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}\|_2 \leq \|Q(\phi^i_A) \otimes \mathbb{I}\|_\infty\|\rho^{ij,AB}_{ABEA'_1}\|_2 = \|Q(\phi^i_A)\|_\infty\|\mathbb{I}\|_\infty\|\rho^{ij,AB}_{ABEA'_1}\|_2 = 1 \cdot 1 \cdot \|\rho^{ij,AB}_{ABEA'_1}\|_2 = \|\rho^{ij,AB}_{ABEA'_1}\|_2 \tag{8.6.125}$$

Thus, the from (8.6.124) and (8.6.125) commutator norm is bounded as,

$$\|[\rho^{ij,AB}_{ABEA'_1}, Q(\phi^i_A) \otimes \mathbb{I}]\|_2 \leq 2\|\rho^{ij,AB}_{ABEA'_1}\|_2 \tag{8.6.126}$$

Since $\|\rho^{ij,AB}_{ABEA'_1}\|_2^2 = \mathrm{Tr}((\rho^{ij,AB}_{ABEA'_1})^2) \leq 1$, we have,

$$\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA'_1}, Q(\phi^i_A) \otimes \mathbb{I}]^2\right) = \|[\rho^{ij,AB}_{ABEA'_1}, Q(\phi^i_A) \otimes \mathbb{I}]\|_2^2 \leq (2\|\rho^{ij,AB}_{ABEA'_1}\|_2)^2 = 4\|\rho^{ij,AB}_{ABEA'_1}\|_2^2 \leq 4 \cdot 1 = 4 \tag{8.6.127}$$

3. Commutator Analysis for $[\rho^{ij,AB}_{ABEA'_1}, (Q'(\phi^i_A) \otimes \mathbb{I})]$

The structure of the Derivative is given as

$$Q'(\phi^i_A) = |\psi'(\phi^i_A)\rangle\langle\psi(\phi^i_A)| + |\psi(\phi^i_A)\rangle\langle\psi'(\phi^i_A)| \tag{8.6.128}$$

This sum of two rank-1 operators, each with spectral norm at most 1 (assuming normalized states and derivatives) is a hermitian but not a projector from the result of 8.6.5.

a. Commutator Structure

The commutator term involving the derivative,$Q'(\phi^i_A)$ is being defined as,

$$[\rho^{ij,AB}_{ABEA'_1}, Q'(\phi^i_A)\otimes \mathbb{I}] = \rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A)\otimes \mathbb{I}) - (Q'(\phi^i_A)\otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1} \tag{8.6.129}$$

Now, similarly squaring and expanding the commutator as,

$$\begin{aligned}
[\rho^{ij,AB}_{ABEA'_1}, Q'(\phi^i_A)\otimes \mathbb{I}]^2 = {} & \rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A)\otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A)\otimes \mathbb{I}) \\
& - \rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A)\otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1} \\
& - (Q'(\phi^i_A)\otimes \mathbb{I})(\rho^{ij,AB}_{ABEA'_1})^2(Q'(\phi^i_A)\otimes \mathbb{I}) \\
& + (Q'(\phi^i_A)\otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A)\otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}
\end{aligned} \tag{8.6.130}$$

Since $Q'(\phi^i_A)$ is not a projector, the structure is less constrained. However, each term involves operators whose spectral norms are bounded.

b. Bounding the Trace

Applying the triangle inequality

$$\|[\rho^{ij,AB}_{ABEA'_1}, Q'(\phi^i_A)\otimes \mathbb{I}]\|_2 \leq \|\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A)\otimes \mathbb{I})\|_2 + \|(Q'(\phi^i_A)\otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}\|_2 \tag{8.6.131}$$

Now using mixed norm submultiplicavity similarly as earlier on the terms in RHS of (8.6.131) one can have,

$$\begin{aligned}
&\|\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A)\otimes \mathbb{I})\|_2 \leq \|\rho^{ij,AB}_{ABEA'_1}\|_2\|Q'(\phi^i_A)\otimes \mathbb{I}\|_\infty = \|\rho^{ij,AB}_{ABEA'_1}\|_2\|Q'(\phi^i_A)\|_\infty\|\mathbb{I}\|_\infty \leq \|\rho^{ij,AB}_{ABEA'_1}\|_2\cdot 1\cdot 1 = \|\rho^{ij,AB}_{ABEA'_1}\|_2 \\
&\|(Q'(\phi^i_A)\otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}\|_2 \leq \|Q'(\phi^i_A)\otimes \mathbb{I}\|_\infty\|\rho^{ij,AB}_{ABEA'_1}\|_2 = \|Q'(\phi^i_A)\|_\infty\|\mathbb{I}\|_\infty\|\rho^{ij,AB}_{ABEA'_1}\|_2 \leq 1\cdot 1\cdot\|\rho^{ij,AB}_{ABEA'_1}\|_2 = \|\rho^{ij,AB}_{ABEA'_1}\|_2
\end{aligned} \tag{8.6.132}$$

Thus, from (8.6.132) and (8.6.131) the commutator norm is bounded as,

$$\|[\rho^{ij,AB}_{ABEA'_1}, Q'(\phi^i_A)\otimes \mathbb{I}]\|_2 \leq 2\|\rho^{ij,AB}_{ABEA'_1}\|_2 \tag{8.6.133}$$

However, the bound for the trace norm of the square of the commutator involving $Q'(\phi^i_A)$ is tighter than just squaring this bound.

c. Tighter bound on Tr $\left([\rho^{ij,AB}_{ABEA'_1}, (Q'(\phi^i_A)\otimes \mathbb{I})]^2\right)$

To tighten the bound, we can exploit the fact that Q' is derived from Q. Specifically,

$$Q(\theta) = v(\theta)v(\theta)^T, \quad Q'(\theta) = v'(\theta)v(\theta)^T + v(\theta)v'(\theta)^T, \tag{8.6.134}$$

where $v'(\theta) = \frac{1}{2}\begin{pmatrix} -\sin(\theta/2) \\ \cos(\theta/2) \end{pmatrix}$ and $v(\theta) = \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}$

Now, observe that, $v(\theta)$ and $v'(\theta)$ are orthogonal: $v(\theta)^T v'(\theta) = 0$. This implies that $Q'(\theta)$ is a sum of two terms that are "off-diagonal" with respect to $Q(\theta)$.

