Modified security analysis of device-independent quantum key distribution with random key basis

Sawan Bhattacharyya,[1,2] Turbasu Chatterjee,[3] Pankaj Agrawal,[2] and Prasenjit Deb[2]

[1]*Department of Computer Science and Engineering, University of Calcutta, Bidhan Nagar, Kolkata - 700106, India*[*]
[2]*Center for Quantum Engineering, Research, and Education,*
TCG CREST, Bidhan Nagar, Kolkata - 700091, India.
[3]*Department of Computer Science, Virginia Tech, Blacksburg, VA 24060, United States*

(Dated: August 19, 2025)

Security analysis is a critical part in any cryptographic protocol, may it be classical or quantum. Without security analysis, one cannot ensure the secrecy of the distributed keys. To perform a conclusive security analysis, it is very often necessary to frame the problem as an optimization problem. However, solving such optimization problems is quite challenging. In this article, we focus on the security analysis of device-independent quantum key distribution (DIQKD) with random key basis protocol. We show that the optimization cost of the existing security analysis can be reduced without compromising the key rate. In particular, we reframe the entire security analysis of this protocol as a strongly convex optimization problem and demonstrate that unlike the original security proof, optimization of Bob's measurement angles for finding a lower bound on Eve's uncertainty about Alice's key generation basis can be done with lesser cost. We derive an explicit form of the pessimistic error that arises while optimizing the measurement angles of both the parties. We also clarify a few parts of the original security proof, making the analysis more rigorous and complete .

I. INTRODUCTION

Cryptographic protocols [Mam24] aid in establishing secure keys between two communicating parties, who later use the keys to access the information that they share between them. However, before implementing such protocols, it is necessary to perform their security analysis for avoiding any kind of potential unauthorized access of information during communication. In particular, through a conclusive security analysis we can find out the necessary and sufficient conditions for generating a positive secret key rate [Ren06] as well as its optimal lower bound. However, finding such bounds often involves optimization of cost functions that are derived by mapping the security analysis to an optimization problem.

Quantum cryptography offers non-trivial advantages over its classical counterpart due to the inherent randomness present in the state of the carrier of quantum information – qubit [NC10]. Moreover, the nonlocality [BCP+14] feature of quantum mechanics enables us to design quantum key distribution protocols that do not rely on any assumptions about the internal working of the quantum devices used in the protocols. Such key distribution protocols are termed as device-independent quantum key distribution (DIQKD) [ABG+07]. Allowing relaxation of the security assumptions made in usual quantum key distribution(QKD) [Ren06], DIQKD has made it possible to distribute secret keys using untrusted devices between two users embedded in an untrusted network. Due to this counterintuitive advantage over classical key distribution, this quantum cryptographic protocol is considered as the future of key distribution.

In this quantum key distribution protocol, the nonlocal state shared between the communicating parties limits the information that can be obtained by any third party. However, certification of nonlocality between the legitimate parties requires verification of violation of Clauser-Horne-Shimony-Holt (CHSH) inequality. For such a purpose, a sufficiently large number of entangled states has to be shared between the parties. This requirement is also necessary for generation of secure keys. However, for the possibility of secure key distribution, mere violation of the CHSH inequality with a large number of shared highly entangled state is not sufficient – the amount of violation must be above a certain threshold that depends on the protocol as well as on the security proofs[PAB+09]. Until the discovery of three variants of DIQKD [SGP+21],[XZZP22],[HST+20] unfortunately, the best known threshold so far was beyond the range achievable by the present day state-of-the-art experiments despite of the methods, e.g., heralding-type solutions [GPS10],[CM11],[KMS+20], local precertification [CS12],[MSMG+16], local Bell tests [LPT+13], and so on, proposed to improve the robustness of DIQKD. One such variant of this quantum key distribution protocol is DIQKD with random key basis [SGP+21]. As the name suggests, instead of a single measurement setting, more

[*] sawanbhattacharyyakv@gmail.com

than one measurement settings are used for generating secret keys. The key idea is to add additional difficulty for the eavesdropper by increasing the uncertainty about the measurement settings used to generate secret keys. In this article, we focus on the security analysis of this quantum key distribution protocol. We find that the optimization technique used in the existing security analysis to find secret key rate can be simplified, thereby reducing its computational cost. By framing the problem as a strongly convex optimization problem, we show that Bob's measurement angles can be optimized with lesser computational cost. We also find a closed form of the pessimistic error arising while optimizing the measurement angles of the parties Alice and Bob. Moreover, we address few other points in the original security analysis that demand clarification.

The rest of the article is arranged as follows. In Section (II), we briefly describe the quantum key distribution protocol that we analyze. Section (III) contains the results and discussions. Finally, in Section (IV), we conclude our work.

II. THE PROTOCOL

In the original DIQKD protocol [ABG+07], Eve knows the predetermined key generation basis. In particular, let Alice's and Bob's measurement settings are denoted by $X \in \{0,1\}$ and $Y \in \{0,1,2\}$, respectively. The corresponding outcomes are given by $A_X \in \{0,1\}$ and $B_Y \in \{0,1\}$, respectively. They derive the secret key from the measurement rounds in which the settings $A = 0$ and $B = 0$, respectively, are chosen. The CHSH score is determined using the remaining measurement combinations. Unlike this setting, in DIQKD with random key basis protocol, the secret is generated using both the measurement settings of Alice. Therefore, to obtain correlated outcomes, Bob needs an extra measurement setting, which is the proposed modification of the original protocol. Hence, in this variant of DIQKD protocol, Alice and Bob choose the settings $X = Y = 0$ and $X = Y = 1$ for generating secret key. The key intuition behind this protocol follows from the fact that the requirement of incompatible measurements for violation of the CHSH inequality necessitates the Alice's key generation measurements to be incompatible when the CHSH score $S > 2$. This condition imposes additional difficulty on Eve because now she has to guess the secret key from two randomly chosen incompatible measurements. For CHSH-based quantum key distribution protocols like DIQKD, different measurements will generate different amount of side information for Eve [SGP+21]. Therefore, Eve will fail to maximize her side information with at least one of the measurements, which in turn, provides an advantage over protocols based only on a single key-generating measurement. Due to her ignorance about which measurement setting will be chosen by Alice for key generation *a priori*, she will not be able to tailor her attack to the measurement in each round individually.

The protocol can be divided into five parts – measurements, sifting, parameter estimation, one-way error correction and verification, and finally privacy amplification. The measurement step is carried out with asymptotically large number of rounds N. In each measurement round, Alice and Bob choose their measurement settings $X \in \{0,1\}$ and $Y \in \{0,1,2,3\}$, respectively, following the probability distributions as defined: $P(X=0) = p$, $P(X=1) = 1-p$, $P(Y=0) = qp$, $P(Y=1) = q(1-p)$ and $P(Y=2) = P(Y=3) = (1-q)/2$, where $0 \leq p, q \leq 1$. Once Alice and Bob perform measurements with their respective devices and settings, they obtain outcomes $A_X \in \{0,1\}$ and $B_Y \in \{0,1\}$, respectively. They begin the second step of the protocol by announcing their measurement inputs over an authenticated public channel. Then using the inputs $Y \in \{0,1\}$ and $X = Y$ and the corresponding outcomes, they generate a pair of raw key of size $\sim q(p^2 + (1-p)^2)N$. Alice and Bob use the measurement settings $X \in \{0,1\}$ and $Y \in \{2,3\}$ to extract a pair of parameter estimation data of size $\sim (1-q)N$. All other measurement data are discarded. The third step of this key distribution protocol aims at computing and estimating the CHSH value. To do so, Alice and Bob reveal their measurement outcomes from the parameter estimation data set. The CHSH parameter can be expressed as,

$$S = \max\{0, C_{12} - C_{02} - C_{03} - C_{13}\} \tag{II.1}$$

where $C_{XY} = P(A_X = B_Y|X,Y) - P(A_X \neq B_Y|X,Y)$ quantifies the correlations between measurement outcomes A_X and B_Y. This step comprises the determining factor for continuation of the entire process. If $S > S_{tol}$, then Alice and Bob proceed to the next step, otherwise, they abort the protocol. Here S_{tol} is a predetermined threshold value of the violation of the CHSH inequality. A violation of the inequality beyond the threshold value confirms that the generated raw keys are secure. After getting this confirmation, Alice and Bob perform one-way error correction and verification, followed by privacy amplification.

Just like other DIQKD protocols, the security of this protocol can be quantified using the asymptotic secret key rate. However, due to the use of two measurement key bases for key generation, the key rate is defined in a way different from the existing DIQKD protocols,

$$K_\infty = p_s r_\infty \tag{II.2}$$

where $p_s := p^2 + (1-p)^2$ is the probability of Alice and Bob having matching key bases in the limit $q \to 1$. This key rate is the ratio of the extractable secret key length to the total number of measurement rounds N, which generally tends to infinity, i.e., $N \to \infty$. The second product term in the above equation is the secret fraction [SGP+21] and it can be written in terms of entropic functions as,

$$\begin{aligned} r_\infty \;:=\; & \lambda H(A_0|E) + (1-\lambda)H(A_1|E) \\ & -\lambda h(Q_{A_0B_0}) - (1-\lambda)h(Q_{A_1B_1}) \end{aligned} \tag{II.3}$$

where $\lambda := p^2/p_s$ and $h(x) := -x\log(x) - (1-x)\log(1-x)$ is the binary entropy function. The quantity in the argument of the binary function quantifies the quantum bit-error rate (QBER) for X, Y and is given by $Q_{A_XB_Y} := P(A_X \neq B_Y|X,Y)$. The quantum side-information that Eve can gather just before the error-correction step is given by E. From the secret fraction we can evaluate how much information Eve can get about Alice's and Bob's measurement outcomes. The terms, conditional von Neumann entropy, comprising the first line of Eq.(3) measure the amount of uncertainty Eve has about Alice's measurement outcomes given her side-information E. The entropic terms in the second line of the equation quantify the information that Eve can steal during the error-correction step by decoding her side-information. If we denote $U = \lambda H(A_0|E) + (1-\lambda)H(A_1|E)$, then Alice and Bob can generate a positive secret key rate, provided they can put a reliable lower bound on U using solely the CHSH violation that they observe. This task is quiet challenging, however, it has been shown that by employing a family of device-independent entropic uncertainty relations[BCC+10][CBTW17], one can have[SGP+21],

$$\lambda H(A_0|E) + (1-\lambda)H(A_1|E) \geq C^\star(S) \tag{II.4}$$

where $C^\star$ is a function of the CHSH score S. Using U as a performance metric, it has been found that the uncertainty of Eve in this protocol is always higher than the original protocol for all $S \in (2, 2\sqrt{2}]$ [SGP+21]. In fact, given a fixed S, for $\lambda = \frac{1}{2}$, the bound on U converges to its nearly optimal value, which is the fundamental upper limit of Eve's uncertainty [SGP+21]. In this work, our goal is to reduce the optimization cost of establishing a lower bound on Eve's uncertainty without compromising the secret key rate.

III. RESULTS AND DISCUSSIONS

Computation of the function $C^\star(S)$ involves a refined version of Pinsker's inequality [Wil13], semi-definite optimization (SDP)[Sik12], and an ϵ-net approximation [SGP+21]. In the original security analysis of the protocol, to find a lower bound on the function $C^\star(S)$, three optimization steps have been followed – optimization of the joint state ρ between Alice and Bob for fixed measurement angles on both sides, optimization of Bob's angle, and optimization of Alice's angles. All the steps in the original analysis assume the worst-case scenario so that the final value is a strict lower bound on the function. Here, we provide an improvised security analysis of the protocol.

It has been found that Alice's and Bob's measurement angles ϕ_A^i and ϕ_B^j appear as constraints in the optimization problem – the SDP used to find out a lower bound on the function $C^\star(S)$. The work of [SGP+21] showed that the Alice's angle ϕ_A^i can be optimized using ϵ - net approach. Here, we show that one can use a similar approach to optimize the Bob angle ϕ_B^j, thus eliminating the polytope optimization used in the original protocol. The polytope optimization step introduced in the original security analysis initialised the polytope $\mathcal{P}$ as a rectangle. $\mathcal{P}$ contain the unit circle [SGP+21] of all the point that depends on Bob's angle through *sin* and *cos* function. In this work, we extend the idea of the $\epsilon-$net approach to optimize Bob's angle, thereby making the overall optimization process less costly.

The standard cryptographic security analysis assume that the measurement angles of Alice and Bob are independent variables. To begin with, let us divide the possible range of each angle, $[0, \frac{\pi}{2}]$, into a set of discrete points, $\phi_{A_k}^i$ for Alice and $\phi_{B_l}^j$ for Bob, and choose the midpoint of each segment as k and l. Recognizing that the optimal angle for either Alice or Bob could lie anywhere within the continuous range $[0, \frac{\pi}{2}]$, a pessimistic error approach has been adopted. This involves iteratively refining the angle for one party (either Alice or Bob) while keeping the other's angle fixed. In each iteration, the SDP is being evaluated at the midpoints of the current segments for both angles. The segment that yields the smallest SDP value is then chosen for further subdivision. This process continues until a point corresponding to a local minimum is reached for the chosen angle. During this minimization for one angle, a pessimistic error estimate is subtracted from the SDP result to account for the continuous nature of the angle. Once the optimal angle for one party is determined through this iterative refinement, this optimal value is then fixed. The same iterative optimization process is then applied to find the optimal angle for the other party, now considering the previously determined optimal angle as a constant. This sequential optimization allows them to find a potentially optimal solution for the SDP by iteratively minimizing with respect to one angle at a time.

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 $\{\phi^i_{A_k}\}_{k=1}^{S_A} \subset I$ and $\{\phi^j_{B_l}\}_{l=1}^{S_B} \subset I$ such that for any $(\phi^i_A, \phi^j_B) \in I \times I$, there exist $\phi^i_{A_k}$ and $\phi^j_{B_l}$ satisfying:

$$|\phi^i_A - \phi^i_{A_k}| \leq \epsilon_0 \quad |\phi^j_B - \phi^j_{B_l}| \leq \epsilon_0 \tag{III.1}$$

where S_A and S_B are the number of segments in the interval I. Each segment is centralized around the angles $\phi^i_{A_k}$ and $\phi^j_{B_l}$ for k^{th} and l^{th} segments respectively. The pessimistic error term was introduced in the original protocol [SGP+21] as a function of change Δ in ϵ_0 and $\phi^i_A\,(\phi^j_B)$. In this work, a closed form of the same is derived.

Let us consider that $f(\phi)$ is a solution of the optimization problem for given angle. Each discrete angle $\phi^i_{A_k}$ and $\phi^j_{B_l}$ is being separated by a distance of $2\epsilon_0$, thus representing two segments of the same width as,

$$\begin{aligned} I_A &= \left[\phi^i_{A_k} - \epsilon_0, \phi^i_{A_k} + \epsilon_0\right] \\ I_B &= \left[\phi^j_{B_l} - \epsilon_0, \phi^j_{B_l} + \epsilon_0\right] \end{aligned} \tag{III.2}$$

The pessimistic error terms $\Delta(\epsilon_0, \phi^i_A)$ and $\Delta(\epsilon_0, \phi^j_B)$ provide an upper bound on the absolute difference between the values of the function at any point within these segments and at the representative discrete point. Hence, from the definition of $f(\phi)$, we have,

$$\begin{aligned} &|f(\phi^i_A) - f(\phi^i_{A_k})| \leq \Delta\left(\epsilon_0, \phi^i_A\right) \\ &|f(\phi^j_B) - f(\phi^j_{B_l})| \leq \Delta\left(\epsilon_0, \phi^j_B\right), \qquad \forall \phi^i_A \in I_A \text{ and } \forall \phi^j_B \in I_B \end{aligned} \tag{III.3}$$

A. Modification of the optimization problem accommodating Lipschitz continuity

Proving the below mentioned theorem, we first modify the optimization problem.

Theorem : If $f(\phi)$ is a solution of an well-defined optimization problem, then it is Lipschitz continuous, i.e., there exists a constant $L_f > 0$ such that,

$$\|f(\phi_1) - f(\phi_2)\| \leq L_f \|\phi_1 - \phi_2\|. \tag{III.4}$$

The proof is provided in the Appendix section F 1. To make the objective function continuously differentiable with respect to both the density operator and the measurement angles, the squared Frobenius norm is introduced. The squared Frobenius norm is smooth and continuously differentiable at every point[PP08]. To make the objective function strongly convex, it is necessary to add a strongly convex regularisation term. Therefore, the optimization problem in the original security analysis of the protocol now becomes,

$$\boxed{\begin{aligned} n^*(S_{ij}) = \inf \quad & \lambda \left\|\rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]\right\|_F^2 + (1-\lambda)\left\|\rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]\right\|_F^2 + \frac{\mu}{2}\|\rho^{ij}_{AB}\|_F^2 \\ \text{s.t.} \quad & \mathrm{Tr}\left(\rho^{ij}_{AB}\, CHSH(\phi^i_A, \phi^j_B)\right) = S_{ij} \\ & \phi^i_A, \phi^j_B \in [0, \pi/2], \\ & \rho^{ij}_{AB} \succeq 0 \\ & \mathrm{Tr}(\rho^{ij}_{AB}) = 1 \end{aligned}} \tag{III.5}$$

The modified version of the optimization problem can be reformulated into an SDP using Schur complements

[BTN01]. Let us first express each Frobenius norm terms into the corresponding inner product form as,

$$\begin{aligned} n^*(S_{ij}) = -\text{maximize} \quad & \lambda \text{Tr}({\rho_0}^{\dagger}\rho_0) + (1-\lambda)\text{Tr}({\rho_1}^{\dagger}\rho_1) + \frac{\mu}{2}\text{Tr}({\rho_2}^{\dagger}\rho_2) \\ \text{s.t.} \quad & \text{Tr}\left(\rho^{ij}_{AB}\, CHSH(\phi^i_A, \phi^j_B)\right) = S_{ij} \\ & \phi^i_A, \phi^j_B \in [0, \pi/2], \\ & \rho^{ij}_{AB} \succeq 0 \\ & \text{Tr}(\rho^{ij}_{AB}) = 1 \end{aligned} \tag{III.6}$$

where $\rho_0 = \rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]$, $\rho_1 = \rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]$ and $\rho_2 = \rho^{ij}_{AB}$
Each inner product (ρ_k, ρ_k) for $k \in \{0, 1, 2\}$ in the above equation is quadratic in ρ_k. The standard SDP formulation requires that the objective function to be linear in its decision variables and the constraints to be in the form of linear matrix inequalities. Let $t_k \geq (\rho_k, \rho_k)$, then using Schur's complement we get,

$$\begin{pmatrix} t_k & \text{vec}(\rho_k)^{\dagger} \\ \text{vec}(\rho_k) & 1 \end{pmatrix} \succeq 0 \tag{III.7}$$

Since CHSH correlation operator is fixed for a given value of the angles ϕ^i_A and ϕ^j_B, the standard SDP formulation of the optimization problem would be,

$$\boxed{\begin{aligned} n^*(S_{ij}) = -\text{maximize} \quad & \lambda t_0 + (1-\lambda) t_1 + \frac{\mu}{2} t_2 \\ \text{s.t.} & \begin{pmatrix} t_0 & \text{vec}(\rho_0)^{\dagger} \\ \text{vec}(\rho_0) & 1 \end{pmatrix} \succeq 0, \begin{pmatrix} t_1 & \text{vec}(\rho_1)^{\dagger} \\ \text{vec}(\rho_1) & 1 \end{pmatrix} \succeq 0, \begin{pmatrix} t_2 & \text{vec}(\rho_2)^{\dagger} \\ \text{vec}(\rho_2) & 1 \end{pmatrix} \succeq 0 \\ & \text{Tr}\left(\rho^{ij}_{AB}\, CHSH(\phi^i_A, \phi^j_B)\right) = S_{ij}, \\ & \phi^i_A, \phi^j_B \in [0, \pi/2], \\ & \rho^{ij}_{AB} \succeq 0 \\ & \text{Tr}(\rho^{ij}_{AB}) = 1 \end{aligned}} \tag{III.8}$$

B. Bounding the pessimistic error terms $\Delta\left(\epsilon_0, \phi^i_A\right)$ and $\Delta\left(\epsilon_0, \phi^j_B\right)$

The solution of the modified optimization problem, $n^*(S_{ij})$, is thus established to be Lipschitz continuous. Now one can proceed towards formulating a closed form for the pessimistic error terms, $\Delta\left(\epsilon_0, \phi^i_A\right)$ and $\Delta\left(\epsilon_0, \phi^j_B\right)$. The solution of the SDP gives an optimal value for the k^{th} and l^{th} segment centered around $\phi^i_{A_k}$ and $\phi^j_{B_l}$ for Alice and Bob, respectively. The solution of the SDP is based on fixed ϕ^i_A and ϕ^j_B. Without loss of generality, we consider the function f to be the solution of the optimization problem. Let denote two functions analogous to the solution of the modified optimization problem for Alice's and Bob's angles as,

$$f_A(\phi^i_A) = R_A, f_B(\phi^j_B) = R_B \tag{III.9}$$

where R_A and R_B give the optimal value of the SDP for the given parameterized angles.
Now, since the functions defined above are Lipschitz continuous, one can have the following relation for each segment, where $\phi^i_{A_k}$ and $\phi^j_{B_l}$ are the centres of the segments k and l, respectively.

$$\begin{aligned} & |f_A(\phi^i_A) - f_A(\phi^i_{A_k})| \leq L_A |\phi^i_A - \phi^i_{A_k}| \\ & |f_B(\phi^j_B) - f_B(\phi^j_{B_l})| \leq L_B |\phi^j_B - \phi^j_{B_l}|, \qquad \forall \phi^i_A \in I_A \text{ and } \forall \phi^j_B \in I_B \end{aligned} \tag{III.10}$$

The above inequalities put upper bounds on the absolute deviation in the solution of the optimization problem. The supremum of the above functions are given as,

$$M^i_A = sup_{\phi^i_A \in I_A} f_A(\phi^i_A) \text{ and } M^j_B = sup_{\phi^j_B \in I_B} f_B(\phi^j_B) \tag{III.11}$$

Similarly, the infimum of the solutions are being given as,

$$m^i_A = inf_{\phi^i_A \in I_A} f_A(\phi^i_A) \text{ and } m^j_B = inf_{\phi^j_B \in I_B} f_B(\phi^j_B) \tag{III.12}$$

The Lipschitz continuity of the solution functions implies that they are continuous in the given interval. Therefore, the supremums are the corresponding maximums and the infimum are the corresponding minimums. The supremum and infimum guarantee that no number smaller than M^i_A and M^j_B can serve as an upper-bound, and equivalently, no number larger than m^i_A and m^j_B can act as lower bound for $f_A(\phi^i_A)$ and $f_B(\phi^j_B)$, respectively. The quantities that we are interested in bounding are,

$$sup_{\phi^i_A \in I_A} \left| f_A(\phi^i_A) - f_A(\phi^i_{A_k}) \right| \quad \text{and} \quad sup_{\phi^j_B \in I_B} \left| f_B(\phi^j_B) - f_B(\phi^j_{B_l}) \right| \tag{III.13}$$

where $\phi^i_{A_k}$ and $\phi^j_{B_l}$ represent the center of the segments k and l, respectively. We find that the deviation gets maximised at the boundary of the respective segments (*Theorem 3* in Appendix F 3). Hence, we have,

$$\begin{aligned} &sup_{\phi^i_A \in I_A} |f_A(\phi^i_A) - f_A(\phi^i_{A_k})| = L_A \epsilon_0 \\ &sup_{\phi^j_B \in I_B} |f_B(\phi^j_B) - f_B(\phi^j_{B_l})| = L_B \epsilon_0, \qquad \forall \phi^i_A \in I_A \text{ and } \forall \phi^j_B \in I_B \end{aligned} \tag{III.14}$$

And from the definition of the pessimistic error terms $\Delta\left(\epsilon_0, \phi^i_A\right)$ and $\Delta\left(\epsilon_0, \phi^j_B\right)$, we get,

$$\Delta\left(\epsilon_0, \phi^i_A\right) = L_A \epsilon_0 \quad \text{and} \quad \Delta\left(\epsilon_0, \phi^j_B\right) = L_B \epsilon_0 \tag{III.15}$$

Since the functions $f_A(\phi^i_A)$ and $f_B(\phi^j_B)$ are differentiable in the intervals I_A and I_B, respectively, we determine the values of L_A and L_B using the following Taylor series as follows,

$$\begin{aligned} f_A(\phi^i_A) &= \sum_{n=0}^{\infty} \frac{f^n_A(\phi^{i_*}_A)}{n!} (\phi^i_A - \phi^{i_*}_A)^n \\ f_B(\phi^j_B) &= \sum_{m=0}^{\infty} \frac{f^m_B(\phi^{j_*}_B)}{m!} (\phi^j_B - \phi^{j_*}_B)^m, \qquad \forall \phi^{i_*}_A \in I_A \text{ and } \forall \phi^{j_*}_B \in I_B \end{aligned} \tag{III.16}$$

Now in this case, linear approximations of $f_A(\phi^i_A)$ near $\phi^i_{A_k}$ and $f_B(\phi^j_B)$ near $\phi^j_{B_l}$ are needed for small ϵ_0. One can neglect the higher-order derivatives terms because $f^1_A(\phi^i_{A_k})(\phi^i_A - \phi^i_{A_k})$ and $f^1_B(\phi^j_{B_l})(\phi^j_B - \phi^j_{B_l})$ dominate as $|\phi^i_A - \phi^i_{A_k}|$ and $|\phi^j_B - \phi^j_{B_l}|$ tends to 0 for k^{th} and l^{th} segments of Alice and Bob, respectively. The necessity of incorporating the higher-order terms in the error analysis arises when ϵ_0 is nonnegligible, the underlying functions exhibit strong non-linearity, or when a more precise and tighter bound on the error is desired. In such scenarios, only linear approximations are insufficient to accurately capture the function's behaviour within the interval of width $2\epsilon_0$. The contributions from higher derivatives that are neglected in the first-order analysis become non-negligible and must be accounted for to achieve the required accuracy in the error estimation. Thus, using the first order linear approximation, one can get an approximation of $f_A(\phi^i_A)$ and $f_A(\phi^j_B)$ near $\phi^i_{A_k}$ and $\phi^j_{B_l}$ as,

$$\begin{aligned} &f_A(\phi^i_A) = f_A(\phi^i_{A_k}) + f^1_A(\phi^i_{A_k})(\phi^i_A - \phi^i_{A_k}) + \mathcal{O}((\phi^i_A - \phi^i_{A_k})^2) \\ &|f_A(\phi^i_A) - f_A(\phi^i_{A_k})| \leq |f^1_A(\phi^i_{A_k})||(\phi^i_A - \phi^i_{A_k})| + \mathcal{O}((\phi^i_A - \phi^i_{A_k})^2) \end{aligned} \tag{III.17}$$

Similarly,

$$|f_B(\phi^j_B) - f_B(\phi^j_{B_l})| \leq |f^1_B(\phi^j_{B_l})||(\phi^j_B - \phi^j_{B_l})| + \mathcal{O}((\phi^j_B - \phi^j_{B_l})^2) \tag{III.18}$$

The higher-order terms, represented by $\mathcal{O}((\phi^i_A - \phi^i_{A_k})^2)$ and $\mathcal{O}((\phi^j_B - \phi^j_{B_l})^2)$, 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 essentially underestimate the actual deviations $|f_A(\phi^i_A) - f_A(\phi^i_{A_k})|$ 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 we get,

$$\begin{aligned} sup_{\phi^i_A \in I_A} |f_A(\phi^i_A) - f_A(\phi^i_{A_k})| \leq |f^1_A(\phi^i_{A_k})||(\phi^i_A - \phi^i_{A_k})| \\ sup_{\phi^j_B \in I_B} |f_B(\phi^j_B) - f_B(\phi^j_{B_l})| \leq |f^1_B(\phi^j_{B_l})||(\phi^j_B - \phi^j_{B_l})| \end{aligned} \tag{III.19}$$

The above relations hold good $\forall\ \phi^i_A, \phi^j_B \in I$ if,

$$\Delta\left(\epsilon_0, \phi^i_A\right) = max_{\phi^i_A \in I} |f^1_A(\phi^i_{A_k})| \epsilon_0 \quad \text{and} \quad \Delta\left(\epsilon_0, \phi^j_B\right) = max_{\phi^i_A \in I} |f^1_B(\phi^j_{B_l})| \epsilon_0 \tag{III.20}$$

where $L_A = max_{\phi^i_A \in I} |f^1_A(\phi^i_A)|$ and $L_B = max_{\phi^i_A \in I} |f^1_B(\phi^j_B)|$. Thus, we evaluate the pessimistic errors involved in the $\epsilon-$net approximation used for optimizing Alice's and Bob's angles.

C. Alice's and Bob's measurement angles

The work of [SGP+21] shows that without loss of generality one can can reduce the problem to two-qubit space $\mathbb{C}_{4\times 4}$ of Alice's and Bob's subsystem. From the result of [BS10] we can decompose a projector of higher dimension into projectors of dimension 2 × 2 acting on either Alice's or Bob's state. Now the work[SGP+21] proposed that the angles for Alice and Bob for choosing the projector and hence the observables are obtained from the *spectrum* of $P^{0|0}_{\mathcal{X}} + P^{0|1}_{\mathcal{X}}$ and $P^{0|2}_{\mathcal{Y}} + P^{0|3}_{\mathcal{Y}}$, respectively (C.2). Using this result, we find the explicit expression for Alice's and Bob's measurement angles as,

$$\phi^i_A = 2 \arccos\left(\frac{\lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}}}{2}\right) \quad \text{and} \quad \phi^j_B = 2 \arccos\left(\frac{\lambda^1_{\mathcal{Y}} - \lambda^2_{\mathcal{Y}}}{2}\right), \tag{III.21}$$

where ϕ^i_A and ϕ^j_B are the angles corresponding to ij^{th} two qubit space marked by i^{th} and j^{th} indices of Alice and Bob, respectively. $\lambda^1_{\mathcal{X}}, \lambda^2_{\mathcal{X}}$ and $\lambda^1_{\mathcal{Y}}, \lambda^2_{\mathcal{Y}}$ are the eigen value of $P^{0|0}_{\mathcal{X}} + P^{0|1}_{\mathcal{X}}$ and $P^{0|2}_{\mathcal{Y}} + P^{0|3}_{\mathcal{Y}}$, respectively. We find that such a mapping is unique (See *Lemma 1* in appendix C).

D. Dependence of CHSH operator on Alice's and Bob's angles

The optimisation problem presented in equation (F.19) exhibits a dependence on Alice's angle, denoted as ϕ^i_A, through the channel Λ_1 as defined by equations (B.2), (C.1), (C.2), and (C.3). Additionally, it depends on the CHSH operator, as specified in equations (C.26) and (F.19). 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 optimisation 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 variations in the parameters.

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. Let us define the maximum of the spectral norms of the differences in the CHSH operator as,

$$\begin{aligned}\delta_p = \max\Big(& ||CHSH(\phi^i_A, \phi^j_B) - CHSH(\phi^i_A + \epsilon, \phi^j_B)||_\infty, \\ & ||CHSH(\phi^i_A, \phi^j_B) - CHSH(\phi^i_A, \phi^j_B + \epsilon)||_\infty, \\ & ||CHSH(\phi^i_A, \phi^j_B) - CHSH(\phi^i_A + \epsilon, \phi^j_B + \epsilon)||_\infty\Big)\end{aligned} \tag{III.22}$$

Evaluating each term in the above equation F 4, we finally get,

$$\begin{aligned}\delta_p \leq \max\Bigg(& \left\|\epsilon\left(\cos\left(\phi^i_A\right)\sigma_x - \sin\left(\phi^i_A\right)\sigma_z\right)\right\|_\infty \cdot \left\|2\left(Q\left(\phi^j_B\right) - \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\right)\right\|_\infty, \\ & \left\|2\left(Q\left(\phi^i_A\right) - \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right)\right\|_\infty \cdot \left\|\epsilon\left(\cos\left(\phi^j_B\right)\sigma_x - \sin\left(\phi^j_B\right)\sigma_z\right)\right\|_\infty, \\ & \left\|\epsilon\left(\cos\left(\phi^i_A\right)\sigma_x - \sin\left(\phi^i_A\right)\sigma_z\right)\right\|_\infty \cdot \left\|\epsilon\left(\cos\left(\phi^j_B\right)\sigma_x - \sin\left(\phi^j_B\right)\sigma_z\right)\right\|_\infty\Bigg) \\ \leq \max & \left(\epsilon \cdot 2\sin\left(\frac{\phi^j_B}{2}\right), \quad 2\cos\left(\frac{\phi^i_A}{2}\right)\cdot\epsilon, \quad \epsilon^2\right) \\ \leq \quad & 2\epsilon_0\end{aligned} \tag{III.23}$$

The equality holds for the spectral norm of Kronecker products, therefore, this upper bound is exact. Now in the interval I, $\sin\left(\frac{\phi^j_B}{2}\right)$ and $\cos\left(\frac{\phi^i_A}{2}\right)$ functions are monotonically increasing and decreasing respectively. The first term $2\sin\left(\frac{\phi^j_B}{2}\right)$ is increasing from 0 to 1.414 and the second term $2\cos\left(\frac{\phi^i_A}{2}\right)$ is decreasing from 2ϵ to 1.414. Thus, the maximum deviation of the CHSH operator is achieved when ϕ^i_A is fixed at 0 but ϕ^j_B is increased by ϵ_0.