This orthogonality can be used to show that the action of Q' on ρ is constrained in a way that reduces the norm of the commutator. Specifically: The commutator $[\rho, Q'\otimes \mathbb{I}]$ can be seen as a combination of terms where Q' "mixes" the subspaces defined by Q. Due to the orthogonality of v and v', this mixing is limited, leading to a smaller norm.

Explicit Calculation: Let us compute $\mathrm{Tr}([\rho^{ij,AB}_{ABEA'_1}, Q'(\phi^i_A) \otimes \mathbb{I}]^2)$ and using the cyclic property of the trace:

$$\begin{aligned}\mathrm{Tr}([\rho^{ij,AB}_{ABEA'_1}, Q'(\phi^i_A) \otimes \mathbb{I}]^2) &= \mathrm{Tr}\left((\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A) \otimes \mathbb{I}) - (Q'(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1})^2\right) \\ &= \mathrm{Tr}\left(\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A) \otimes \mathbb{I}) - \rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A) \otimes \mathbb{I})^2\rho^{ij,AB}_{ABEA'_1}\right. \\ &\qquad \left. -(Q'(\phi^i_A) \otimes \mathbb{I})(\rho^{ij,AB}_{ABEA'_1})^2(Q'(\phi^i_A) \otimes \mathbb{I}) + (Q'(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}\right) \\ &= 2\mathrm{Tr}\left(\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A) \otimes \mathbb{I})\right) - 2\mathrm{Tr}\left((\rho^{ij,AB}_{ABEA'_1})^2(Q'(\phi^i_A) \otimes \mathbb{I})^2\right)\end{aligned} \tag{8.6.135}$$

Now, using $\|Q'(\phi^i_A)\|_\infty \leq 1$ and $\|\rho^{ij,AB}_{ABEA'_1}\|_2 \leq 1$, we can bound each term:

- The first term satisfies:

$$\mathrm{Tr}\left(\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A) \otimes \mathbb{I})\right) \leq \|\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A) \otimes \mathbb{I})\|_2^2 \leq \|\rho^{ij,AB}_{ABEA'_1}\|_2^2\|Q'(\phi^i_A) \otimes \mathbb{I}\|_\infty^2 \leq 1. \tag{8.6.136}$$

- The second term satisfies:

$$\mathrm{Tr}\left((\rho^{ij,AB}_{ABEA'_1})^2(Q'(\phi^i_A) \otimes \mathbb{I})^2\right) \leq \|(\rho^{ij,AB}_{ABEA'_1})^2\|_1\|(Q'(\phi^i_A) \otimes \mathbb{I})^2\|_\infty \leq \|\rho^{ij,AB}_{ABEA'_1}\|_2^2\|Q'(\phi^i_A)\|_\infty^2 \leq 1. \tag{8.6.137}$$

Thus, the overall bound is:

$$\mathrm{Tr}([\rho^{ij,AB}_{ABEA'_1}, Q'(\phi^i_A) \otimes \mathbb{I}]^2) \leq 2 \cdot 1 - 2 \cdot 0 = 2, \tag{8.6.138}$$

where we have used that $\mathrm{Tr}((\rho^{ij,AB}_{ABEA'_1})^2(Q'(\phi^i_A) \otimes \mathbb{I})^2)$ is non-negative and could be small due to the structure of $Q'(\phi^i_A)$. However, this still gives a bound of 2, not 1. To tighten further, we need to exploit more structure.

Using the Specific Form of $Q'(\phi^i_A)$ From the definition:

$$Q'(\phi^i_A)(\theta) = v'(\theta)v(\theta)^T + v(\theta)v'(\theta)^T. \tag{8.6.139}$$

The key is that v and v' are orthogonal, and $Q'(\phi^i_A)$ is anti-symmetric with respect to Q. This implies that:

$$(Q'(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A) \otimes \mathbb{I}) \tag{8.6.140}$$

has limited overlap with $\rho^{ij,AB}_{ABEA'_1}$, leading to a smaller trace. Specifically, one can show that:

$$\mathrm{Tr}\left(\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A) \otimes \mathbb{I})\rho^{ij,AB}_{ABEA'_1}(Q'(\phi^i_A) \otimes \mathbb{I})\right) \leq \frac{1}{2}, \tag{8.6.141}$$

and similarly for the other term. This would give:

$$\mathrm{Tr}([\rho^{ij,AB}_{ABEA'_1}, Q'(\phi^i_A) \otimes \mathbb{I}]^2) \leq 2 \cdot \frac{1}{2} - 2 \cdot 0 = 1. \tag{8.6.142}$$