E. Dependence of objective function on Alice's angle

The objective function in (F.19) is a function of the density operator describing the joint state between Alice and Bob and Alice's angle ϕ^i_A. To analyse the dependency of the solution of the optimisation problem on ϕ^i_A, we reinterpret the Frobenius norm terms in (F.19) using dual norms. Since the Frobenius norm is self-dual, this framework retains the norm structure while emphasising its role as a maximiser of inner products. Now, as the objective function depends only upon Alice's angle, it is sufficient to consider only the perturbation of Alice's angle ϕ^i_A. Moreover, in the objective function, only one term is a function of the Alice's angle; therefore, we consider only that term and analyse the perturbation. The absolute change in the objective function can be written as,

$$\left|h(\lambda, \phi^i_A, \rho^{ij}_{AB}) - h(\lambda, \phi^i_A + \epsilon, \rho^{ij}_{AB})\right| \tag{III.24}$$

The first term can be expanded in the following way,

$$\begin{aligned}h(\lambda, \phi^i_A, \rho^{ij}_{AB}) = (1-\lambda)\mathrm{Tr}\Bigg(\Bigg[&\rho^{ij}_{AB} - \left(\frac{1}{2}\{\rho^{ij}_{AB}, Q(\phi^i_A)\otimes\mathbb{I}\} - \frac{1}{2}[\rho^{ij}_{AB}, Q(\phi^i_A)\otimes\mathbb{I}]\right) Q(\phi^i_A)\otimes\mathbb{I} \\ & - \left(\frac{1}{2}\{\rho^{ij}_{AB}, (\mathbb{I} - Q(\phi^i_A)\otimes\mathbb{I})\} - \frac{1}{2}[\rho^{ij}_{AB}, (\mathbb{I} - Q(\phi^i_A)\otimes\mathbb{I})]\right)\{\mathbb{I} - Q(\phi^i_A)\otimes\mathbb{I}\}\Bigg]^2\Bigg)\end{aligned} \tag{III.25}$$

In a similar way, we can expand the second term and finally get,

$$\boxed{\begin{array}{l}\quad\left|h(\lambda, \phi^i_A, \rho^{ij}_{AB}) - h(\lambda, \phi^i_A + \epsilon_0, \rho^{ij}_{AB})\right| \\ \leq (1-\lambda)[4\epsilon_0 + \mathcal{O}(\epsilon_0^2)]\end{array}} \tag{III.26}$$

The tighter bound of $4\epsilon_0$ in the previous work is based upon assumption that ρ^{ij}_{AB} and $Q(\phi^i_A)\otimes\mathbb{I}$ commute. Having evaluated the change in the objective function with respect to Alice's angle, we can now lower bound the maximal error in k^{th} segment centered around $\phi^i_{A_k}$ and get the modified optimization problem as,

$$\boxed{\begin{aligned} n^*(S_{ij}) \geq \inf \quad & \lambda \left\|\rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]\right\|_F^2 + (1-\lambda)\left\|\rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]\right\|_F^2 \\ & + \frac{\mu}{2}\|\rho^{ij}_{AB}\|_F^2 \ - 2(1-\lambda)\epsilon_0 \\ \text{s.t.} \quad & \mathrm{Tr}\left(\rho^{ij}_{AB}\, CHSH(\phi^i_A, \phi^j_B)\right) = S_{ij} \ - 2\epsilon_0 \\ & \phi^i_A, \phi^j_B \in [0, \pi/2], \\ & \rho^{ij}_{AB} \succeq 0 \\ & \mathrm{Tr}(\rho^{ij}_{AB}) = 1 \end{aligned}} \tag{III.27}$$

F. Convex lower bound of $C^*(S)$

The optimization problem defining $n^*(S)$ is given as,

$$\begin{aligned} & n^*(S_{ij}) = \inf\left\{\lambda \left\|\rho^{ij}_{AB} - \Lambda_0\left[\rho^{ij}_{AB}\right]\right\|_F^2 + (1-\lambda)\left\|\rho^{ij}_{AB} - \Lambda_1\left[\rho^{ij}_{AB}\right]\right\|_F^2 + \frac{\mu}{2}\left\|\rho^{ij}_{AB}\right\|_F^2\right\} \\ & s.t.\ \mathrm{Tr}\left(\rho^{ij}_{AB}\, CHSH(\phi^i_A, \phi^j_B)\right) = S_{ij}. \end{aligned} \tag{III.28}$$

Now, the given objective function involving the Frobenius norm $\|\cdot\|_F^2$ is strongly convex in ρ. The term $\frac{\mu}{2}\left\|\rho^{ij}_{AB}\right\|_F^2$ introduces μ-strong convexity, and the linear constraints of the optimization problem preserve the strong convexity on the feasible set. For fixed S_{ij}, the objective function is strongly convex in ρ. The parameterized problem's solution function $n^*(S_{ij})$ inherits convexity. Moreover, strong convexity implies quadratic dependence on the perturbations in S_{ij}. Thus, $n^*(S)$ is strongly convex in S for $S \in (2, 2\sqrt{2}]$ For the measure η with,

$$\int_2^{2\sqrt{2}} \eta(dS') = 1, \quad \eta \geq 0, \quad \int_2^{2\sqrt{2}} \eta(dS')S' = S \tag{III.29}$$

Now application of Jensen's inequality for strongly convex functions yields,

$$\int_2^{2\sqrt{2}} \eta(dS')n^*(S') \geq n^*\left(\int_2^{2\sqrt{2}} \eta(dS')S'\right) + \frac{\mu}{2}\,\mathrm{Var}_\eta(S') \tag{III.30}$$

Since $\mathrm{Var}_\eta(S') \geq 0$,

$$\int_2^{2\sqrt{2}} \eta(dS')n^*(S') \geq n^*(S). \tag{III.31}$$

Substituting $\overline{C}(S') = n^*(S')$ into the original inequality, we get,

$$C^*(S) \geq \int_2^{2\sqrt{2}} \eta(dS')n^*(S') \geq n^*(S) \tag{III.32}$$

Therefore, $\overline{C}(S) = n^*(S)$ is a valid convex lower bound.

IV. CONCLUSION

In most of the QKD protocols, Alice and Bob single out one of their measurement settings for key generation. This strategy helps them improving the key rate of those protocols. In the original DIQKD protocol [ABG+07], Alice and Bob chose a single measurement setting $\{A_0, B_0\}$ with high probability for key generation, whereas, other measurement settings with low probability for testing the channels. However, the security analysis of this protocol shows that there exists optimal attack such that $H(A_0|E) \leq H(A_1|E)$ [PAB+09]. In other words, as Eve has the information about Alice's measurement settings, she could focus on minimizing her uncertainty about the key generating measurement A_0 at the cost of having higher uncertainty about the other setting. To circumvent this problem (to increase Eve's uncertainty about the key generation measurement setting), DIQKD with random key basis is proposed. In this variant of DIQKD protocol, as Alice uses two measurement settings for key generation, the uncertainty of Eve about Alice's measurement settings is now a combination of conditional entropies, $\lambda H(A_0|E) + (1-\lambda)H(A_1|E)$, where λ is the probability for choosing the measurement A_0 and E is Eve's side information. The main challenge in the security analysis of this protocol is to find a reliable lower bound on this quantity using only the observed CHSH violation. The authors in[SGP+21] have framed the problem onto an optimization problem and found a lower bound $C^\star(S)$, such that $\lambda H(A_0|E) + (1-\lambda)H(A_1|E) \geq C^\star(S)$. The results produced in this manuscript show that the optimization cost involved in the security analysis of the DIQKD protocol can be reduced. In the original work, polytope optimization has been used for finding Bob's optimal angle in context to the main optimization problem. Our results show that reframing the security analysis as a strongly convex optimization problem can help in optimizing Bob's measurement angles using the same approach as used for Alice, i.e., ϵ-net approximation. Now, as polytope optimization is costlier than ϵ-net approximation, substituting the earlier with the later in the main problem results to lesser cost for the security analysis. Evaluation of the pessimistic error involved in the $\epsilon-$net approximation is an important task because in every iteration the error is being subtracted from the CHSH score S. However, so far, any explicit expression for the pessimistic error has not been noticed in any literature. Here, we have derived an explicit expression for the pessimistic error. It is found to be the maximised value of the product of two terms, namely, desired precision of the angles ϕ^i_A and ϕ^j_B and the first-order derivative of the solution functions of the optimization problem. Our result shows how Alice and Bob can determine their measurement angles ϕ^i_A and ϕ^j_B from the spectrum of the projectors they chose for checking CHSH violation and generating secret keys. The dependence of CHSH operator on the angles of the parties were not clearly explained in the original security analysis. Our results also show how the operator varies with the change in the angles of Alice and Bob, determining the instance that results to maximum deviation in the CHSH operator. Moreover, from the results derived in this manuscript, it can be concluded that a convex lower bound on the uncertainty of Eve about Alice's key generation measurements can be derived analytically. We believe that our work has made the security analysis of the DIQKD with random key basis more clear and complete.

V. ACKNOWLEDGMENT

SB would like to acknowledge the University of Calcutta for providing the opportunity to work at TCG CREST. He is also grateful to team of CQuERE, TCG CREST, for fruitful discussion.

[ABG+07] 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.

[AL70] Huzihiro Araki and Elliott H. Lieb. Entropy inequalities. *Communications In Mathematical Physics*, 18(2):160–170, June 1970.

[BCC+10] 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.

[BCP+14] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, Apr 2014.

[BS10] 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.

[BTN01] 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.

[BV04] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[CBTW17] Patrick J. Coles, Mario Berta, Marco Tomamichel, and Stephanie Wehner. Entropic uncertainty relations and their applications. *Rev. Mod. Phys.*, 89:015002, Feb 2017.

[CM11] Marcos Curty and Tobias Moroder. Heralded-qubit amplifiers for practical device-independent quantum key distribution. *Phys. Rev. A*, 84:010304, Jul 2011.

[CS12] Adán Cabello and Fabio Sciarrino. Loophole-free bell test based on local precertification of photon's presence. *Phys. Rev. X*, 2:021010, Jun 2012.

[GPS10] 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.

[Hol19] Alexander S. Holevo. *Quantum Systems, Channels, Information*. De Gruyter, Berlin, Boston, 2019.

[HST$^+$20] 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.

[KMS$^+$20] 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.

[LPT$^+$13] 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.

[Mam24] Z.Z. Mammeri. *Cryptography: Algorithms, Protocols, and Standards for Computer Security*. Wiley, 2024.

[MSMG$^+$16] 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.

[NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[PAB$^+$09] 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.

[PP08] K. B. Petersen and M. S. Pedersen. The matrix cookbook, October 2008. Version 20081110.

[Ras15] Ioan Rasa. Concavity of some entropies, 2015.

[Ren06] Renato Renner. Security of quantum key distribution, 2006.

[SGP$^+$21] 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.

[Sik12] Jamie William Jonathon Sikora. Analyzing quantum cryptographic protocols using optimization techniques. 2012.

[TBB08] B.S. Thomson, A.M. Bruckner, and J.B. Bruckner. *Elementary Real Analysis*. Number v. 1. www.classicalrealanalysis.com., 2008.

[Wat18] J. Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.

[Wil13] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2013.

[XZZP22] 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.

APPENDIX

Appendix A: Framework

We first describe the framework for phrasing the security analysis as an optimization problem. The aim of the analysis is to establish a lower bound on the quantity $C^*(S)$ [SGP+21], where S is the observed CHSH value for the shared entangled state between the parties Alice and Bob, such that

$$\lambda H(A_0|E)_{\rho_{A'BEA_0}} + (1-\lambda)H(A_1|E)_{\rho_{A'BEA_1}} \geq C^*(S), \tag{A.1}$$

where A_X is Alice's outcome for her measurement setting $X \in \{0,1\}$ and $H(A_X|E)_{\rho_{A'BEA_X}}$ signifies Eve's uncertainty about Alice's outcomes given her side-channel quantum information E. The protocol assumes the measurement setting for Bob to be $Y \in \{0,1,2,3\}$. Let $\mathcal{X}$ and $\mathcal{Y}$ be the Hilbert spaces defining Alice's and Bob's subsystems, respectively. Now consider $O_x^{\mathcal{X}} \in \mathrm{Herm}(\mathcal{X})$ for $x \in \{0,1\}$ and $O_y^{\mathcal{Y}} \in \mathrm{Herm}(\mathcal{Y})$ for $y \in \{0,1,2,3\}$, respectively, to be the observables corresponding to Alice's and Bob's measurement settings. The observables can be explicitly written as $\mathcal{A} = \{O_0^{\mathcal{X}}, O_1^{\mathcal{X}}\}$ and $\mathcal{B} = \{O_0^{\mathcal{Y}}, O_1^{\mathcal{Y}}, O_2^{\mathcal{Y}}, O_3^{\mathcal{Y}}\}$. These can be written in spectral form as,

$$O_x^{\mathcal{X}} = \sum_i \lambda_i^x P_i^x, \quad O_y^{\mathcal{Y}} = \sum_j \lambda_j^y P_j^y \tag{A.2}$$

where λ_i^x and λ_j^y are eigenvalues corresponding to the projectors P_i^x and P_j^y, respectively. As the measurement outcomes are dichotomic, we can also express the observables as,

$$O_x^{\mathcal{X}} = P_{\mathcal{X}}^{0|x} - P_{\mathcal{X}}^{1|x}, \quad O_y^{\mathcal{Y}} = P_{\mathcal{Y}}^{0|y} - P_{\mathcal{Y}}^{1|y} \tag{A.3}$$

The measurements settings chosen by Alice and Bob to check the CHSH violation are $\{A_0, A_1\}$ and $\{B_2, B_3\}$, respectively. Therefore, the correlation measurement operators corresponding to the chosen measurement settings can be defined as,

$$\begin{aligned} C^{O_0^{\mathcal{X}}} &= P_{\mathcal{X}}^{0|0} - P_{\mathcal{X}}^{1|0} \quad \text{and} \quad C^{O_1^{\mathcal{X}}} = P_{\mathcal{X}}^{0|1} - P_{\mathcal{X}}^{1|1}, \\ C^{O_2^{\mathcal{Y}}} &= P_{\mathcal{Y}}^{0|2} - P_{\mathcal{Y}}^{1|2} \quad \text{and} \quad C^{O_3^{\mathcal{Y}}} = P_{\mathcal{Y}}^{0|3} - P_{\mathcal{Y}}^{1|3} \end{aligned} \tag{A.4}$$

The correlation function defined in the main text, $C_{XY} = P(A_X = B_Y|X,Y) - P(A_X \neq B_Y|X,Y)$, is related to the correlation measurement operators via $C_{XY} = \mathrm{Tr}(\rho\, P_{\mathcal{X}}^{a|x} \otimes P_{\mathcal{Y}}^{b|y})$, where ρ is the state on which Alice and Bob perform the measurement. Hence, the CHSH operator can be written in terms of correlation measurement operators as [SGP+21],

$$\begin{aligned} CHSH \;:=\; & C^{O_1^{\mathcal{X}}} \otimes C^{O_2^{\mathcal{Y}}} - C^{O_0^{\mathcal{X}}} \otimes C^{O_2^{\mathcal{Y}}} \\ - \; & C^{O_0^{\mathcal{X}}} \otimes C^{O_3^{\mathcal{Y}}} - C^{O_1^{\mathcal{X}}} \otimes C^{O_3^{\mathcal{Y}}} \end{aligned} \tag{A.5}$$

Each term in the operator can be expressed in terms of the projectors defined above as [Wil13],

$$\begin{aligned} C^{O_x^{\mathcal{X}}} \otimes C^{O_y^{\mathcal{Y}}} &= (P_{\mathcal{X}}^{0|x} - P_{\mathcal{X}}^{1|x}) \otimes (P_{\mathcal{Y}}^{0|y} - P_{\mathcal{Y}}^{1|y}) \\ &= (P_{\mathcal{X}}^{0|x} \otimes P_{\mathcal{Y}}^{0|y}) - (P_{\mathcal{X}}^{0|x} \otimes P_{\mathcal{Y}}^{1|y}) - (P_{\mathcal{X}}^{1|x} \otimes P_{\mathcal{Y}}^{0|y}) + (P_{\mathcal{X}}^{1|x} \otimes P_{\mathcal{Y}}^{1|y}) \\ &= (P_{\mathcal{X}}^{0|x} \otimes P_{\mathcal{Y}}^{0|y} + P_{\mathcal{X}}^{1|x} \otimes P_{\mathcal{Y}}^{1|y}) - (P_{\mathcal{X}}^{0|x} \otimes P_{\mathcal{Y}}^{1|y} + P_{\mathcal{X}}^{1|x} \otimes P_{\mathcal{Y}}^{0|y}) \end{aligned} \tag{A.6}$$

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 Ψ_{ABE}. A more particular description of the state may include a fourth component – Alice's outcome A_X, which is initially associated with a pure quantum state ψ, and hence the joint state between Alice, Bob, and Eve can be expressed as Ψ_{ABEA_X}. 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. However, in this protocol, Eve can not determine beforehand the measurement setting used by Alice to generate the secret key – Alice now uses two random bases for key generation.

The purpose of optimising the objective function defined in (A.11) is to find a lower bound on the uncertainty of Eve about Alice's measurement outcomes. The conditional von Neumann entropy function defined in the Eq.(A.1) 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 [Hol19] between Alice and Eve. Similarly, the von Neumann entropy can be expressed 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 $I(A_X;B)_{\rho_{A'BEA_X}}$ and $H(A_X|B)_{\rho_{A'BEA_X}}$ are respectively the mutual information and conditional entropy between Alice and Bob. Using the above two relations, Eq. (A.1) can be rewritten as,

$$\begin{aligned} C^*(S) \leq \lambda &\left[I(A_0;B)_{\rho_{A'BEA_0}} + H(A_0|B)_{\rho_{A'BEA_0}} - \chi(A_0;E)\right] \\ &+ (1-\lambda)\left[I(A_1;B)_{\rho_{A'BEA_1}} + H(A_1|B)_{\rho_{A'BEA_1}} - \chi(A_1;E)\right] \end{aligned} \tag{A.7}$$

Under the assumption of optimal error correction ($H(A_X|B_Y)_{\rho_{A'BEA_X}} \to 0$), the above inequality simplifies to:

$$C^*(S) \leq \lambda\left[I(A_0;B)_{\rho_{A'BEA_0}} - \chi(A_0;E)\right] + (1-\lambda)\left[I(A_1;B)_{\rho_{A'BEA_1}} - \chi(A_1;E)\right]. \tag{A.8}$$

The quantum-classical state obtained after Alice's measurement is given by,

$$\rho_{A'BEA_X} = \sum_{i\in\{0,1\}} \rho_{\psi_i} \otimes \mathrm{Tr}_A((P_{\mathcal{X}}^{i|x} \otimes \mathbb{1}_{BE})\Psi_{ABE}(P_{\mathcal{X}}^{i|x} \otimes \mathbb{1}_{BE})^*) \tag{A.9}$$

where A', B, E, A_X are Alice's subsystem, Bob's subsystem, Eve's subsystem, and Alice's measurement outcome associated with pure quantum state ψ, respectively. Using this state, if the conditional Von Neumann entropy $H(A_X|E)_{\rho_{A'BEA_X}}$ is calculated, then we find that it essentially captures the change in entropy of the joint state of Alice and Bob after and before Alice's measurement,

$$\begin{aligned} H(A_X|E)_{\rho_{A'BEA_X}} &= H(A_XE)_{\rho_{A'BEA_X}} - H(E)_{\rho_{ABEA'_X}} \\ &= H(A'B)_{\rho_{A'BEA_X}} - H(AB)_{\rho_{ABEA'_X}} \\ &= \Delta H_X \end{aligned} \tag{A.10}$$

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), i.e., $\rho_{AB} = \mathrm{Tr}_{EA'_X}(\rho_{ABEA'_X})$. Now one can formally define the optimization problem as

$$\boxed{\begin{array}{rll} C^*(S) = \inf & \lambda H(A_0|E)_{\rho_{A'BEA_0}} + (1-\lambda)H(A_1|E)_{\rho_{A'BEA_1}} \\ \text{s.t.} & \mathrm{Tr}(\rho_{AB}\cdot \mathrm{CHSH}) = S \\ & \rho_{AB} \succeq 0 \\ & \mathrm{Tr}(\rho_{AB}) = 1 \\ \text{or equivalently,} & \\ C^*(S) = \inf & \lambda\left[I(A_0;B)_{\rho_{A'BEA_0}} - \chi(A_0;E)\right] + (1-\lambda)\left[I(A_1;B)_{\rho_{A'BEA_1}} - \chi(A_1;E)\right] \\ \text{s.t.} & \mathrm{Tr}(\rho_{AB}\cdot \mathrm{CHSH}) = S \\ & \rho_{AB} \succeq 0 \\ & \mathrm{Tr}(\rho_{AB}) = 1 \end{array}} \tag{A.11}$$

Appendix B: Reformulation of change in entropy ΔH_X in terms of Alice and Bob subsystems

After formally defining the optimization problem, our first task is to reformulate the problem in terms of the Alice's and Bob's subsystems A and B. The problem defined in (A.11) is defined in terms of Alice's outcome A_X and Eve's subsystem E. This 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's subsystem can be described by a vector space, whose dimension is unknown. 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 makes no sense in the DIQKD framework.