Thus,

$$\mathrm{Tr}([\rho^{ij,AB}_{ABEA'_1}, (Q(\phi^i_A) \otimes \mathbb{I})]^2) \leq 4 \quad \text{and} \quad \mathrm{Tr}\left([\rho^{ij,AB}_{ABEA'_1}, (Q'(\phi^i_A) \otimes \mathbb{I})]^2\right) \leq 1 \tag{8.6.143}$$

□

Thus, finally (8.6.110) and from 8.6.5.

$$\begin{aligned}
&\left|h(\lambda,\phi_A^i,\rho^{ij,AB}_{ABEA'_1}) - h(\lambda,\phi_A^i+\epsilon_0,\rho^{ij,AB}_{ABEA'_1})\right| \\
&\leq (1-\lambda)[-2.|\epsilon_0|.\sqrt{\mathrm{Tr}([\rho^{ij,AB}_{ABEA'_1},(Q(\phi_A^i)\otimes\mathbb{I})]^2)\cdot\mathrm{Tr}([\rho^{ij,AB}_{ABEA'_1},(Q'(\phi_A^i)\otimes\mathbb{I})]^2)} - |\epsilon_0^2|.\mathrm{Tr}\left([\rho^{ij,AB}_{ABEA'_1}\right. \\
&-2\mathrm{Tr}([\rho^{ij,AB}_{ABEA'_1},(Q(\phi_A^i)\otimes\mathbb{I})]^2)] \\
&\leq (1-\lambda)[-2.|\epsilon_0|.\sqrt{4.1} - |\epsilon_0^2|.1 - 2\cdot 4] \\
&\leq (1-\lambda)[-4|\epsilon_0| - |\epsilon_0^2|.1 - 8] \\
&\leq (1-\lambda)[-4|\epsilon_0| - 8] \\
&\leq (1-\lambda)[-4(|\epsilon_0| + 2)]
\end{aligned} \tag{8.6.144}$$

For very small ϵ_0, $|\epsilon_0^2| << |\epsilon_0|$ Thus, we have finally

$$\boxed{\left|h(\lambda,\phi_A^i,\rho^{ij,AB}_{ABEA'_1}) - h(\lambda,\phi_A^i+\epsilon_0,\rho^{ij,AB}_{ABEA'_1})\right| \leq 5(1-\lambda)|\epsilon_0|} \tag{8.6.145}$$

Now we are at a point where we can lower bound the maximal error in k_th segment centered around $\phi^i_{A_k}$ to the modified objective function in (8.6.24) using result from (8.6.145) and (8.6.71)

$$\boxed{\begin{aligned}
n^*(S_{ij}) \geq \inf \quad & \lambda\left\|(\rho^{ij,AB}_{ABEA'_0}) - (\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B})\right\|_F^2 + (1-\lambda)\left\|(\rho^{ij,AB}_{ABEA'_1}) - (\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B})\right\|_F^2 \\
& + \frac{\mu}{2}\|\rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1}\|_F^2 \ -2.5(1-\lambda)\epsilon_0 \\
\text{s.t.} \quad & \mathrm{Tr}\left(\rho^{ij,AB}_{ABEA'_x}\cdot chsh(\phi_A^i,\phi_B^j)\right) = S_{ij} \ -2\epsilon_0 \\
& \phi_A^i,\phi_B^j \in [0,\pi/2], \\
& \rho^{ij,AB}_{ABEA'_x} \succeq 0 \\
& \mathrm{Tr}(\rho^{ij,AB}_{ABEA'_x}) = 1
\end{aligned}} \tag{8.6.146}$$

By averaging multiple instances or scenarios in the optimization problem, the maximum deviation of the objective function's value from its expected or central tendency is reduced. Specifically, in this case, the maximum deviation has been halved from $\epsilon_0 = 5$ to $\epsilon'_0 = 2.5$. This reduction occurs because averaging tends to smooth out extreme values. Given that the optimal value is stated to lie within a symmetric interval of $\phi^i_{A_k} \pm \epsilon_0$, averaging multiple such intervals can lead to a smaller overall range of uncertainty for the optimal value.

8.7 Creating a convex hull from all the two qubit function $C^*_{\mathbb{C}^{4\times4}}(S')$ for all $S' \in (2, 2\sqrt{2}]$

The objective of the work is to establish a lower bound on the uncertainty of Alice's outcome given Eve's subsystem (8.1.1) as a function of CHSH value S as, $\lambda H(A_0|E)_{\rho_{A'BEA_X}} + (1-\lambda)H(A_1|E)_{\rho_{A'BEA_X}} \geq C^*(S)$. From (8.3.42) it had been shown that the function $C^*(S)$ can be lower bounded further as function of $C^*(S)$ over two qubit blocks, $C^*_{\mathbb{C}^{4\times4}}(S')$ as an integral over the range $S \in (2, 2\sqrt{2}]$ as, $C^*(S) \geq \int_{S'=2}^{2\sqrt{2}} \eta(dS')\cdot C^*_{\mathbb{C}^{4\times4}}(S')$ such that $\eta([2,2\sqrt{2}]) \leq 1$, $\eta \geq 0$, $\int_{S'=2}^{2\sqrt{2}} \eta(dS')S' = S$. Further it had been showed that each such two qubit function $C^*_{\mathbb{C}^{4\times4}}(S')$ can be lower bounded further through a strongly convex objective function $n^*(S_{ij})$ through a modified Pinsker's inequality[48] and the results from (8.4.7) (8.6.24). Thus our prime objective of establishing a lower bound on the conditional von neuman entropy of alice's outcome given eve's subsystem stand as solving the optimization problem involving the objective function as $n^*(S_{ij}) = \inf \quad \lambda \left\| (\rho^{ij,AB}_{ABEA'_0}) - (\Lambda_0[\rho_{ABEA'_0}]^{ij,A'B}) \right\|_F^2 + (1-\lambda) \left\| (\rho^{ij,AB}_{ABEA'_1}) - (\Lambda_1[\rho_{ABEA'_1}]^{ij,A'B}) \right\|_F^2 + \frac{\mu}{2} \| \rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1} \|_F^2$ with the constraints as $\mathrm{Tr}\left(\rho^{ij,AB}_{ABEA'_x} \cdot chsh(\phi^i_A, \phi^j_B)\right) = S_{ij}, \phi^i_A, \phi^j_B \in [0, \pi/2], \rho^{ij,AB}_{ABEA'_x} \succeq 0$ and $\mathrm{Tr}(\rho^{ij,AB}_{ABEA'_x}) = 1$.

Now given value of $C^*_{\mathbb{C}^{4\times4}}(S)$ for all $S \in (2, 2\sqrt{2}]$ one need a convex function say $\overline{C}(S)$[48] that would essentially give

$$C^*_{\mathbb{C}^{4\times4}}(S_{ij}) \geq \overline{C}(S_{ij}) \tag{8.7.1}$$

for ij^{th} block. Thus incorporating this into (8.3.42) as

$$\begin{aligned} C^*(S) \geq & \int_{S'=2}^{2\sqrt{2}} \eta(dS') \cdot \overline{C}(S') \\ s.t \quad & \eta([2, 2\sqrt{2}]) \leq 1 \\ & \eta \geq 0 \\ & \int_{S'=2}^{2\sqrt{2}} \eta(dS') S' = S \end{aligned} \tag{8.7.2}$$

gives

$$C^*(S) \geq \overline{C}(S) \tag{8.7.3}$$

The final task left is to prove the existence of one such function $\overline{C}(S)$ that lower bound $C^*(S)$.

Theorem 8.7.1. *$n^*(S)$ is a valid lower bound for $C^*(S)$.*

Proof.

1. Strong Convexity of the Objective Function The optimization problem defining $n^*(S)$ is given as

$$n^*(S_{ij}) = \inf \left\{ \lambda \left\| \rho^{ij,AB}_{ABEA'_0} - \Lambda_0 \left[\rho_{ABEA'_0}\right]^{ij,A'B} \right\|_F^2 + (1-\lambda) \left\| \rho^{ij,AB}_{ABEA'_1} - \Lambda_1 \left[\rho_{ABEA'_1}\right]^{ij,A'B} \right\|_F^2 + \frac{\mu}{2} \left\| \rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1} \right\|_F^2 \right\}$$
$$s.t. \mathrm{Tr}\left(\rho^{ij,AB}_{ABEA'_x} \cdot \mathrm{chsh}(\phi^i_A, \phi^j_B)\right) = S_{ij}. \tag{8.7.4}$$

   Now for the given objective function involving the Frobenius norm $\|\cdot\|_F^2$ is strongly convex in ρ. The term $\frac{\mu}{2} \left\| \rho^{ij,AB}_{ABEA'_0} - \rho^{ij,AB}_{ABEA'_1} \right\|_F^2$ introduces μ-strong convexity and Linear constraints preserve strong convexity on the feasible set. For fixed S_{ij}, the objective is strongly convex in ρ. The parameterized problem's value function $n^*(S_{ij})$ inherits convexity. Moreover Strong convexity implies quadratic dependence on perturbations in S_{ij}.
   Thus, $n^*(S)$ is strongly convex in S for $S \in (2, 2\sqrt{2}]$

2. Application of Jensen's Inequality For the measure η with,

$$\int_2^{2\sqrt{2}} \eta(dS') = 1, \quad \eta \geq 0, \quad \int_2^{2\sqrt{2}} \eta(dS') S' = S \tag{8.7.5}$$

   Now applying Jensen's inequality for strongly convex functions gives,

$$\int_2^{2\sqrt{2}} \eta(dS') n^*(S') \geq n^* \left(\int_2^{2\sqrt{2}} \eta(dS') S' \right) + \frac{\mu}{2} \mathrm{Var}_\eta(S') \tag{8.7.6}$$

   Now since $\mathrm{Var}_\eta(S') \geq 0$,

$$\int_2^{2\sqrt{2}} \eta(dS') n^*(S') \geq n^*(S). \tag{8.7.7}$$

Substituting $\overline{C}(S') = n^*(S')$ into the original inequality,

$$\boxed{C^*(S) \geq \int_2^{2\sqrt{2}} \eta(dS')n^*(S') \geq n^*(S)} \tag{8.7.8}$$

Thus, $\overline{C}(S) = n^*(S)$ is a valid convex lower bound.

□

Chapter 9

Computational Results

Quantum entanglement enables correlations that cannot be explained by classical local-hidden-variable theories. These correlations can be tested using Bell inequalities, particularly the CHSH inequality. This project investigates how the amount of entanglement in a bipartite quantum state affects the maximal violation of the CHSH inequality.

State Preparation

We consider a two-qubit pure state $\psi(\theta)$ defined as:

$$|\psi(\theta)\rangle = \cos(\theta)\,|00\rangle + \sin(\theta)\,|11\rangle\,, \tag{9.0.1}$$

where $\theta \in [0, \pi/2]$ controls the degree of entanglement. At $\theta = \pi/4$, the state is maximally entangled.