The global state ρ is a pure state; one may consider it as a bipartite state where A_X and E are constitute one part, and A and B the other. Now, from the result of *Theorem 2.c* of [AL70], one can easily find that the local entropy production ΔH_X on the two sides of the bipartition is essentially the same,

$$\Delta H_X = H(A'B)_{\rho_{A'BEA_X}} - H(AB)_{\rho_{ABEA'_X}} \tag{B.1}$$

Hence, using Eq. (A.10), one can perform the security analysis in terms of accessible Alice and Bob subsystems.[SGP$^+$21] 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 measurement outcome 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 classical register storing the Alice's measurement outcomes. The work of [SGP$^+$21] shows that this transformation can indeed be defined through pinching channels defined for each observable as,

$$\begin{aligned}
\Lambda_0[\rho_{ABEA'_0}] &= (P^{0|0}_{\mathcal{X}} \otimes \mathbb{1})\rho_{ABEA'_0}(P^{0|0}_{\mathcal{X}} \otimes \mathbb{1}) + (P^{1|0}_{\mathcal{X}} \otimes \mathbb{1})\rho_{ABEA'_0}(P^{1|0}_{\mathcal{X}} \otimes \mathbb{1}) = \rho_{A'BEA_0} \\
\Lambda_1[\rho_{ABEA'_1}] &= (P^{0|1}_{\mathcal{X}} \otimes \mathbb{1})\rho_{ABEA'_1}(P^{0|1}_{\mathcal{X}} \otimes \mathbb{1}) + (P^{1|1}_{\mathcal{X}} \otimes \mathbb{1})\rho_{ABEA'_1}(P^{1|1}_{\mathcal{X}} \otimes \mathbb{1}) = \rho_{A'BEA_1}
\end{aligned} \tag{B.2}$$

Now let us redefine the reduced joint state of Alice and Bob as,

$$\begin{aligned}
\mathrm{Tr}_{EA_X}(\Lambda[\rho_{ABEA'_X}]) &= \rho_{A'B} \\
\mathrm{Tr}_{EA_X}(\rho_{ABEA'_X}) &= \rho_{AB}
\end{aligned} \tag{B.3}$$

1. Pinching Channel

The pinching channel Λ is defined as:

$$\Lambda[\rho] = \sum_i P_i \rho P_i \tag{B.4}$$

where $\{P_i\}$ is a set of orthogonal projectors and $\rho \in$ is a density operator.

Properties of Pinching Channels

1. Idempotence: The pinching channel is idempotent, i.e.,

$$\Lambda[\Lambda[\rho]] = \Lambda[\rho]. \tag{B.5}$$

Proof:

$$\Lambda[\Lambda[\rho]] = \sum_i P_i \left(\sum_j P_j \rho P_j \right) P_i. \tag{B.6}$$

Using the orthogonality of the projectors ($P_i P_j = \delta_{ij} P_i$), this simplifies to:

$$\Lambda[\Lambda[\rho]] = \sum_i P_i \rho P_i = \Lambda[\rho]. \tag{B.7}$$

Thus, Λ is idempotent.

2. Any pinching channel acts as an identity operator if ρ commutes with all P_i, i.e.,

$$\Lambda[\rho] = \rho, \qquad \forall \rho \quad if [\rho, P_i] = 0 \tag{B.8}$$

Proof: If ρ commutes with P_i, then $P_i \rho = \rho P_i$. Substituting into the definition of Λ:

$$\Lambda[\rho] = \sum_i P_i \rho P_i = \sum_i P_i P_i \rho = \sum_i P_i \rho = \rho, \tag{B.9}$$

where one can used $P_i^2 = P_i$ (since P_i is a projector) and $\sum_i P_i = I$.

3. Pinching channel does not increase trace distance between any two states, i.e., for any two states ρ and σ,

$$\|\Lambda[\rho] - \Lambda[\sigma]\|_1 \leq \|\rho - \sigma\|_1. \tag{B.10}$$

Proof: The trace distance is defined as:

$$\|\rho - \sigma\|_1 = Tr|\rho - \sigma|, \tag{B.11}$$

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:

$$\|\Lambda[\rho] - \Lambda[\sigma]\|_1 \leq \|\rho - \sigma\|_1. \tag{B.12}$$

4. 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:

$$\mathrm{Tr}(P_i\Lambda[\rho]) = \mathrm{Tr}\left(P_i \sum_j P_j \rho P_j\right) = \mathrm{Tr}(P_i\rho P_i) = \mathrm{Tr}(P_i\rho) \tag{B.13}$$

one can use the trace's cyclic property and $P_iP_j = \delta_{ij}P_i$. Thus, the diagonal terms are preserved.

5. Pinching channel increases the von Neumann entropy, ie.,

$$S(\Lambda[\rho]) \geq S(\rho), \tag{B.14}$$

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:

$$S(\Lambda[\rho]) \geq S(\rho). \tag{B.15}$$

6. The eigenvalues of $\Lambda[\rho]$ are majorized 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:

$$\lambda(\Lambda[\rho]) \prec \lambda(\rho). \tag{B.16}$$

This means that the eigenvalues of $\Lambda[\rho]$ are more uniformly distributed than those of ρ.

7. For any measurable function f, the pinching channel satisfies:

$$f(\Lambda[\rho]) = \Lambda[f(\rho)]. \tag{B.17}$$

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:

$$f(\Lambda[\rho]) = \Lambda[f(\rho)]. \tag{B.18}$$

In quantum information theory, the dual (adjoint) of a map Φ, denoted Φ^*, is defined via the Hilbert-Schmidt inner product:

$$\mathrm{Tr}(A \cdot \Phi[\rho]) = \mathrm{Tr}(\Phi^*[A] \cdot \rho), \quad \forall A, \rho \in L(\mathcal{V}), \tag{B.19}$$

where A and ρ are Hermitian operators on a complex Euclidean space $\mathcal{V}$. More generally, for any linear map Φ, the adjoint satisfies:

$$\mathrm{Tr}(A \cdot \Phi[B]) = \mathrm{Tr}(\Phi^*[A] \cdot B), \quad \forall A, B \in L(\mathcal{V}), \tag{B.20}$$

where A and B are Hermitian operators on a complex Euclidean space $\mathcal{V}$. Now since pinching channel Λ are self adjoint from (B.20) one can have,

$$\mathrm{Tr}(A \cdot \Lambda[\rho]) = \mathrm{Tr}(\rho \cdot \Lambda^*[A]) = \mathrm{Tr}(\rho \cdot \Lambda[A]) \quad \forall A, \rho \in L(\mathcal{V}) \tag{B.21}$$

2. Entropy production

Now, one can find a mathematical expression for the production of entropy ΔH_X as

$$\begin{aligned}\Delta H_X &= H(A'B)_{\rho_{A'BEA_X}} - H(AB)_{\rho_{ABEA'_X}} \\ &= H(\mathrm{Tr}_{EA_X}(\Lambda[\rho_{ABEA'_X}])) - H(\mathrm{Tr}_{EA'_X}(\rho_{ABEA'_X})) \\ &= H(\Lambda[\rho_{AB}]) - H(\rho_{AB}) \end{aligned} \tag{B.22}$$

Using Eqs. ((B.5),(B.21)) and putting $A = \log_2(\Lambda[\rho_{AB}])$ and $B = \rho_{AB}$, we get,

$$\begin{aligned}\Delta H_X &= -\mathrm{Tr}(\Lambda[\rho_{AB}]\log_2(\Lambda[\rho_{AB}])) + \mathrm{Tr}(\rho_{AB}\log_2(\rho_{AB})) \\ &= -\mathrm{Tr}(\rho_{AB}\Lambda^*(\log_2(\Lambda[\rho_{AB}]))) + \mathrm{Tr}(\rho_{AB}\log_2(\rho_{AB})) \\ &= -\mathrm{Tr}(\rho_{AB}\log_2(\Lambda[\rho_{AB})) + \mathrm{Tr}(\rho_{AB}\log_2(\rho_{AB})) \\ &= \mathrm{Tr}(\rho_{AB}\log_2(\rho_{AB})) - Tr(\rho_{AB}\log_2(\Lambda[\rho_{AB})) \\ &= \mathrm{Tr}(\rho_{AB}(\log_2(\rho_{AB}) - \log_2(\Lambda[\rho_{AB}]))) \\ &= D(\rho_{AB}||(\Lambda[\rho_{AB}])) \end{aligned} \tag{B.23}$$

Here the relative entropy D is defined as $D(\rho||\sigma) = \mathrm{Tr}(\rho(\log(\rho) - \log(\sigma)))$ [Wil13]. Our main objective function in (A.11) is a convex combination of conditional von Neumann entropies. Using (A.10) and (B.23) the objective function can thus be reformulated as,

$$\begin{aligned} C^*(S) = \inf \quad & \lambda D(\rho_{AB}||(\Lambda_0[\rho_{AB})) + (1-\lambda)D(\rho_{AB}||(\Lambda_1[\rho_{AB}])) \\ \text{s.t.} \quad & \mathrm{Tr}\,(\rho_{AB}\cdot \mathrm{CHSH}) = S \\ & \rho_{AB} \succeq 0 \\ & \mathrm{Tr}(\rho_{AB}) = 1 \end{aligned} \tag{B.24}$$

Appendix C: Reduction of the optimization problem to two qubit space $\mathbb{C}_{4\times4}$ of Alice and Bob

Let $\rho_{A'BEA_x}, \rho_{ABEA'_x}, \in \mathcal{D}(\mathcal{P})$ for Hilbert space $\mathcal{P}$ over which the density operator space is defined. The dimension of $D(\mathcal{P})$ is unknown as the dimension of Eve's subsystem E is unknown according to the standard assumptions of security analysis. Analyzing the subsystem of Alice and Bob in a larger, unknown dimension is not only impractical but also complicates the analysis – larger systems are inherently more susceptible to attacks due to their increased complexity. The work of [SGP$^+$21] shows that without loss of generality, one can can reduce the problem to two-qubit space $\mathbb{C}_{4\times4}$ of Alice and Bob. From the result of [BS10], one can decompose a projector of higher dimension into projectors of dimension 2 × 2 acting on either Alice's or Bob's local states. Therefore, the four pairs of projectors as in (F.7) can be decomposed accordingly. Let L^z_S be the set of pairwise-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 $L^z_S = \{L^0_A, L^1_A, L^2_B, L^3_B\}$. Now let us consider the following projector of dimension 2 × 2 and being parameterized by angle θ

$$Q(\theta) = \begin{pmatrix} \cos^2\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) \\ \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) & \sin^2\left(\frac{\theta}{2}\right) \end{pmatrix} \tag{C.1}$$

Having defined the local projectors as above, one can decompose $P^{0|0}_{\mathcal{X}}, P^{0|1}_{\mathcal{X}}, P^{0|2}_{\mathcal{Y}}$ and $P^{0|3}_{\mathcal{Y}}$ as,

$$\begin{aligned} P^{0|0}_{\mathcal{X}} = \bigoplus_i Q(0) \oplus L^0_A, \quad P^{0|1}_{\mathcal{X}} = \bigoplus_i Q(\phi^i_A) \oplus L^1_A \\ P^{0|2}_{\mathcal{Y}} = \bigoplus_j Q(0) \oplus L^2_B, \quad P^{0|3}_{\mathcal{Y}} = \bigoplus_j Q(\phi^j_B) \oplus L^3_B \end{aligned} \tag{C.2}$$

The remaining projectors can be obtained using the completeness property and thus,

$$\begin{aligned} P^{1|0}_{\mathcal{X}} = \mathbb{1} - P^{0|0}_{\mathcal{X}}, \quad P^{1|1}_{\mathcal{X}} = \mathbb{1} - P^{0|1}_{\mathcal{X}} \\ P^{1|2}_{\mathcal{Y}} = \mathbb{1} - P^{0|2}_{\mathcal{Y}}, \quad P^{1|3}_{\mathcal{Y}} = \mathbb{1} - P^{0|3}_{\mathcal{Y}} \end{aligned} \tag{C.3}$$

The angles $\{\phi^i_A\}$ and $\{\phi^j_B\}$ can be obtained from the spectrums of $P^{0|0}_{\mathcal{X}} + P^{0|1}_{\mathcal{X}}$ and $P^{0|2}_{\mathcal{Y}} + P^{0|3}_{\mathcal{Y}}$, respectively. As the projectors are block-diagonal operators, they are expressed as direct sum of commuting and noncommuting parts. The commuting parts contribute discrete and angle-independent eigenvalues, while the noncommuting parts contribute angle-dependent eigenvalues. Thus, the angles ϕ^i_A, ϕ^j_B can be uniquely determined from the non-integer eigenvalues in the spectrum. Using Eq. (C.1,C.2), we compute the angles ϕ^i_A and ϕ^i_A from the eigenvalues as,

$$\begin{aligned} \cos(\phi_A/2) &= \frac{\lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}}}{2} \text{or, } \phi_A = 2\arccos\left(\frac{\lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}}}{2}\right). \\ \cos(\phi_B/2) &= \frac{\lambda^1_{\mathcal{Y}} - \lambda^2_{\mathcal{Y}}}{2} \text{or, } \phi_A = 2\arccos\left(\frac{\lambda^1_{\mathcal{Y}} - \lambda^2_{\mathcal{Y}}}{2}\right). \end{aligned} \tag{C.4}$$

where $\lambda^1_{\mathcal{X}} = 1 + \cos(\phi^i_A/2)$, $\lambda^2_{\mathcal{X}} = 1 - \cos(\phi^i_A/2)$ $\left(P^{0|2}_{\mathcal{Y}} + P^{0|3}_{\mathcal{Y}}\right)$, $\lambda^1_{\mathcal{Y}} = 1 + \cos(\phi^j_B/2)$, and $\lambda^2_{\mathcal{Y}} = 1 - \cos(\phi^j_B/2)$. The mapping $\phi^m_N \mapsto 1 \pm \cos(\phi^m_N/2)$ where $N \in \{\mathcal{X}, \mathcal{Y}\}$ and $m \in \{i, j\}$is bijective for $\phi^m_N \in [0, 2\pi)$ when restricted to non-integer eigenvalues. Moreover the inverse function *arccos* is unique in $[0, \pi]$.

Lemma 1: The mapping $\phi^m_N \mapsto 1 \pm \cos(\phi^m_N/2)$, where $N \in \{\mathcal{X}, \mathcal{Y}\}$ and $m \in \{i, j\}$ is bijective for $\phi^m_N \in [0, \pi]$

Proof: Let $h(\phi^m_N) = 1 \pm \cos(\phi^m_N/2)$ where $N \in \{\mathcal{X}, \mathcal{Y}\}$ and $m \in \{i, j\}$

1. *Proof of injectivity:* Let us assume that $\phi^1_N \neq \phi^2_N \in [0, \pi]$ where ϕ^1_N, ϕ^2_N are arbitrarily two angles obtained from spectrum of the sum of either $\left(P^{0|0}_{\mathcal{X}} + P^{0|1}_{\mathcal{X}}\right)$ or $\left(P^{0|2}_{\mathcal{Y}} + P^{0|3}_{\mathcal{Y}}\right)$

$$\cos(\phi^1_N/2) \neq \cos(\phi^2_N/2) \text{or, } \lambda^1_{\mathcal{X}} \neq \lambda^2_{\mathcal{X}} \quad \text{or} \quad (\lambda^1_{\mathcal{Y}} \neq \lambda^2_{\mathcal{Y}}) \text{or, } h(\phi^1_N) \neq h(\phi^2_N) \tag{C.5}$$

Thus f is an injective function.