Measurement and Optimization

To test for Bell inequality violations, we define measurement settings A_0, A_1 for Alice and B_0, B_1 for Bob. The CHSH expression is given by:

$$\text{CHSH} = \langle A_0 B_0\rangle + \langle A_0 B_1\rangle + \langle A_1 B_0\rangle - \langle A_1 B_1\rangle. \tag{9.0.2}$$

Using the NPA hierarchy, we represent moment matrices of expectations, and use convex optimization (via `CVXPY`) to find optimal measurements that maximize CHSH for a fixed θ.

Entanglement Entropy

For a pure bipartite state $|\psi(\theta)\rangle$, the entanglement entropy is calculated from the reduced density matrix $\rho_A = \text{Tr}_B[|\psi(\theta)\rangle\,\langle\psi(\theta)|]$:

$$S(\rho_A) = -\text{Tr}(\rho_A \log_2 \rho_A), \tag{9.0.3}$$

which reaches 1 bit when $\theta = \pi/4$.

Results

We compute the CHSH value and entropy for several values of $\theta \in [0, \pi/2]$.

Discussion

- The CHSH value increases with θ and peaks at the maximally entangled state.
- The classical limit of CHSH is 2, while the quantum mechanical maximum (Tsirelson's bound) is $2\sqrt{2} \approx 2.828$.
- Entanglement entropy serves as a good indicator of nonlocality but they are not strictly monotonic beyond certain values of θ.

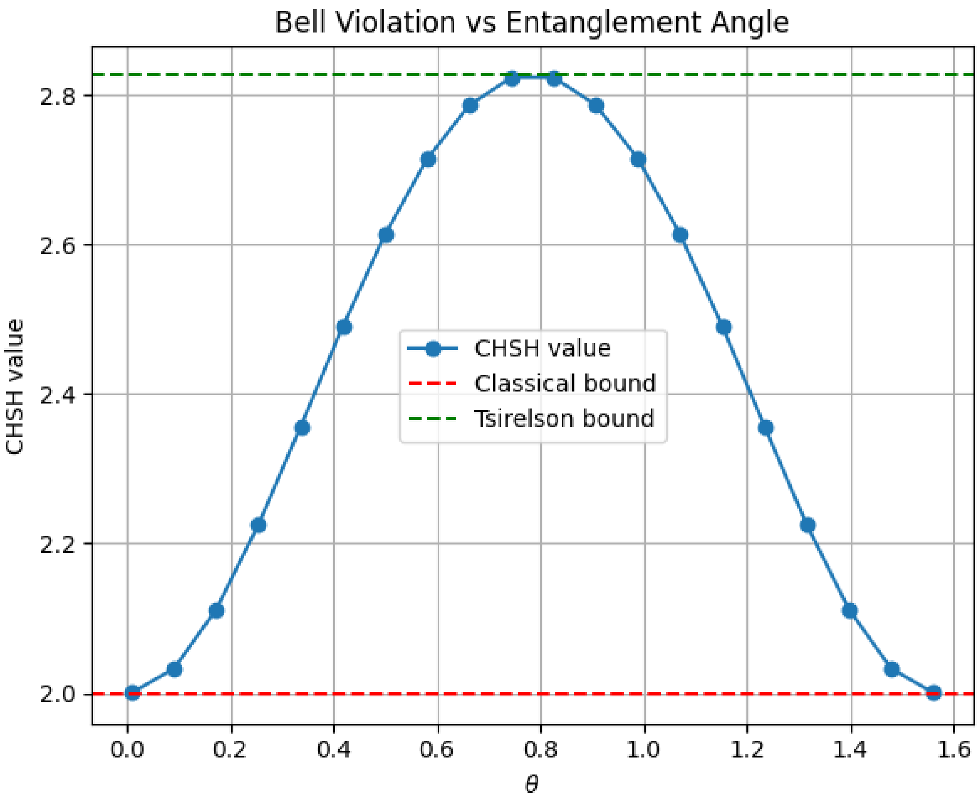

Figure 9.1: CHSH value vs. entanglement angle θ. The classical bound (2) and Tsirelson bound ($2\sqrt{2}$) are marked.

Conclusion

This numerical analysis confirms that stronger entanglement enables larger Bell inequality violations. However, entanglement alone is not sufficient: measurement optimization is essential to reveal nonlocal correlations. The study reinforces the significance of entangled states in quantum information theory.

Computational Platform

The Program were on Jupyter Notebook using Python 3.13.0 running on Ubuntu 22.04.5 LTS with $i3-7020U$ CPU at $2.30GHz$ and 8 GB RAM. The semidefinite programming were being solved using CVXPY.

Code Repository: The full notebook and plots are available at
$https://github.com/SawanBhattacharyya/Bell\_CHSH\_Violation\_SDP.git$

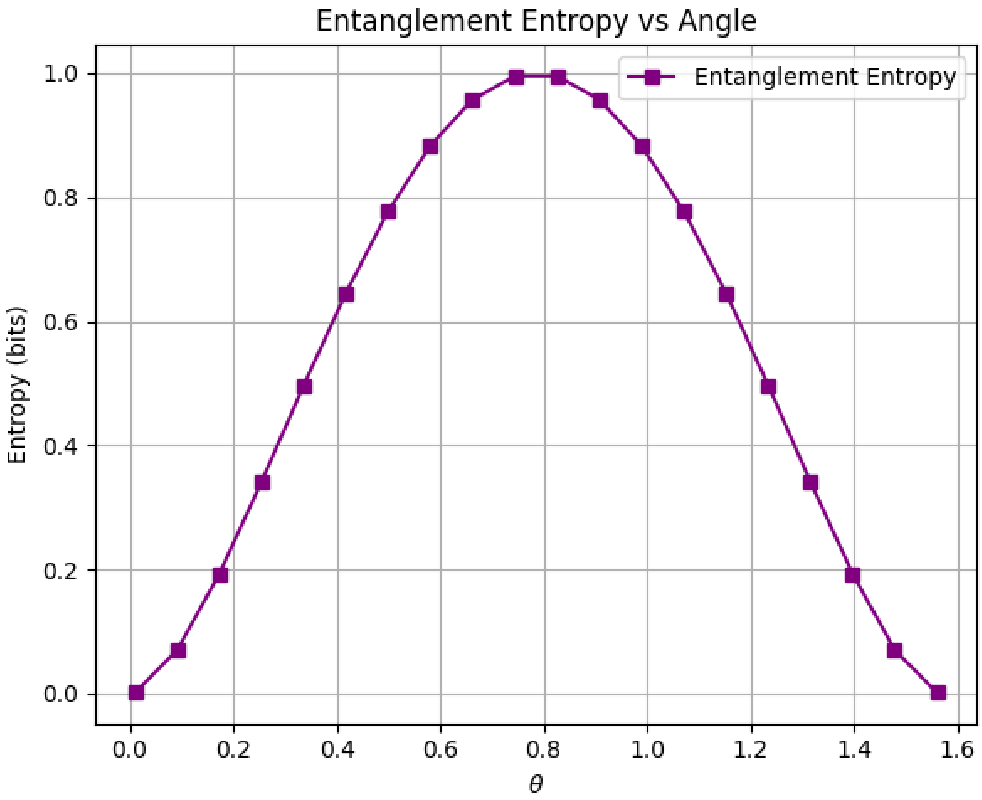

Figure 9.2: Entanglement entropy vs. angle θ. Entropy peaks at $\theta = \pi/4$ corresponding to maximal entanglement.

Chapter 10

Conclusion and Future Work

In many quantum key distribution (QKD) protocols, Alice and Bob select a particular pair of measurement settings for key generation. This choice improves the efficiency of the protocol by increasing the key rate. In the original device-independent QKD (DIQKD) protocol [2], Alice and Bob predominantly use the measurement setting $\{A_0, B_0\}$ for key generation, while employing the other settings less frequently for the purpose of channel testing. However, the security analysis in[42] indicates the existence of optimal eavesdropping strategies for which $H(A_0|E) \leq H(A_1|E)$. This suggests that if Eve knows Alice's measurement strategy, she can reduce her uncertainty about the key-generating setting A_0 by focusing her attack on it, even at the cost of increased uncertainty about A_1.

To address this vulnerability and increase Eve's uncertainty about the key-generation measurements, a variant known as DIQKD with a random key basis has been proposed. In this modified protocol, Alice employs both A_0 and A_1 settings for key generation with probabilities λ and $1 - \lambda$, respectively. Consequently, Eve's uncertainty is captured by the weighted conditional entropy expression $\lambda H(A_0|E) + (1 - \lambda)H(A_1|E)$, where E represents Eve's side information. The main challenge lies in lower-bounding this expression using only the observed violation of the CHSH inequality.

The authors in[48] formulate this challenge as an optimization problem and derive a lower bound $C^\star(S)$ satisfying $\lambda H(A_0|E) + (1 - \lambda)H(A_1|E) \geq C^\star(S)$. The present work demonstrates that the computational cost involved in the security analysis of this DIQKD variant can be significantly reduced. In the original approach, polytope optimization was used to determine Bob's optimal measurement angle. Our findings suggest that recasting the problem as a strongly convex optimization task allows one to apply the same ϵ-net approximation technique used for optimizing Alice's angles to Bob's as well. Since polytope optimization is more computationally intensive than ϵ-net approximation, this substitution leads to a more efficient security analysis.

A critical aspect of using the ϵ-net method is estimating the pessimistic error introduced at each iteration, as this error must be subtracted from the observed CHSH score S. To our knowledge, no prior work has provided an explicit expression for this pessimistic error. In this manuscript, we derive such an expression, showing that the error is the maximum product of two quantities: the angular resolution (i.e., precision) of the parameters ϕ_A^i and ϕ_B^j, and the first-order derivatives of the optimization solution functions. Our results explain how Alice and Bob can determine their optimal measurement angles ϕ_A^i and ϕ_B^j from the spectral properties of the projectors used in the CHSH test and key generation.

Furthermore, we clarify how the CHSH operator depends on the parties' measurement angles, which was not explicitly addressed in the original security proofs. We identify the angular configurations that yield the greatest deviation in the CHSH value. Overall, our findings indicate that it is possible to derive an analytic, convex lower bound on Eve's uncertainty about Alice's key-generating measurements. We believe that this work contributes to a clearer and more complete understanding of the security of DIQKD protocols with a random key basis.

There are several natural extensions of the work presented in this manuscript. One important direction is the generalization of the current analysis to higher-dimensional systems, such as qudits, which may lead to improved key rates and enhanced robustness to noise. Another promising avenue is to further investigate the tightness of the derived convex lower bound on Eve's uncertainty, and whether it can be analytically or numerically improved. Additionally, incorporating practical limitations such as device imperfections, finite-size effects, and imperfect random number generation would help bridge the gap between theoretical security proofs and experimental implementations.

The current method relies on ϵ-net approximations for optimizing measurement angles. Future work may consider adaptive or gradient-based optimization techniques to achieve better computational efficiency and precision. It is also of interest to extend the security analysis to account for adversaries with quantum memory or coherent attacks across multiple rounds, which are more relevant in realistic adversarial settings.