2. *Proof of Surjectivity:* The eigenvalues $\lambda^1_{\mathcal{X}}, \lambda^2_{\mathcal{X}}$ or $(\lambda^1_{\mathcal{Y}}, \lambda^2_{\mathcal{Y}}) \in [0, 2]$ with $\lambda^1_{\mathcal{X}} + \lambda^2_{\mathcal{X}} = 2$ or $(\lambda^1_{\mathcal{Y}} + \lambda^2_{\mathcal{Y}})$ and $\lambda^1_{\mathcal{X}} > \lambda^2_{\mathcal{X}}$ or $(\lambda^1_{\mathcal{Y}} > \lambda^2_{\mathcal{Y}})$ there exists a unique $\phi^m_N \in [0, \pi]$ where $m \in \{i, j\}$ such that the eigenvalues corresponding to ϕ^m_N are $\lambda^1_{\mathcal{X}}, \lambda^2_{\mathcal{X}}$ or $(\lambda^1_{\mathcal{Y}}, \lambda^2_{\mathcal{Y}})$ for $N = A$ or B respectively.
The condition of $\lambda^1_{\mathcal{X}} + \lambda^2_{\mathcal{X}} = 2$ or $(\lambda^1_{\mathcal{Y}} + \lambda^2_{\mathcal{Y}})$ arises from a fundamental result of matrix algebra, *trace of a matrix is the sum of its eigenvalues* and using this result we get,

$$\begin{aligned} \mathrm{Tr}(L) &= 1 + \cos^2(\phi^i_A/2) + \sin^2(\phi^i_A/2) \\ &= 2 \\ &= \lambda^1_{\mathcal{X}} + \lambda^2_{\mathcal{X}} \end{aligned} \tag{C.6}$$

and equivalently

$$\lambda^1_{\mathcal{Y}} + \lambda^2_{\mathcal{Y}} = 2 \tag{C.7}$$

Given the above three criteria, one can have,

$$\begin{aligned} c_{\mathcal{X}} &= \frac{\lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}}}{2} \\ c_{\mathcal{Y}} &= \frac{\lambda^1_{\mathcal{Y}} - \lambda^2_{\mathcal{Y}}}{2} \end{aligned} \tag{C.8}$$

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 $\lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}} > 0$ and $\lambda^1_{\mathcal{X}} = 2 - \lambda^2_{\mathcal{X}}$. Thus,

$$c_{\mathcal{X}} = \frac{2 - \lambda^2_{\mathcal{X}} - \lambda^2_{\mathcal{X}}}{2} = \frac{2 - 2\lambda^2_{\mathcal{X}}}{2} = 1 - \lambda^2_{\mathcal{X}} \tag{C.9}$$

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_A^i*:* Now using (C.4) and (C.8) one can have,

$$\phi_A = 2\arccos(c_{\mathcal{X}}) \tag{C.10}$$

Since $c_{\mathcal{X}} \in (0,1)$, $\arccos(c_{\mathcal{X}}) \in (0, \frac{\pi}{2})$. Therefore, $\phi_A^i = 2\arccos(c_{\mathcal{X}}) \in (0,\pi)$.

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,

$$\begin{aligned} \lambda^1_{\mathcal{X}}, \lambda^2_{\mathcal{X}} &= 1 \pm \cos(\phi_A^i/2) \\ \lambda^1_{\mathcal{Y}}, \lambda^2_{\mathcal{Y}} &= 1 \pm \cos(\phi_B^j/2) \end{aligned} \tag{C.11}$$

Substituting $\phi_A^i = 2\arccos(c_{\mathcal{X}})$, one can get:

$$\lambda^1_{\mathcal{X}}, \lambda^2_{\mathcal{X}} = 1 \pm \cos(\arccos(c_{\mathcal{X}})) = 1 \pm c_{\mathcal{X}} \tag{C.12}$$

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,

$$\begin{aligned} \lambda^1_{\mathcal{X}} &= 1 + \frac{\lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}}}{2} \text{or, } 2\lambda^1_{\mathcal{X}} = 2 + \lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}} \text{or, } \lambda^1_{\mathcal{X}} = 2 - \lambda^2_{\mathcal{X}} \\ \lambda^2_{\mathcal{X}} &= 1 - \frac{\lambda^1_{\mathcal{X}} - \lambda^2_{\mathcal{X}}}{2} \text{or, } 2\lambda^2_{\mathcal{X}} = 2 - \lambda^1_{\mathcal{X}} + \lambda^2_{\mathcal{X}} \text{or, } \lambda^2_{\mathcal{X}} = 2 - \lambda^1_{\mathcal{X}} \end{aligned} \tag{C.13}$$

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_A^i are $\lambda^1_{\mathcal{X}}$ and $\lambda^2_{\mathcal{X}}$. Similarly by substituting $\phi_B^j = 2arccos(c_{\mathcal{Y}})$ in (C.11) we prove the surjectivity of the mapping. Since the map is both injective and surjective, this completes the proof that it is bijective.

□

1. Computing lower bound of $C^*(S)$

Having established the functional relationship between the angles ϕ_A^i and ϕ_B^j and characterized the spectral properties of the operators $P^{0|0}_{\mathcal{X}} + P^{0|1}_{\mathcal{X}}$ (or equivalently, $P^{0|2}_{\mathcal{Y}} + P^{0|3}_{\mathcal{Y}}$), we now proceed to find a lower bound of the function $C^*(S)$. Let $\Lambda_{\mathcal{X}}$ and $\Lambda_{\mathcal{Y}}$ be the Pinching channels that act on Alice's and Bob's subsystems and decompose Alice's and Bob's subsystems into block structures as would be obtained from the effect of the projectors in Eq. (C.2).

Now let $\Lambda_{\mathcal{X}\mathcal{Y}} = \Lambda_{\mathcal{X}} \otimes \Lambda_{\mathcal{Y}}$ be the channel acting on the combined state $\rho_{ABEA'_X}$, resulting to a joint state of Alice and Bob corresponding to the angles ϕ_A^i and ϕ_B^j and a commuting part obtained from either $L^0_A \otimes L^1_A$ or $L^2_B \otimes L^3_B$. 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 $\rho_{ABEA'_X}$,

$$P = \{\rho^{ij}_{ABEA'_X} \mid Q(\phi_A^i) \otimes Q(\phi_B^j)(\rho_{ABEA'_X}) = \rho^{ij}_{ABEA'_X}\} \tag{C.14}$$

Thus,

$$\Lambda_{\mathcal{X}\mathcal{Y}}[\rho_{ABEA'_X}] = \bigoplus_{ij} (\eta_{ij}\rho^{ij}_{ABEA'_X}) \oplus (\eta_{\text{comm}}\rho^{\text{comm}}_{ABEA'_X}) \tag{C.15}$$

where η_{ij} are normalization factors obtained by normalizing each state $\rho^{ij}_{ABEA'_X} \in P$. Each of the states depends on the operators $Q(\phi_A^i) \otimes Q(\phi_B^j)$ and consequently on the angles obtained from respective spectra. The $\rho^{\text{comm}}_{ABEA'_X}$ are the projected blocks that commute with either Alice's subsystem or Bob's subsystem. Now, using the monotonicity property of relative entropy, we get from Eq. (B.24),

$$\begin{aligned}
C^*(S) = \inf \quad & \lambda D\left(\rho_{AB} \,||\, \Lambda_0[\rho_{AB}]\right) + (1-\lambda) D\left(\rho_{AB} \,||\, \Lambda_1[\rho_{AB}]\right) \\
& \geq \lambda D\left(\Lambda_{\mathcal{XY}}[\rho_{AB}] \,||\, \Lambda_0 \circ \Lambda_{\mathcal{XY}}[\rho_{AB}]\right) \\
& + (1-\lambda) D\left(\Lambda_{\mathcal{XY}}[\rho_{AB}] \,||\, \Lambda_1 \circ \Lambda_{\mathcal{XY}}[\rho_{AB}]\right) \\
& \geq \sum_{ij} \eta_{ij} \left[\lambda D\left(\rho^{ij}_{AB} \,||\, \Lambda_0[\rho^{ij}_{AB}]\right)\right. \\
& \qquad \left. + (1-\lambda) D\left(\rho^{ij}_{AB} \,||\, \Lambda_1[\rho^{ij}_{AB}]\right)\right] \\
& \qquad + \eta_{\text{comm}} \left[\lambda D\left(\rho^{\text{comm}}_{AB} \,||\, \Lambda_0[\rho^{\text{comm}}_{AB}]\right)\right. \\
& \qquad \left. + (1-\lambda) D\left(\rho^{\text{comm}}_{AB} \,||\, \Lambda_1[\rho^{\text{comm}}_{AB}]\right)\right] \\
& \text{Tr}(\rho_{AB}\, CHSH) = S \\
& \rho_{AB} \succeq 0 \\
& \text{Tr}(\rho_{AB}) = 1
\end{aligned} \tag{C.16}$$

where, $\rho^{ij}_{AB} = \text{Tr}_{EA'_X}(\rho^{ij}_{ABEA'_X})$ and $\Lambda[\rho^{ij}_{AB}] = \text{Tr}_{EA_X}(\Lambda[\rho^{ij}_{ABEA'_X}])$

$\rho^{\text{comm}}_{AB} = \text{Tr}_{EA'_X}(\rho^{\text{comm}}_{ABEA'_X})$ and $\Lambda[\rho^{\text{comm}}_{AB}] = \text{Tr}_{EA_X}(\Lambda[\rho^{\text{comm}}_{ABEA'_X}])$

When Alice's observables commute with that of and Bob, the joint probabilities of the measurement outcomes can be factorized into local outcome probabilities. Using this fact, one can can similarly analyze the state $\rho^{\text{comm}}_{ABEA'_X}$, where $X \in \{0, 1\}$ represents the two measurement outcomes. Due to the classical nature of these probability distributions, we can can disregard the contribution from the part $\rho^{\text{comm}}_{ABEA'_X}$ in the violation of CHSH inequality, as it will lead to $\text{Tr}(\rho^{\text{comm}}_{AB}\, CHSH) \leq 2$. Now, from the constraint of the optimization problem in (B.24), we can establish the same for each state $\rho^{ij}_{ABEA'_X} \in \mathcal{P}$ as,

$$\begin{aligned}
& \text{Tr}(\rho_{AB}\, CHSH) = S \\
\Leftrightarrow \quad & \text{Tr}(\rho^{ij}_{AB}\, CHSH) = S_{ij} \quad \text{and} \quad \sum_{ij} \eta_{ij} S_{ij} = S
\end{aligned} \tag{C.17}$$

Having established the constraint for each state $\rho^{ij}_{ABEA'_X} \in \mathcal{P}$, we can define the optimization problem for finding the lower bound of the function $C^*(S)$ as,

$$\begin{aligned}
C^*(S) \geq \inf \quad & \sum_{ij} \eta_{ij} \left[\lambda D\left(\rho^{ij}_{AB} \,||\, \Lambda_0[\rho^{ij}_{AB}]\right) + (1-\lambda) D\left(\rho^{ij}_{AB} \,||\, \Lambda_1[\rho^{ij}_{AB}]\right)\right] \\
\text{s.t.} \quad & \text{Tr}((\rho^{ij}_{AB}\, CHSH) = S_{ij} \\
& \rho^{ij}_{AB} \succeq 0 \\
& \text{Tr}(\rho^{ij}_{AB}) = 1 \\
& \sum_{ij} \eta_{ij} \leq 1 \\
& \sum_{ij} \eta_{ij}\, S_{ij} = S
\end{aligned} \tag{C.18}$$

The state $\rho_{ABEA'_X}$ is first passed through the channel $\Lambda_{\mathcal{XY}}$, which essentially decomposes the state into blocks along the principal diagonal as described in (C.16), before passing it through the pinching channel. Therefore, our objective function in (C.18) is essentially lower bounded by the minimum values in each such block. Thus,

$$\left[\lambda D\left(\rho^{ij}_{AB} \,||\, \Lambda_0[\rho^{ij}_{AB}]\right) + (1-\lambda) D\left(\rho^{ij}_{AB} \,||\, \Lambda_1[\rho^{ij}_{AB}]\right)\right] \geq C^*_{\mathrm{C}^{4\times 4}}(S_{ij}) \tag{C.19}$$

where,

$$\begin{aligned} C^*_{\mathbb{C}^{4\times4}}(S_{ij}) = \inf \quad & \left[\lambda D\left(\rho^{ij}_{AB}\,||\,\Lambda_0[\rho^{ij}_{AB}]\right) + (1-\lambda) D\left(\rho^{ij}_{AB}\,||\,\Lambda_1[\rho^{ij}_{AB}]\right)\right] \\ \text{s.t.} \quad & \mathrm{Tr}(\rho^{ij}_{AB}\; CHSH) = S_{ij} \\ & \rho^{ij}_{AB} \succeq 0 \\ & \mathrm{Tr}(\rho^{ij}_{AB}) = 1 \end{aligned} \tag{C.20}$$

Thus using (C.18),(C.19) and (C.20), one can lower bound the function $C^*(S)$ as,

$$\begin{aligned} C^*(S) \geq \sum_{ij} \eta_{ij} \quad \inf \quad & \left(C^*_{\mathbb{C}^{4\times4}}(S_{ij})\right) \\ \text{s.t.} \quad & \mathrm{Tr}(\rho^{ij}_{AB}\; CHSH) = S_{ij} \\ & \rho^{ij}_{AB} \succeq 0 \\ & \mathrm{Tr}(\rho^{ij}_{AB}) = 1 \\ & \sum_{ij} \eta_{ij} \leq 1 \\ & \sum_{ij} \eta_{ij} S_{ij} = S \end{aligned} \tag{C.21}$$

Now (C.21) is independent of the angles ϕ^i_A and ϕ^j_B as each states obtained from (C.14) are already optimized in (C.20). Therefore, we can reduce the optimization of $C^*(S)$ to optimizing only over to η_{ij} as,

$$\boxed{\begin{aligned} C^*(S) \geq & \int_{S'=2}^{2\sqrt{2}} C^*_{\mathbb{C}^{4\times4}}(S') \cdot \eta(dS') \\ \geq & \int_{S'=2}^{2\sqrt{2}} \eta(dS') \cdot C^*_{\mathbb{C}^{4\times4}}(S') \\ s.t \quad & \eta([2, 2\sqrt{2}]) \leq 1 \\ & \eta \geq 0 \\ & \int_{S'=2}^{2\sqrt{2}} \eta(dS') S' = S \end{aligned}} \tag{C.22}$$

Here, essentially one single block is considered and integrated over positive sub-normalized weights η in the interval $S' = (2, 2\sqrt{2}]$ with the conditions,

$$C^*_{\mathbb{C}^{4\times4}}(S_{ij}) = 0 \;\; \forall S_{ij} \leq 0 \tag{C.23}$$

2. Reformulating the CHSH operator in explicit matrix form

Having established the optimization problem in terms of two-qubit space $\mathbb{C}_{4\times4}$ of Alice and Bob, we now find an explicit matrix representation of the CHSH operator in (F.13) using (F.14),

$$\begin{aligned} CHSH = & \; C^{O_1^{\mathcal{X}}} \otimes C^{O_2^{\mathcal{Y}}} - C^{O_0^{\mathcal{X}}} \otimes C^{O_2^{\mathcal{Y}}} - C^{O_0^{\mathcal{X}}} \otimes C^{O_3^{\mathcal{Y}}} - C^{O_1^{\mathcal{X}}} \otimes C^{O_3^{\mathcal{Y}}} \\ = & \left[(P^{0|1}_{\mathcal{X}} \otimes P^{0|2}_{\mathcal{Y}} + P^{1|1}_{\mathcal{X}} \otimes P^{1|2}_{\mathcal{Y}}) - (P^{0|1}_{\mathcal{X}} \otimes P^{1|2}_{\mathcal{Y}} + P^{1|1}_{\mathcal{X}} \otimes P^{0|2}_{\mathcal{Y}}) \right] \\ - & \left[(P^{0|0}_{\mathcal{X}} \otimes P^{0|2}_{\mathcal{Y}} + P^{1|0}_{\mathcal{X}} \otimes P^{1|2}_{\mathcal{Y}}) - (P^{0|0}_{\mathcal{X}} \otimes P^{1|2}_{\mathcal{Y}} + P^{1|0}_{\mathcal{X}} \otimes P^{0|2}_{\mathcal{Y}}) \right] \\ - & \left[(P^{0|0}_{\mathcal{X}} \otimes P^{0|3}_{\mathcal{Y}} + P^{1|0}_{\mathcal{X}} \otimes P^{1|3}_{\mathcal{Y}}) - (P^{0|0}_{\mathcal{X}} \otimes P^{1|3}_{\mathcal{Y}} + P^{1|0}_{\mathcal{X}} \otimes P^{0|3}_{\mathcal{Y}}) \right] \\ - & \left[(P^{0|1}_{\mathcal{X}} \otimes P^{0|3}_{\mathcal{Y}} + P^{1|1}_{\mathcal{X}} \otimes P^{1|3}_{\mathcal{Y}}) - (P^{0|1}_{\mathcal{X}} \otimes P^{1|3}_{\mathcal{Y}} + P^{1|1}_{\mathcal{X}} \otimes P^{0|3}_{\mathcal{Y}}) \right] \end{aligned} \tag{C.24}$$

Now using the explicit matrix form of the projectors in (C.2) and (C.3), the CHSH operator can further be decomposed as,

$$\begin{aligned} CHSH = &\Big[\Big(Q(\phi_A^i)\otimes Q(0)\Big) + \Big((\mathbb{1}-Q(\phi_A^i))\otimes(\mathbb{1}-Q(0))\Big) - \Big(Q(\phi_A^i)\otimes(\mathbb{1}-Q(0))\Big) - \Big((\mathbb{1}-Q(\phi_A^i))\otimes Q(0)\Big)\Big] \\ &-[(Q(0)\otimes Q(0)) + (\mathbb{1}-Q(0))\otimes(\mathbb{1}-Q(0))) - (Q(0)\otimes(\mathbb{1}-Q(0))) - (\mathbb{1}-Q(0))\otimes Q(0))] \\ &-\Big[\Big(Q(0)\otimes Q(\phi_B^j)\Big) + \Big((\mathbb{1}-Q(0))\otimes(\mathbb{1}-Q(\phi_B^j))\Big) - \Big(Q(0)\otimes(\mathbb{1}-Q(\phi_B^j))\Big) - \Big((\mathbb{1}-Q(0))\otimes Q(\phi_B^j)\Big)\Big] \\ &-\Big[\Big(Q(\phi_A^i)\otimes Q(\phi_B^j)\Big) + \Big((\mathbb{1}-Q(\phi_A^i))\otimes(\mathbb{1}-Q(\phi_B^j))\Big) - \Big(Q(\phi_A^i)\otimes(\mathbb{1}-Q(\phi_B^j))\Big) - \Big((\mathbb{1}-Q(\phi_A^i))\otimes Q(\phi_B^j)\Big)\Big] \end{aligned} \tag{C.25}$$