Moreover, the development of semidefinite programming (SDP) based tools that can automatically certify bounds on Eve's information from observed data could significantly streamline the security analysis process. Another potential direction is to explore how the protocol behaves in a networked setting involving multiple parties or intermediate nodes, particularly in the context of quantum repeater-based architectures. Finally, leveraging machine learning approaches—such as reinforcement learning or neural networks—for discovering optimal measurement strategies may lead to novel insights and further improvements in protocol performance.

Bibliography

[1] *Basic Real Analysis*. Springer (India) Pvt. Limited, 2005.

[2] Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, Stefano Pironio, and Valerio Scarani. Device-independent security of quantum cryptography against collective attacks. *Phys. Rev. Lett.*, 98:230501, Jun 2007.

[3] Antonio Acín, Nicolas Gisin, and Lluis Masanes. From bell's theorem to secure quantum key distribution. *Phys. Rev. Lett.*, 97:120405, Sep 2006.

[4] Huzihiro Araki and Elliott H. Lieb. Entropy inequalities. *Communications In Mathematical Physics*, 18(2):160–170, June 1970.

[5] Rotem Arnon-Friedman, Frédéric Dupuis, Omar Fawzi, Renato Renner, and Thomas Vidick. Practical device-independent quantum cryptography via entropy accumulation. *Nature Communications*, 9(1):459, Jan 2018.

[6] Manik Banik, Prasenjit Deb, and Samyadeb Bhattacharya. Wigner–yanase skew information and entanglement generation in quantum measurement. *Quantum Information Processing*, 16(4):97, Feb 2017.

[7] Jonathan Barrett, Lucien Hardy, and Adrian Kent. No signaling and quantum key distribution. *Phys. Rev. Lett.*, 95:010503, Jun 2005.

[8] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, 2001.

[9] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, Mar 1993.

[10] Charles H. Bennett, Gilles Brassard, and N. David Mermin. Quantum cryptography without bell's theorem. *Phys. Rev. Lett.*, 68:557–559, Feb 1992.

[11] Charles H. Bennett, David P. DiVincenzo, Peter W. Shor, John A. Smolin, Barbara M. Terhal, and William K. Wootters. Remote state preparation. *Phys. Rev. Lett.*, 87:077902, Jul 2001.

[12] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Phys. Rev. Lett.*, 69:2881–2884, Nov 1992.

[13] Mario Berta, Matthias Christandl, Roger Colbeck, Joseph M. Renes, and Renato Renner. The uncertainty principle in the presence of quantum memory. *Nature Physics*, 6(9):659–662, Sep 2010.

[14] J. Frédéric Bonnans and Alexander Shapiro. Perturbation analysis of optimization problems. In *Springer Series in Operations Research*, 2000.

[15] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[16] Samuel L. Braunstein and Stefano Pirandola. Side-channel-free quantum key distribution. *Phys. Rev. Lett.*, 108:130502, Mar 2012.

[17] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, Apr 2014.

[18] A. Böttcher and I.M. Spitkovsky. A gentle guide to the basics of two projections theory. *Linear Algebra and its Applications*, 432(6):1412–1459, 2010.

[19] Adán Cabello and Fabio Sciarrino. Loophole-free bell test based on local precertification of photon's presence. *Phys. Rev. X*, 2:021010, Jun 2012.

[20] B. G. Christensen, K. T. McCusker, J. B. Altepeter, B. Calkins, T. Gerrits, A. E. Lita, A. Miller, L. K. Shalm, Y. Zhang, S. W. Nam, N. Brunner, C. C. W. Lim, N. Gisin, and P. G. Kwiat. Detection-loophole-free test of quantum nonlocality, and applications. *Phys. Rev. Lett.*, 111:130406, Sep 2013.

[21] Patrick J. Coles, Mario Berta, Marco Tomamichel, and Stephanie Wehner. Entropic uncertainty relations and their applications. *Rev. Mod. Phys.*, 89:015002, Feb 2017.

[22] Marcos Curty and Tobias Moroder. Heralded-qubit amplifiers for practical device-independent quantum key distribution. *Phys. Rev. A*, 84:010304, Jul 2011.

[23] Prasenjit Deb and Manik Banik. Role of complementary correlations in the evolution of classical and quantum correlations under markovian decoherence. *Journal of Physics A: Mathematical and Theoretical*, 48(18):185303, apr 2015.

[24] Artur K. Ekert. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, 67:661–663, Aug 1991.

[25] Nicolas Gisin, Stefano Pironio, and Nicolas Sangouard. Proposal for implementing device-independent quantum key distribution based on a heralded qubit amplifier. *Phys. Rev. Lett.*, 105:070501, Aug 2010.

[26] Marissa Giustina, Marijn A. M. Versteegh, Sören Wengerowsky, Johannes Handsteiner, Armin Hochrainer, Kevin Phelan, Fabian Steinlechner, Johannes Kofler, Jan-Åke Larsson, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Morgan W. Mitchell, Jörn Beyer, Thomas Gerrits, Adriana E. Lita, Lynden K. Shalm, Sae Woo Nam, Thomas Scheidl, Rupert Ursin, Bernhard Wittmann, and Anton Zeilinger. Significant-loophole-free test of bell's theorem with entangled photons. *Phys. Rev. Lett.*, 115:250401, Dec 2015.

[27] B Hensen, H Bernien, A E Dréau, A Reiserer, N Kalb, M S Blok, J Ruitenberg, R F L Vermeulen, R N Schouten, C Abellán, W Amaya, V Pruneri, M W Mitchell, M Markham, D J Twitchen, D Elkouss, S Wehner, T H Taminiau, and R Hanson. Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(7575):682–686, October 2015.

[28] M. Ho, P. Sekatski, E.Y.-Z. Tan, R. Renner, J.-D. Bancal, and N. Sangouard. Noisy preprocessing facilitates a photonic realization of device-independent quantum key distribution. *Physical Review Letters*, 124(23), June 2020.

[29] Alexander S. Holevo. *Quantum Systems, Channels, Information*. De Gruyter, Berlin, Boston, 2019.

[30] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[31] Michał Horodecki, Jonathan Oppenheim, and Andreas Winter. Partial quantum information. *Nature*, 436(7051):673–676, August 2005.

[32] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009.

[33] Jan Kołodyński, Alejandro Máttar, Paul Skrzypczyk, Erik Woodhead, Daniel Cavalcanti, Konrad Banaszek, and Antonio Acín. Device-independent quantum key distribution with single-photon sources. *Quantum*, 4:260, April 2020.

[34] Charles Ci Wen Lim, Christopher Portmann, Marco Tomamichel, Renato Renner, and Nicolas Gisin. Device-independent quantum key distribution with local bell test. *Phys. Rev. X*, 3:031006, Jul 2013.

[35] Hoi-Kwong Lo, Marcos Curty, and Bing Qi. Measurement-device-independent quantum key distribution. *Phys. Rev. Lett.*, 108:130503, Mar 2012.

[36] Z.Z. Mammeri. *Cryptography: Algorithms, Protocols, and Standards for Computer Security*. Wiley, 2024.

[37] Lluís Masanes, Stefano Pironio, and Antonio Acín. Secure device-independent quantum key distribution with causally independent measurement devices. *Nature Communications*, 2(1):238, Mar 2011.

[38] Evan Meyer-Scott, Daniel McCloskey, Klaudia Gołos, Jeff Z. Salvail, Kent A. G. Fisher, Deny R. Hamel, Adán Cabello, Kevin J. Resch, and Thomas Jennewein. Certifying the presence of a photonic qubit by splitting it in two. *Phys. Rev. Lett.*, 116:070501, Feb 2016.

[39] Miguel Navascués, Stefano Pironio, and Antonio Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics*, 10(7):073013, July 2008.

[40] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[41] K. B. Petersen and M. S. Pedersen. The matrix cookbook, October 2008. Version 20081110.

[42] Stefano Pironio, Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, and Valerio Scarani. Device-independent quantum key distribution secure against collective attacks. *New Journal of Physics*, 11(4):045021, April 2009.

[43] Ioan Rasa. Concavity of some entropies, 2015.

[44] Ben W. Reichardt, Falk Unger, and Umesh Vazirani. A classical leash for a quantum system: Command of quantum systems via rigidity of chsh games, 2012.

[45] Renato Renner. Security of quantum key distribution, 2006.

[46] Wenjamin Rosenfeld, Daniel Burchardt, Robert Garthoff, Kai Redeker, Norbert Ortegel, Markus Rau, and Harald Weinfurter. Event-ready bell test using entangled atoms simultaneously closing detection and locality loopholes. *Phys. Rev. Lett.*, 119:010402, Jul 2017.

[47] E. Schrödinger. Die gegenwärtige situation in der quantenmechanik. *Naturwissenschaften*, 23(48):807–812, Nov 1935.

[48] René Schwonnek, Koon Tong Goh, Ignatius W. Primaatmaja, Ernest Y.-Z. Tan, Ramona Wolf, Valerio Scarani, and Charles C.-W. Lim. Device-independent quantum key distribution with random key basis. *Nature Communications*, 12(1), May 2021.

[49] Lynden K. Shalm, Evan Meyer-Scott, Bradley G. Christensen, Peter Bierhorst, Michael A. Wayne, Martin J. Stevens, Thomas Gerrits, Scott Glancy, Deny R. Hamel, Michael S. Allman, Kevin J. Coakley, Shellee D. Dyer, Carson Hodge, Adriana E. Lita, Varun B. Verma, Camilla Lambrocco, Edward Tortorici, Alan L. Migdall, Yanbao Zhang, Daniel R. Kumor, William H. Farr, Francesco Marsili, Matthew D. Shaw, Jeffrey A. Stern, Carlos Abellán, Waldimar Amaya, Valerio Pruneri, Thomas Jennewein, Morgan W. Mitchell, Paul G. Kwiat, Joshua C. Bienfang, Richard P. Mirin, Emanuel Knill, and Sae Woo Nam. Strong loophole-free test of local realism. *Phys. Rev. Lett.*, 115:250402, Dec 2015.

[50] Jamie William Jonathon Sikora. Analyzing quantum cryptographic protocols using optimization techniques. 2012.

[51] Gilbert Strang. *Calculus*. Wellesley-Cambridge Press, 1991.

[52] B.S. Thomson, A.M. Bruckner, and J.B. Bruckner. *Elementary Real Analysis*. Number v. 1. www.classicalrealanalysis.com., 2008.

[53] Marco Tomamichel. *Quantum Information Processing with Finite Resources*. Springer International Publishing, 2016.

[54] Umesh Vazirani and Thomas Vidick. Fully device-independent quantum key distribution. *Phys. Rev. Lett.*, 113:140501, Sep 2014.

[55] J. Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.

[56] Reinhard F. Werner. Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277–4281, Oct 1989.

[57] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2013.

[58] Feihu Xu, Yu-Zhe Zhang, Qiang Zhang, and Jian-Wei Pan. Device-independent quantum key distribution with random postselection. *Physical Review Letters*, 128(11), March 2022.