Solving the tensor products in each term above, one can explicitly write the CHSH operator as,

$$CHSH = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{C.26}$$

where

$$A = \begin{pmatrix} \cos(\phi_A^i) - 1 - \cos(\phi_B^j) - \cos(\phi_A^i)\cos(\phi_B^j) & -\sin(\phi_B^j) - \cos(\phi_A^i)\sin(\phi_B^j) \\ -\sin(\phi_B^j) - \cos(\phi_A^i)\sin(\phi_B^j) & -\cos(\phi_A^i) + 1 + \cos(\phi_B^j) + \cos(\phi_A^i)\cos(\phi_B^j) \end{pmatrix} \tag{C.27}$$

$$B = \begin{pmatrix} \sin(\phi_A^i) - \sin(\phi_A^i)\cos(\phi_B^j) & -\sin(\phi_A^i)\sin(\phi_B^j) \\ -\sin(\phi_A^i)\sin(\phi_B^j) & \sin(\phi_A^i)\cos(\phi_B^j) - \sin(\phi_A^i) \end{pmatrix} \tag{C.28}$$

$$C = \begin{pmatrix} \sin(\phi_A^i) - \sin(\phi_A^i)\cos(\phi_B^j) & -\sin(\phi_A^i)\sin(\phi_B^j) \\ -\sin(\phi_A^i)\sin(\phi_B^j) & \sin(\phi_A^i)\cos(\phi_B^j) - \sin(\phi_A^i) \end{pmatrix} \tag{C.29}$$

$$D = \begin{pmatrix} -\cos(\phi_A^i) + 1 + \cos(\phi_B^j) + \cos(\phi_A^i)\cos(\phi_B^j) & \sin(\phi_B^j) + \cos(\phi_A^i)\sin(\phi_B^j) \\ \sin(\phi_B^j) + \cos(\phi_A^i)\sin(\phi_B^j) & \cos(\phi_A^i) - 1 - \cos(\phi_B^j) - \cos(\phi_A^i)\cos(\phi_B^j) \end{pmatrix} \tag{C.30}$$

Here individual operators A, B, C and D are Hermitian and $B = C^*$ implies that CHSH operator in (C.26) is also Hermitian. Now from *Lemma 1*, the mapping $\phi_N^m \mapsto 1 \pm \cos(\phi_N^m/2)$ 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 in the subinterval of $[0, \pi/2]$. In the first quadrant, both *sin* and *cosine* functions are positive; therefore, restricting the arguments of these functions to the first quadrant ensures that the CHSH operator in (C.26) does not change. Moreover, since both *sin* and *cosine* functions have unique values in the first quadrant, the optimization problem in Eq. (C.20) for each block can be reformulated with the explicit matrix representation of CHSH operator in (C.26) as,

$$\boxed{\begin{aligned} C^*_{\mathbb{C}^{4\times4}}(S_{ij}) = \inf \quad & \Big[\lambda D\Big(\rho_{AB}^{ij}\,||\,\Lambda_0[\rho_{AB}^{ij}]\Big) + (1-\lambda)D\Big(\rho_{AB}^{ij}\,||\,\Lambda_1[\rho_{AB}^{ij}]\Big)\Big] \\ \text{s.t.} \quad & \mathrm{Tr}\left(\rho_{AB}^{ij}\cdot\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right) = S_{ij} \\ & \rho_{AB}^{ij} \succeq 0 \\ & \mathrm{Tr}(\rho_{AB}^{ij}) = 1 \\ & \phi_A^i, \phi_B^j \in [0, \pi/2] \end{aligned}} \tag{C.31}$$

Appendix D: Formulation of the objective function in terms of trace norm using 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 block and expressed the CHSH operator in its explicit matrix form (C.26), we now focus on finding a lower bound on the relative entropies described in Eq. (C.31) through a modified form of Pinsker's inequality[SGP+21]. The modified version of Pinsker's inequality, as shown in the original work, is stated in the following theorem,

Theorem 1: Let Q be a projector, not necessarily of rank 1, defines the action of a pinching channel Λ acting on a state ρ as $\Lambda[\rho] = Q\rho Q + (\mathbb{1} - Q)\rho(\mathbb{1} - Q)$. Then the quantum relative entropy between ρ and $\Lambda[\rho]$ is lower bounded as,

$$D(\rho||\Lambda[\rho]) \geq \log_2(2) - h\left(\frac{1 - ||\rho - \Lambda[\rho]||_1}{2}\right) \tag{D.1}$$

where $h(p) = -\sum_{i=0}^{1} p_i \log_2(p_i)$ is the binary Shannon entropy.

Proof : The pinching channel Λ is defined as,

$$\begin{aligned}\Lambda[\rho] &= Q\rho Q + (\mathbb{1} - Q)\rho(\mathbb{1} - Q) \\ &= Q\rho Q + (\rho - Q\rho)(\mathbb{1} - Q) \\ &= Q\rho Q + (\rho - \rho Q - Q\rho + Q\rho Q) \\ &= 2Q\rho Q + \rho - \{Q, \rho\}\end{aligned} \tag{D.2}$$

Now applying *Theorem 1* in Eq. (C.31), we can have,

$$\begin{aligned}C^*_{\mathbb{C}^{4\times4}}(S_{ij}) = \inf \quad & \left[\lambda D\left(\rho^{ij}_{AB}\,||\,\Lambda_0[\rho^{ij}_{AB}]\right) + (1-\lambda) D\left(\rho^{ij}_{AB}\,||\,\Lambda_1[\rho^{ij}_{AB}]\right)\right] \\ \geq & \left[\lambda\left(\log_2(2) - h\left(\frac{1 - ||\rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]||_1}{2}\right)\right)\right. \\ & \left. +(1-\lambda)\left(\log_2(2) - h\left(\frac{1 - ||\rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]||_1}{2}\right)\right)\right] \\ \geq & \left[\lambda \log_2(2) - \lambda h\left(\frac{1 - ||\rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]||_1}{2}\right)\right. \\ & \left. +(1-\lambda)\log_2(2) - (1-\lambda) h\left(\frac{1 - ||\rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]||_1}{2}\right)\right] \\ \geq & \left[\log_2(2) - \lambda h\left(\frac{1 - ||\rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]||_1}{2}\right)\right. \\ & \left. -(1-\lambda) h\left(\frac{1 - ||\rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]||_1}{2}\right)\right] \\ \text{s.t.} \quad \operatorname{Tr} & \left(\rho^{ij}_{AB} \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix}\right) = S_{ij}\end{aligned} \tag{D.3}$$

Let us now define two arguments a and b as,

$$\begin{aligned}a &= \frac{1 - \left\|\rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]\right\|_1}{2}, \\ b &= \frac{1 - \left\|\rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]\right\|_1}{2}\end{aligned} \tag{D.4}$$

As $\|\cdot\|_1 \leq 1$ is valid for normalized quantum states, one can have $a, b \in [0, 0.5]$. This ensures that the entropy functions $h(a)$ 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 [Ras15]. Thus from Jensen's inequality [BV04], for $\lambda \in [0, 1]$ and $a, b \in [0, 1]$, we have:

$$\begin{aligned}\lambda h(a) + (1-\lambda) h(b) &\leq h(\lambda a + (1-\lambda) b) \\ -[\lambda h(a) + (1-\lambda) h(b)] &\geq -h(\lambda a + (1-\lambda) b)\end{aligned} \tag{D.5}$$

Finally, using Eqns. (D.3), (D.4), and (D.5), we get the lower the bound on $C^*_{\mathbb{C}^{4\times4}}(S_{ij})$ as,

$$\begin{aligned} C^*_{\mathbb{C}^{4\times4}}(S_{ij}) &\geq \log_2(2) - h(\lambda a + (1-\lambda)b) \\ &\geq \log_2(2) - h\left(\frac{1}{2} - \frac{1}{2}n\right) \end{aligned} \tag{D.6}$$

where

$$n = \left(\lambda \left\|\rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]\right\|_1 + (1-\lambda)\left\|\rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]\right\|_1\right) \tag{D.7}$$

and it represents a convex combination of trace norms. Now, one may assume that there exists a function $n^*(S_{ij}) \geq n$, which gives an upper bound on the given convex combination of trace norms (D.7). Hence, we can get the upper bound by solving the below optimization problem,

$$\boxed{\begin{aligned} n^*(S_{ij}) = \inf \quad & \left[\left(\lambda \left\|\rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]\right\|_1 + (1-\lambda)\left\|\rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]\right\|_1\right)\right] \\ \text{s.t.} \quad & \text{Tr}\left(\rho^{ij}_{AB} \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix}\right) = S_{ij} \\ & \rho^{ij}_{AB} \succeq 0 \\ & \text{Tr}(\rho^{ij}_{AB}) = 1 \\ & \phi^i_A, \phi^j_B \in [0, \pi/2] \end{aligned}} \tag{D.8}$$

Appendix E: Semi definite programming formulation of the objective function for fixed ϕ^i_A and ϕ^j_B

A semidefinite program (SDP) is an optimisation problem of a linear function defined over a positive semidefinite variable, subjected to affine constraints as [Sik12],

$$\begin{aligned} \alpha = \text{maximize} \quad & \langle A, X\rangle \\ \text{s.t.} \quad & \Phi(X) = B \\ & X \in \text{Pos}(\mathcal{X}) \end{aligned} \tag{E.1}$$

where,

$$\begin{aligned} & X \in \mathbb{C}^{\Sigma}, \quad Y \in \mathbb{C}^{\Omega} \\ & A \in \text{Herm}(X), \quad B \in \text{Herm}(Y) \\ & \Phi : \text{Herm}(X) \rightarrow \text{Herm}(Y) \\ & (A, B, \Phi) \text{ problem's data} \\ & \langle A, X\rangle \text{ is the objective function} \\ & \Phi(X) = B, \text{ and } X \in \text{Pos}(\mathcal{X}) \\ & \alpha \text{ is the optimal value} \end{aligned} \tag{E.2}$$

The trace norm of a quadratic matrix M can be represented in the form of a trace of two additional matrices P and Q [BV04] as,

$$\begin{aligned} ||M||_1 = inf \quad & \frac{1}{2}\text{Tr}(P+Q) \\ s.t. \quad & \begin{pmatrix} P & M \\ M^* & Q \end{pmatrix} \succeq 0 \end{aligned} \tag{E.3}$$

Now the terms $\rho_{AB}^{ij} - \Lambda_0[\rho_{AB}^{ij}]$ and $\rho_{AB}^{ij} - \Lambda_1[\rho_{AB}^{ij}]$ in Eq. (D.8) can be decomposed using Eq. (D.2) and Eq. (B.2) as,

$$\begin{aligned}
&\rho_{AB}^{ij} - \Lambda_0[\rho_{AB}^{ij}] \\
&\quad = \rho_{AB}^{ij} - \left(2Q(0)\otimes\mathbb{I}\,\rho_{AB}^{ij}\,Q(0)\otimes\mathbb{I} + \rho_{AB}^{ij} - \{Q(0)\otimes\mathbb{I}, \rho_{AB}^{ij}\}\right) \\
&\quad = \rho_{AB}^{ij} - \left(2Q(0)\otimes\mathbb{I}\,\rho_{AB}^{ij}\,Q(0)\otimes\mathbb{I} + \rho_{AB}^{ij} - \left(Q(0)\otimes\mathbb{I}\,\rho_{AB}^{ij} - \rho_{AB}^{ij}\,Q(0)\otimes\mathbb{I}\right)\right) \\
&\quad = -2Q(0)\otimes\mathbb{I}\,\rho_{AB}^{ij}\,Q(0)\otimes\mathbb{I} + \left(Q(0)\otimes\mathbb{I}\,\rho_{AB}^{ij} + \rho_{AB}^{ij}\,Q(0)\otimes\mathbb{I}\right) \\
&\quad = \left(Q(0)\otimes\mathbb{I}\,\rho_{AB}^{ij} + \rho_{AB}^{ij}\,Q(0)\otimes\mathbb{I}\right) - 2Q(0)\otimes\mathbb{I}\,\rho_{AB}^{ij}\,Q(0)\otimes\mathbb{I}
\end{aligned} \tag{E.4}$$

In a similar fashion, we can also expand the other term in Eq. (D.8) as,

$$\rho_{AB}^{ij} - \Lambda_1[\rho_{AB}^{ij}] = \left(Q(\phi_A^i)\otimes\mathbb{I}\,\rho_{AB}^{ij} + \rho_{AB}^{ij}\,Q(\phi_A^i)\otimes\mathbb{I}\right) - 2Q(\phi_A^i)\otimes\mathbb{I}\,\rho_{AB}^{ij}\,Q(\phi_A^i)\otimes\mathbb{I} \tag{E.5}$$

The SDP formulation of the objective function can now be formally done using the results of (D.8),(E.1),(E.2),(E.3),(E.4), and (E.5) as,

$$\boxed{\begin{aligned}
n^*(S_{ij}) = -\text{maximize}\quad & \frac{\lambda}{2}\langle (P_0 + Q_0), X_0\rangle + \frac{(1-\lambda)}{2}\langle (P_1 + Q_1), X_1\rangle \\
\text{s.t.}\quad & \mathrm{Tr}\left(\rho_{AB}^{ij}\begin{bmatrix} A & B \\ C & D\end{bmatrix}\right) = S_{ij}, \\
& \begin{pmatrix} P_0 & M_0 \\ M_0^* & Q_0 \end{pmatrix} \succeq 0, \quad \begin{pmatrix} P_1 & M_1 \\ M_1^* & Q_1 \end{pmatrix} \succeq 0, \\
& Q(\phi_A^i) = \begin{bmatrix} q_{11} & q_{12} \\ q_{12}^* & q_{22}\end{bmatrix}, \quad q_{11} + q_{22} = 1, \quad q_{12}^2 \leq q_{11}q_{22}, \\
& \rho_{AB}^{ij} \succeq 0, \quad \mathrm{Tr}(\rho_{AB}^{ij}) = 1
\end{aligned}} \tag{E.6}$$

where

$$\begin{aligned}
M_0 = M_0^* &= \left(Q(0)\otimes\mathbb{I}\,\rho_{AB}^{ij} + \rho_{AB}^{ij}\,Q(0)\otimes\mathbb{I}\right) - 2Q(0)\otimes\mathbb{I}\,\rho_{AB}^{ij}\,Q(0)\otimes\mathbb{I} \\
M_1 = M_1^* &= \left(Q(\phi_A^i)\otimes\mathbb{I}\,\rho_{AB}^{ij} + \rho_{AB}^{ij}\,Q(\phi_A^i)\otimes\mathbb{I}\right) - 2Q(\phi_A^i)\otimes\mathbb{I}\,\rho_{AB}^{ij}\,Q(\phi_A^i)\otimes\mathbb{I}
\end{aligned} \tag{E.7}$$

The projector $Q(\phi_A^i)$ follows the positive semi-definite condition (C.1).

Appendix F: Optimization of the angles ϕ_A^i and ϕ_B^j using ϵ - net approximation

Alice's and Bob's angles ϕ_A^i and ϕ_B^j appear as constraints in the optimisation problem in Eq. (D.8). The objective function has been formulated in standard SDP form in Eq. (E.6) taking the assumption of fixed angles. In this section, we optimize the angles ϕ_A^i and ϕ_B^j using ϵ - net approach. The work of [SGP+21] showed that the Alice's angle ϕ_A^i can be optimised using this approach. Here, we show that we can use the same approach to optimise Bob's angle ϕ_B^j, thus eliminating the polytope optimisation in the security analysis as done [SGP+21].

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:

$$|\phi_A^i - \phi_{A_k}^i| \leq \epsilon_0 \quad |\phi_B^j - \phi_{B_l}^j| \leq \epsilon_0 \tag{F.1}$$

where S_A and S_B are the number of segments in the interval I for Alice and Bob, respectively. Each segment is centralized around the angles $\phi_{A_k}^i$ and $\phi_{B_l}^j$ for k^{th} and l^{th} segments, respectively. The values of both the angles ϕ_A^i and ϕ_B^j

are needed to solve any instance of the SDP and hence, it is solved for each discrete points $\phi^i_{A_k}$ and $\phi^j_{B_l}$. An error term known as pessimistic error [SGP+21] is being subtracted from each SDP's result. The error term accounts for the variation of the optimal value. The pessimistic error term Δ is introduced in the original work[SGP+21] as a function of ϵ_0 and the angle ϕ^i_A orϕ^j_B. Iteratively, the segment that gives the smallest value of the objective function in (D.8) is chosen until a point is reached that corresponds to the global minima. Here, we try to derive a closed form of the pessimistic error.

Let us consider that $f(\phi)$ is a solution of the optimization problem in (D.8) for a given value of the angle ϕ. Now, the discrete angles $\phi^i_{A_k}$ and $\phi^j_{B_l}$ corresponding to Alice's and Bob's measurement settings are being separated by a distance of $2\epsilon_0$. Thus each angle represents a segment of the same width as,

$$I_A = \left[\phi^i_{A_k} - \epsilon_0, \phi^i_{A_k} + \epsilon_0\right] \quad \text{and} \quad I_B = \left[\phi^j_{B_l} - \epsilon_0, \phi^j_{B_l} + \epsilon_0\right] \tag{F.2}$$

Hence, the pessimistic error terms $\Delta(\epsilon_0, \phi^i_A)$ and $\Delta(\epsilon_0, \phi^j_B)$ provide an upper bound on the absolute difference between the values of the function at any point within these segments and the optimal one, i.e.,

$$\begin{aligned} &|f(\phi^i_A) - f(\phi^i_{A_k})| \leq \Delta\left(\epsilon_0, \phi^i_A\right) \\ &|f(\phi^i_B) - f(\phi^j_{B_l})| \leq \Delta\left(\epsilon_0, \phi^j_B\right), \qquad \forall \phi^i_A \in I_A \text{ and } \forall \phi^j_B \in I_B \end{aligned} \tag{F.3}$$

The iterative process involves selecting the segment that yields the minimum value of the objective function in Eq. (D.8), continuing until the global minima is reached.

1. Lipschitz Continuity of $f(\phi)$

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

$$\|f(y) - f(x)\| \leq L\|y - x\| \tag{F.4}$$

for all $y \in S$ sufficiently near x.[TBB08]

Theorem 2: If $f(\phi)$ is the solution of a well-behaved optimisation problem, then it is Lipschitz continuous

Problem setup To prove our claim, we first formally define the optimisation problem. Let $f(\phi)$ be defined as the solution to the following optimization problem,

$$\begin{aligned} f(\phi) = &\min g(x, \phi) \\ &\text{s.t. } x \in \mathcal{X} \end{aligned} \tag{F.5}$$

where $\mathcal{X} \subseteq \mathbb{R}^n$ is a compact and convex set, $g(x, \phi)$ is a continuously differentiable function in both x and ϕ, and $g(x, \phi)$ is strongly convex in x with modulus $\mu > 0$, i.e., for all $x_1, x_2 \in \mathcal{X}$ and $\phi \in \Phi$,

$$g(x_1, \phi) \geq g(x_2, \phi) + \nabla_x g(x_2, \phi)^T (x_1 - x_2) + \frac{\mu}{2}\|x_1 - x_2\|^2 \tag{F.6}$$

and $\nabla_x g(x, \phi)$ is Lipschitz continuous in ϕ with constant L_ϕ, i.e.,

$$\|\nabla_x g(x, \phi_1) - \nabla_x g(x, \phi_2)\| \leq L_\phi \|\phi_1 - \phi_2\|. \tag{F.7}$$

The optimisation problem is well-behaved in the sense that the solution $f(\phi)$ exists and is unique for all $\phi \in \Phi$ and the optimality condition $\nabla_x g(f(\phi), \phi) = 0$ holds.

Proof of Lipschitz Continuity of $f(\phi)$ The aim of the proof is to show that $f(\phi)$ is Lipschitz continuous, that is there exists a constant $L_f > 0$ such that:

$$\|f(\phi_1) - f(\phi_2)\| \leq L_f \|\phi_1 - \phi_2\|. \tag{F.8}$$

Let us consider that for any two $\phi_1, \phi_2 \in \Phi$ there are two solutions $f(\phi_1)$ and $f(\phi_2)$. Therefore, we have,

$$\nabla_x g(f(\phi_1), \phi_1) = 0 \quad \text{and} \quad \nabla_x g(f(\phi_2), \phi_2) = 0. \tag{F.9}$$

Now since $g(f(\phi), \phi)$ is strongly convex in $f(\phi)$ with modulus $\mu > 0$, one can can have

$$g(f(\phi_1), \phi_1) \geq g(f(\phi_2), \phi_2) + \nabla_x g(f(\phi_2), \phi_2)^T (f(\phi_1) - f(\phi_2)) + \frac{\mu}{2} \|f(\phi_1) - f(\phi_2)\|^2 \tag{F.10}$$

and thus from Eq. F.9 we get,

$$g(f(\phi_1), \phi_2) \geq g(f(\phi_2), \phi_2) + \frac{\mu}{2} \|f(\phi_1) - f(\phi_2)\|^2. \tag{F.11}$$

Now, as the the lower bound set up by the quadratic norm is symmetric in terms of the angles, we have,

$$g(f(\phi_2), \phi_1) \geq g(f(\phi_1), \phi_1) + \frac{\mu}{2} \|f(\phi_2) - f(\phi_1)\|^2. \tag{F.12}$$

Adding the two inequalities F.11 and F.12 and rearranging the terms we get,

$$\mu \|f(\phi_1) - f(\phi_2)\|^2 \leq g(f(\phi_1), \phi_2) - g(f(\phi_1), \phi_1) + g(f(\phi_2), \phi_1) - g(f(\phi_2), \phi_2). \tag{F.13}$$

Using the above equation and considering the fact that $\nabla_x g(x, \phi)$ in ϕ is Lipschitz continuous, one can have the following two inequalities

$$\begin{aligned} |g(f(\phi_1), \phi_2) - g(f(\phi_1), \phi_1)| &\leq L_\phi \|f(\phi_1)\| \|\phi_2 - \phi_1\|, \\ |g(f(\phi_2), \phi_1) - g(f(\phi_2), \phi_2)| &\leq L_\phi \|f(\phi_2)\| \|\phi_1 - \phi_2\|. \end{aligned} \tag{F.14}$$

Since the feasible solution set $\mathcal{X}$ is compact, F.13 and F.14 $\|f(\phi)\|$ jointly yield,

$$\mu \|f(\phi_1) - f(\phi_2)\|^2 \leq 2 L_\phi M \|\phi_1 - \phi_2\|. \tag{F.15}$$

where M is a constant. Hence, simple algebra shows,

$$\|f(\phi_1) - f(\phi_2)\| \leq \sqrt{\frac{2 L_\phi M}{\mu}} \|\phi_1 - \phi_2\| \tag{F.16}$$

Thus, $f(\phi)$ is Lipschitz continuous with constant,

$$L_f = \sqrt{\frac{2 L_\phi M}{\mu}} \tag{F.17}$$

□

2. Modification of the optimization problem

In order to make the objective function in (D.8) continuously differentiable with respect to both the density operator ρ_{AB}^{ij} and the parameters ϕ_A^i, we introduce the squared Frobenius norm. The squared Frobenius norm is smooth and continuously differentiable at every point [PP08]. For any operator X, this norm is defined as,

$$\begin{aligned} \|X\|_F &= \sqrt{\mathrm{Tr}(X^* X)} \\ \|X\|_F^2 &= \mathrm{Tr}(X^* X) \end{aligned} \tag{F.18}$$

where X^* is the Hermitian conjugate of X. Replacing the l_1 norm in the original objective function with the squared Frobenius norm and adding a convex regularization term $\frac{\mu}{2} ||\rho_{AB}^{ij}||_F^2$, we define the modified optimisation problem as,

$$\boxed{\begin{aligned} n^*(S_{ij}) = \inf \quad & \lambda \left\|\rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]\right\|^2_F + (1-\lambda)\left\|\rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]\right\|^2_F + \frac{\mu}{2}\|\rho^{ij}_{AB}\|^2_F \\ \text{s.t.} \quad & \mathrm{Tr}\left(\rho^{ij}_{AB}\, CHSH(\phi^i_A, \phi^j_B)\right) = S_{ij} \\ & \phi^i_A, \phi^j_B \in [0, \pi/2], \\ & \rho^{ij}_{AB} \succeq 0 \\ & \mathrm{Tr}(\rho^{ij}_{AB}) = 1 \end{aligned}} \tag{F.19}$$

This modified version of the optimisation problem can be reformulated into an SDP using Schur complements. Let us first decompose each Frobenius norm term into the corresponding inner product form as,

$$\begin{aligned} n^*(S_{ij}) = -\text{maximize} \quad & \lambda \mathrm{Tr}({\rho_0}^\dagger \rho_0) + (1-\lambda)\mathrm{Tr}({\rho_1}^\dagger \rho_1) + \frac{\mu}{2}\mathrm{Tr}({\rho_2}^\dagger \rho_2) \\ \text{s.t.} \quad & \mathrm{Tr}\left(\rho^{ij}_{AB}\, CHSH(\phi^i_A, \phi^j_B)\right) = S_{ij} \\ & \phi^i_A, \phi^j_B \in [0, \pi/2], \\ & \rho^{ij}_{AB} \succeq 0 \\ & \mathrm{Tr}(\rho^{ij}_{AB}) = 1 \end{aligned} \tag{F.20}$$

where $\rho_0 = \rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]$, $\rho_1 = \rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]$ and $\rho_2 = \rho^{ij}_{AB}$
Here, each inner product term (ρ_k, ρ_k) for $k \in \{0, 1, 2\}$ in (F.20) is quadratic in ρ_k. The standard SDP formulation requires that the objective function is linear in its decision variables and the constraints are in the form of linear matrix inequalities. Let $t_k \geq (\rho_k, \rho_k)$, then using Schur's complement[BTN01] one can have,

$$\begin{pmatrix} t_k & \mathrm{vec}(\rho_k)^\dagger \\ \mathrm{vec}(\rho_k) & 1 \end{pmatrix} \succeq 0 \tag{F.21}$$

Since $CHSH$ is fixed for given values of ϕ^i_A and ϕ^j_B, the standard SDP formulation would be as,

$$\boxed{\begin{aligned} n^*(S_{ij}) = -\text{maximize} \quad & \lambda t_0 + (1-\lambda) t_1 + \frac{\mu}{2} t_2 \\ \text{s.t.} \quad & \begin{pmatrix} t_0 & \mathrm{vec}(\rho_0)^\dagger \\ \mathrm{vec}(\rho_0) & 1 \end{pmatrix} \succeq 0 \\ & \begin{pmatrix} t_1 & \mathrm{vec}(\rho_1)^\dagger \\ \mathrm{vec}(\rho_1) & 1 \end{pmatrix} \succeq 0 \\ & \begin{pmatrix} t_2 & \mathrm{vec}(\rho_2)^\dagger \\ \mathrm{vec}(\rho_2) & 1 \end{pmatrix} \succeq 0 \\ & \mathrm{Tr}\left(\rho^{ij}_{AB}\, CHSH(\phi^i_A, \phi^j_B)\right) = S_{ij}, \\ & \phi^i_A, \phi^j_B \in [0, \pi/2], \\ & \rho^{ij}_{AB} \succeq 0 \\ & \mathrm{Tr}(\rho^{ij}_{AB}) = 1 \end{aligned}} \tag{F.22}$$

In the below, we formally prove that the modified optimisation problem satisfies the conditions for being Lipschitz continuous.

Lemma 4: The modified objective function is continuously differentiable with respect to the density operator ρ^{ij}_{AB} and the parameter ϕ^i_A.
Proof: The squared Frobenius norm $\|X\|^2_F = \mathrm{Tr}(X^\dagger X)$ is smooth and infinitely differentiable. As the channels Λ_0 and Λ_1 are linear maps, the action $\Lambda_x[\rho]$ is linear in ρ. The regularization term $\frac{\mu}{2}\|\rho\|^2_F$ is quadratic and smooth. Hence,

the objective function in the modified optimization problem is continuously differentiable with respect to both the density operator ρ^{ij}_{AB} and the parameters ϕ^i_A. □

Lemma 5: The modified objective function is strongly convex in ρ^{ij}_{AB}

Proof: The objective function in Eq.(F.19) includes the term $\frac{\mu}{2}\|\rho^{ij}_{AB}|^2_F$, which is μ-strongly convex. For any ρ^{ij}_{AB} and ρ^{kl}_{AB}:

$$n\left(\rho^{ij}_{AB},(\phi^i_A)\right) \geq n\left(\rho^{kl}_{AB},(\phi^i_A)\right) + \nabla_\rho n\left(\rho^{kl}_{AB},(\phi^i_A)\right)^T(\rho^{ij}_{AB}-\rho^{kl}_{AB}) + \frac{\mu}{2}\|(\rho^{ij}_{AB}-\rho^{kl}_{AB})\|^2_F. \tag{F.23}$$

To rigorously establish the definition of strong convexity for the function, let us first calculate the gradient of the objective function $\nabla_\rho n\left(\rho^{kl}_{AB},(\phi^i_A)\right)$. The Fréchet derivative of $||A||^2_F$ is $2A$[Wat18]. Thus,

$$\begin{aligned}
&\nabla_\rho n\left(\rho^{kl}_{AB},(\phi^i_A)\right)\\
&= \nabla_\rho\left(\lambda\left\|\rho^{kl}_{AB}-\Lambda_0[\rho^{kl}_{AB}]\right\|^2_F + (1-\lambda)\left\|\rho^{kl}_{AB}-\Lambda_1[\rho^{kl}_{AB}]\right\|^2_F + \frac{\mu}{2}\|\rho^{kl}_{AB}\|^2_F\right)\\
&= \nabla_\rho\left(\lambda\left\|\rho^{kl}_{AB}-\Lambda_0[\rho^{kl}_{AB}]\right\|^2_F\right) + \nabla_\rho\left((1-\lambda)\left\|\rho^{kl}_{AB}-\Lambda_1[\rho^{kl}_{AB}]\right\|^2_F\right)\\
&\quad + \nabla_\rho\left(\frac{\mu}{2}\|\rho^{kl}_{AB}\|^2_F\right)\\
&= \lambda 2\left(\rho^{kl}_{AB}-\Lambda_0[\rho^{kl}_{AB}]\right)\cdot\nabla_\rho\left(\rho^{kl}_{AB}-\Lambda_0[\rho^{kl}_{AB}]\right)\\
&\quad + (1-\lambda)2\left(\rho^{kl}_{AB}-\Lambda_1[\rho^{kl}_{AB}]\right)\cdot\nabla_\rho\left(\rho^{kl}_{AB}-\Lambda_1[\rho^{kl}_{AB}]\right) + \mu\rho\\
&= \lambda 2\left(\rho^{kl}_{AB}-\Lambda_0[\rho^{kl}_{AB}]\right)\cdot(\mathbb{I}-\Lambda_0^*) + (1-\lambda)2\left(\rho^{kl}_{AB}-\Lambda_1[\rho^{kl}_{AB}]\right)\cdot(\mathbb{I}-\Lambda_1^*) + \mu\rho\\
&= \lambda 2\left(\rho^{kl}_{AB}-\Lambda_0[\rho^{kl}_{AB}]-\Lambda_0^*\left(\rho^{kl}_{AB}-\Lambda_0[\rho^{kl}_{AB}]\right)\right)\\
&\quad + (1-\lambda)2\left(\rho^{kl}_{AB}-\Lambda_1[\rho^{kl}_{AB}]-\Lambda_1^*\left(\rho^{kl}_{AB}-\Lambda_1[\rho^{kl}_{AB}]\right)\right) + \mu\rho\\
&= \lambda 2\left(\rho^{kl}_{AB}-\Lambda_0[\rho^{kl}_{AB}]\right) + (1-\lambda)2\left(\rho^{kl}_{AB}-\Lambda_1[\rho^{kl}_{AB}]\right) + \mu\rho
\end{aligned} \tag{F.24}$$

Hence the objective function $n\left(\rho^{ij}_{AB},(\phi^i_A)\right)$ is strongly convex with modulus $\mu > 0$.

□

Lemma 6: The feasible set $\mathcal{D}$ of all density operator and the interval $I = [0, \frac{\pi}{2}]$ is compact and 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, $\mathrm{Tr}\left(\rho^{ij}_{AB}\, CHSH(\phi^i_A,\phi^j_B)\right) = S_{ij}$, $\phi^i_A, \phi^j_B \in [0,\pi/2]$, $\rho^{ij}_{AB} \succeq 0$, $\mathrm{Tr}(\rho^{ij}_{AB}) = 1$

a) Proof of convexity: For $\rho_1, \rho_2 \succeq 0$ and $\lambda \in [0,1]$,

$$\rho = \lambda\rho_1 + (1-\lambda)\rho_2 \succeq 0, \quad \mathrm{Tr}(\rho) = \lambda\mathrm{Tr}(\rho_1) + (1-\lambda)\mathrm{Tr}(\rho_2) = 1. \tag{F.25}$$

Thus the set of density operators $\rho \succeq 0$ with $\mathrm{Tr}(\rho) = 1$ is convex.[Wat18],[Wil13]

- Convexity of Angle Intervals: For $\phi_1, \phi_2 \in [0, \pi/2]$,

$$\lambda\phi_1 + (1-\lambda)\phi_2 \in [0,\pi/2] \tag{F.26}$$

  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 $\mathrm{Tr}(\rho_2 \cdot CHSH(\phi^i_A,\phi^j_B)) = S_{ij}$,

$$\mathrm{Tr}\left((\lambda\rho_1 + (1-\lambda)\rho_2)\cdot CHSH(\phi^i_A,\phi^j_B)\right) = \lambda S_{ij} + (1-\lambda)S_{ij} = S_{ij}. \tag{F.27}$$

Thus the constraint $\mathrm{Tr}(\rho \cdot CHSH(\phi_A^i, \phi_B^j)) = S_{ij}$ 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)[Wat18].

  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.[Wat18]

  The interval $[0, \pi/2]$ is bounded in $\mathbb{R}$.

□

Lemma 7: $\nabla_\rho n\left(\rho_{AB}^{ij}, (\phi_A^i)\right)$ is Lipschitz continuous in ϕ_A^i for $x \in \{0, 1\}$

Proof: From (F.24), one can have the gradient of $\nabla_\rho n\left(\rho_{AB}^{ij}, (\phi_A^i)\right)$. For a given segment ϕ_A^i parameterizes Λ_0, Λ_1 smoothly because for a particular segment its fixed, then $\nabla_\rho n\left(\rho_{AB}^{ij}, (\phi_A^i)\right)$ depends smoothly on ϕ_A^i. Since $\cos(\phi_A^i)$, $\sin(\phi_A^i)$, and their derivatives are bounded, $\nabla_\rho n\left(\rho_{AB}^{ij}, (\phi_A^i)\right)$ is Lipschitz continuous in ϕ_A^i as,

$$\|\nabla_\rho n\left(\rho_{AB}^{ij}, (\phi_A^i)\right) - \nabla_\rho n\left(\rho_{AB}^{kl}, (\phi_A^k)\right)\| \leq L_{\phi_A}\|\phi_A^i - \phi_A^k\|. \tag{F.28}$$

where ρ_{AB}^{ij} and ρ_{AB}^{kl} are being two arbitrary states and ϕ_A^i and ϕ_A^k are associated angles.
Hence $\nabla_\rho n\left(\rho_{AB}^{ij}, (\phi_A^i)\right)$ is lipschitz continuous in ϕ_A^i for $x \in \{0, 1\}$.

□

Lemma 8: Existence of Unique optima, $\nabla_\rho n\left(n^*(S_{ij}), (\phi_A^i)\right) = 0$
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:

$$\nabla_\rho n\left(n^*(S_{ij}), (\phi_A^i)\right) = 0 \tag{F.29}$$

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,

$$g(tx_a + (1-t)x_b) \leq tg(x_a) + (1-t)g(x_b) \tag{F.30}$$

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 9: 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 (F.30).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,

$$y - g(x_a) = \frac{g(x_b) - g(x_a)}{x_b - x_a}(x - x_a) \tag{F.31}$$

Now, let $x = tx_a + (1-t)x_b$. Then,

$$\begin{aligned} x - x_a &= tx_a + (1-t)x_b - x_a \\ &= (t-1)x_a + (1-t)x_b \\ &= (1-t)(x_b - x_a) \end{aligned} \tag{F.32}$$

Substituting this into the equation of the secant line, in (F.31)

$$\begin{aligned} y - g(x_a) &= \frac{g(x_b) - g(x_a)}{x_b - x_a}(1-t)(x_b - x_a) \\ y - g(x_a) &= (1-t)(g(x_b) - g(x_a)) \\ y &= g(x_a) + (1-t)g(x_b) - (1-t)g(x_a) \\ y &= g(x_a) - (1-t)g(x_a) + (1-t)g(x_b) \\ y &= (1-(1-t))g(x_a) + (1-t)g(x_b) \\ y &= tg(x_a) + (1-t)g(x_b) \end{aligned} \tag{F.33}$$

Thus from *Lemma 9*, the inequality in (F.30) 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 $(x_b, g(x_b))$. □

3. 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 optimisation problem in (F.19), $n^*(S_{ij})$, is thus established to be Lipschitz continuous. Now we 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 (F.22) gives an optimal value for the k^{th} and l^{th} segment centered around $\phi_{A_k}^i$ and $\phi_{B_l}^j$, respectively. As the solution of the SDP is based on fixed ϕ_A^i, ϕ_B^j, to find the optimal value of the objective function over the whole interval I, we need a slightly different approach from the one given in [SGP+21] because here we optimize both the angles using the ϵ-net approach. The process begins by solving the SDP for each segment of Alice (Bob), parameterised by discretised angles $\phi_{A_k}^i (\phi_{B_l}^j)$. Subsequently, the segment that produced the minimum 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. Then the optimal value with respect to the angles of 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. Let us consider that $f_A(\phi_A^i)$ and $f_B(\phi_B^j)$ are the solutions of the SDP in (F.22) or the optimization problem in (F.19) for the fixed angles ϕ_A^i and ϕ_B^j. Since it is proved that the solutions are Lipschitz continuous, we have the following relations for each segment where $\phi_{A_k}^i$ and $\phi_{B_l}^j$ are the centres of the segments k and l, respectively.

$$\begin{aligned} &|f_A(\phi_A^i) - f_A(\phi_{A_k}^i)| \leq L_A|\phi_A^i - \phi_{A_k}^i| \\ &|f_B(\phi_B^j) - f_B(\phi_{B_l}^j)| \leq L_B|\phi_B^j - \phi_{B_l}^j|, \qquad \forall \phi_A^i \in I_A \quad \text{and} \quad \forall \phi_B^j \in I_B \end{aligned} \tag{F.34}$$

The above inequalities depict that the absolute value of the difference between the solutions for two different angles is upper bounded by the absolute value of the difference between the respective angles times a constant. The supremum of the functions $f_A(\phi_A^i)$ and $f_B(\phi_B^j)$ are given as,

$$\begin{aligned} M_A^i &= sup_{\phi_A^i \in I_A} f_A(\phi_A^i) \\ M_B^j &= sup_{\phi_B^j \in I_B} f_B(\phi_B^j) \end{aligned} \tag{F.35}$$

Similarly, the infimum are given as,

$$\begin{aligned} m_A^i &= inf_{\phi_A^i \in I_A} f_A(\phi_A^i) \\ m_B^j &= inf_{\phi_B^j \in I_B} f_B(\phi_B^j) \end{aligned} \tag{F.36}$$

The Lipschitz continuity of the functions implies that they are continuous in the given interval. Therefore, the supremums in (F.35) are the corresponding maxima and the infimums in (F.36) are the corresponding minima. The supremum and infimum guarantee that no number smaller than $M_A^i(M_B^j)$ can serve as an upper bound, and equivalently, no number larger than $m_A^i(m_B^j)$ can be the lower bound for $f_A(\phi_A^i)(f_B(\phi_B^j))$. The quantity that one is interested in bounding is,

$$\begin{aligned} & sup_{\phi_A^i \in I_A} \left| f_A(\phi_A^i) - f_A(\phi_{A_k}^i) \right| \\ & sup_{\phi_B^j \in I_B} \left| f_B(\phi_B^j) - f_B(\phi_{B_l}^j) \right| \end{aligned} \tag{F.37}$$

The deviation in (F.34) is being maximised at the boundary of the respective segments.

Theorem 3: The values of $|f_A(\phi_A^i) - f_A(\phi_{A_k}^i)|$ and $\left| f_B(\phi_B^j) - f_B(\phi_{B_l}^j) \right|$ are maximum when $\phi_A^i = \phi_{A_k}^i \pm \epsilon_0$ and $\phi_B^i = \phi_{B_k}^i \pm \epsilon_0$

Proof: Before proving the theorem, we focus on some important properties of the function $f(\phi)$. The results of *Theorem 2*, and *Lemma 4* to *Lemma 8* show that the functions $f_A(\phi_A^i)$ and $f_B(\phi_B^j)$ are Lipschitz continuous in the respective intervals I_A and I_B. The functions are strictly monotonous in the intervals due to the monotonicity of the *sin* and *cosine* functions. Moreover, using the second derivative of the functions, one can find that these are concave in the respective intervals. Having mentioned the properties of the solution functions, we now prove the theorem.

Let us assume that the quantity $|f_A(\phi_A^i) - f_A(\phi_{A_k}^i)|$ achieves its maximum at some $\phi_A^{*i} \in I_A$ such that $|\phi_A^{*i} - \phi_{A_k}^i| < \epsilon_0$. From the definition of the segments, we can write the interval as, $I_A = [\phi_{A_k}^i - \epsilon_0, \phi_{A_k}^i + \epsilon_0]$. 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 $\phi_{A_k}^i + \epsilon_0$. Now, if we consider the difference $|f_A(\phi_A^i) - f_A(\phi_{A_k}^i)|$, its absolute values at the endpoints are,

$$\begin{aligned} & |f_A(\phi_{A_k}^i - \epsilon_0) - f_A(\phi_{A_k}^i)| \\ & |f_A(\phi_{A_k}^i + \epsilon_0) - f_A(\phi_{A_k}^i)| \end{aligned} \tag{F.38}$$

Now, for any $\phi_A^{*i} \in (\phi_{A_k}^i - \epsilon_0, \phi_{A_k}^i + \epsilon_0)$ and $\phi_A^{*i} \neq \phi_{A_k}^i$, due to the strict monotonicity of f_A, we have two cases;

- 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 $\phi_B^j = \phi_{B_l}^j \pm \epsilon_0$.

□

We now combine the results of (F.37), (F.34) and *Theorem 3* as,

$$\begin{aligned} sup_{\phi^i_A \in I_A} |f_A(\phi^i_A) - f_A(\phi^i_{A_k})| &= L_A \epsilon_0 \\ sup_{\phi^j_B \in I_B} |f_B(\phi^j_B) - f_B(\phi^j_{B_l})| &= L_B \epsilon_0 \\ \forall \phi^i_A &\in I_A \\ \text{and } \forall \phi^j_B &\in I_B \end{aligned} \tag{F.39}$$

And from the definition of the pessimistic error terms $\Delta\left(\epsilon_0, \phi^i_A\right)$ and $\Delta\left(\epsilon_0, \phi^j_B\right)$ in (F.3), we have

$$\begin{aligned} \Delta\left(\epsilon_0, \phi^i_A\right) &= L_A \epsilon_0 \\ \Delta\left(\epsilon_0, \phi^j_B\right) &= L_B \epsilon_0 \end{aligned} \tag{F.40}$$

Since $f_A(\phi^i_A)$ and $f_B(\phi^j_B)$ are differentiable in the interval I, the values of L_A and L_B can be evaluated using Taylor series expansion of the functions as,

$$\begin{aligned} f_A(\phi^i_A) &= \sum_{n=0}^{\infty} \frac{f^n_A(\phi^{i_*}_A)}{n!} (\phi^i_A - \phi^{i_*}_A)^n \\ f_B(\phi^j_B) &= \sum_{m=0}^{\infty} \frac{f^m_B(\phi^{j_*}_B)}{m!} (\phi^j_B - \phi^{j_*}_B)^m \\ \forall \phi^{i_*}_A &\in I_A \\ \text{and } \forall \phi^{j_*}_B &\in I_B \end{aligned} \tag{F.41}$$

Now in this case, one needs a linear approximation of $f_A(\phi^i_A)$ near $\phi^i_{A_k}$ and $f_B(\phi^j_B)$ near $\phi^j_{B_l}$ for small ϵ_0. One can neglect the higher-order derivatives terms as $f^1_A(\phi^i_{A_k})(\phi^i_A - \phi^i_{A_k})$ and $f^1_B(\phi^j_{B_l})(\phi^j_B - \phi^j_{B_l})$ dominate as $|\phi^i_A - \phi^i_{A_k}|$ and $|\phi^j_B - \phi^j_{B_l}|$ tend to 0 for k^{th} and l^{th} segments, 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 estimation. Thus, using the first order linear approximation one can get an approximation of $f_A(\phi^i_A)$ and $f_A(\phi^j_B)$ near $\phi^i_{A_k}$ and $\phi^j_{B_l}$ as,

$$\begin{aligned} f_A(\phi^i_A) &= f_A(\phi^i_{A_k}) + f^1_A(\phi^i_{A_k})(\phi^i_A - \phi^i_{A_k}) + \mathcal{O}((\phi^i_A - \phi^i_{A_k})^2) \\ |f_A(\phi^i_A) - f_A(\phi^i_{A_k})| &\leq |f^1_A(\phi^i_{A_k})||(\phi^i_A - \phi^i_{A_k})| + \mathcal{O}((\phi^i_A - \phi^i_{A_k})^2) \end{aligned} \tag{F.42}$$

Similarly,

$$|f_B(\phi^j_B) - f_B(\phi^j_{B_l})| \leq |f^1_B(\phi^j_{B_l})||(\phi^j_B - \phi^j_{B_l})| + \mathcal{O}((\phi^j_B - \phi^j_{B_l})^2) \tag{F.43}$$

The higher-order terms, represented by $\mathcal{O}((\phi^i_A - \phi^i_{A_k})^2)$ and $\mathcal{O}((\phi^j_B - \phi^j_{B_l})^2)$, 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 higher-order positive 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, we get,

$$\begin{aligned} sup_{\phi_A^i \in I_A} |f_A(\phi_A^i) - f_A(\phi_{A_k}^i)| \leq |f_A^1(\phi_{A_k}^i)||(\phi_A^i - \phi_{A_k}^i)| \\ sup_{\phi_B^j \in I_B} |f_B(\phi_B^j) - f_B(\phi_{B_l}^j)| \leq |f_B^1(\phi_{B_l}^j)||(\phi_B^j - \phi_{B_l}^j)| \end{aligned} \tag{F.44}$$

Therefore, using Eqs. (F.44) and (F.40), we get the pessimistic error to be,

$$\begin{aligned} \Delta\left(\epsilon_0, \phi_A^i\right) = max_{\phi_A^i \in I} |f_A^1(\phi_{A_k}^i)|\epsilon_0 \\ \Delta\left(\epsilon_0, \phi_B^j\right) = max_{\phi_A^i \in I} |f_B^1(\phi_{B_l}^j)|\epsilon_0 \end{aligned} \tag{F.45}$$

where $L_A = max_{\phi_A^i \in I} |f_A^1(\phi_A^i)|$ and $L_B = max_{\phi_A^i \in I} |f_B^1(\phi_B^j)|$.

4. Relating the change in ϕ_A^i and ϕ_B^j to the optimization problem through CHSH operator

The feasible set in (F.19) can be relaxed by allowing all the density operators σ_{AB}^{ij} that achieve CHSH violation greater than or equal to S_{ij} [SGP+21],

$$\begin{aligned} S_{ij}^{\phi_A^i \cup \phi_B^j} = \{\sigma \mid \exists\, \phi_A \text{ with } \phi_B = \phi_B^j :\ \mathrm{Tr}[\sigma\, CHSH(\phi_A, \phi_B^j)] \geq S_{ij}\} \\ \cup \{\sigma \mid \exists\, \phi_B \text{ with } \phi_A = \phi_A^i :\ \mathrm{Tr}[\sigma\, CHSH(\phi_A^i, \phi_B)] \geq S_{ij}\}. \end{aligned} \tag{F.46}$$

The above equation takes the union of two sets. The first set accounts for all those density operators σ_{AB}^{ij} in the block specified by index i, j that achieve CHSH violation greater than S_{ij} for fixed ϕ_B^j. The second set similarly accounts for the case when ϕ_A^i is fixed . Now, consider the following set,

$$M_{ij,\epsilon}^{\phi_A^i \cup \phi_B^j} := \bigcup_{|\delta_A| \leq \epsilon_0} S_{ij}^{\phi_A^i + \delta_A \,\cup\, \phi_B^j} \cup \bigcup_{|\delta_B| \leq \epsilon_0} S_{ij}^{\phi_A^i \,\cup\, \phi_B^j + \delta_B} \cup \bigcup_{\substack{|\delta_A| \leq \epsilon_0 \\ |\delta_B| \leq \epsilon_0}} S_{ij}^{\phi_A^i + \delta_A \,\cup\, \phi_B^j + \delta_B}. \tag{F.47}$$

In the aforementioned equation, the analysis focuses on a set of density operators derived by introducing a parameter ϵ to the initial angles ϕ_A^i and ϕ_B^j, both individually and concurrently. The defining characteristic of these modified density operators is that the expected value of the CHSH operator evaluated on them remains greater than or equal to the initial CHSH value S_{ij}.

Given discrete points $\phi_{A_k}^i$ and $\phi_{B_l}^j$ 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 $\phi_{A_k+\epsilon}^i$ is equivalent to the deviation at $\phi_{A_k-\epsilon}^i$, and similarly for $\phi_{B_l}^j$. The spectral norm of the difference between the CHSH operators evaluated at $\phi_{A_k}^i$ and $\phi_{A_k\pm\epsilon}^i$ (and similarly for $\phi_{B_l}^j$ and $\phi_{B_l\pm\epsilon}^j$) is equal to the largest singular value of this difference operator. Let us define the maximum of the spectral norms of the differences in the CHSH operator as,

$$\begin{aligned} \delta_p = \max\Big(& ||CHSH(\phi_A^i, \phi_B^j) - CHSH(\phi_A^i + \epsilon, \phi_B^j)||_\infty, \\ & ||CHSH(\phi_A^i, \phi_B^j) - CHSH(\phi_A^i, \phi_B^j + \epsilon)||_\infty, \\ & ||CHSH(\phi_A^i, \phi_B^j) - CHSH(\phi_A^i + \epsilon, \phi_B^j + \epsilon)||_\infty \Big) \end{aligned} \tag{F.48}$$

This spectral norm corresponds to the largest singular value of these difference operators. For the Hermitian CHSH operator, this largest singular value coincides with the absolute value of the largest eigenvalue of the respective difference operators. The spectral norm can be found as,

$$\begin{aligned}&\left\|CHSH(\phi_A^i,\phi_B^j)-CHSH(\phi_A^i+\epsilon,\phi_B^j)\right\|_\infty\\=&\left\|\begin{pmatrix}\cos(\phi_A^i+\epsilon)-\cos\phi_A^i & \sin(\phi_A^i+\epsilon)-\sin\phi_A^i\\ \sin(\phi_A^i+\epsilon)-\sin\phi_A^i & -\cos(\phi_A^i+\epsilon)-\cos\phi_A^i\end{pmatrix}\otimes\begin{pmatrix}-(1-\cos\phi_B^j) & \sin\phi_B^j\\ \sin\phi_B^j & 1-\cos\phi_B^j\end{pmatrix}\right\|_\infty\end{aligned} \tag{F.49}$$

Now, one can use some simple trigonometric relations to decompose the relation $\Delta\cos=\cos(\phi_A^i+\epsilon)-\cos\phi_A^i$ and $\Delta\sin=\sin(\phi_A^i+\epsilon)-\sin\phi_A^i$.

$$\begin{aligned}&\Delta\cos=\cos(\phi_A^i+\epsilon)-\cos(\phi_A^i)=-2\sin\left(\frac{(\phi_A^i+\epsilon)+\phi_A^i}{2}\right)\sin\left(\frac{(\phi_A^i+\epsilon)-\phi_A^i}{2}\right)=-2\sin\left(\phi_A^i+\frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right)\\&\Delta\sin=\sin(\phi_A^i+\epsilon)-\sin(\phi_A^i)=2\cos\left(\frac{(\phi_A^i+\epsilon)+\phi_A^i}{2}\right)\sin\left(\frac{(\phi_A^i+\epsilon)-\phi_A^i}{2}\right)=2\cos\left(\phi_A^i+\frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right)\\&1-\cos\phi_B^j=2\sin^2\left(\frac{\phi_B^j}{2}\right)\\&\sin\phi_B^j=2\sin\left(\frac{\phi_B^j}{2}\right)\cos\left(\frac{\phi_B^j}{2}\right)\end{aligned} \tag{F.50}$$

From (F.50) and (F.49),

$$\begin{aligned}&\left\|CHSH(\phi_A^i,\phi_B^j)-CHSH(\phi_A^i+\epsilon,\phi_B^j)\right\|_\infty\\=&\left\|\begin{pmatrix}-2\sin\left(\phi_A^i+\frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & 2\cos\left(\phi_A^i+\frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right)\\ 2\cos\left(\phi_A^i+\frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & 2\sin\left(\phi_A^i+\frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right)\end{pmatrix}\otimes\begin{pmatrix}-2\sin^2\left(\frac{\phi_B^j}{2}\right) & 2\sin\left(\frac{\phi_B^j}{2}\right)\cos\left(\frac{\phi_B^j}{2}\right)\\ 2\sin\left(\frac{\phi_B^j}{2}\right)\cos\left(\frac{\phi_B^j}{2}\right) & 2\sin^2\left(\frac{\phi_B^j}{2}\right)\end{pmatrix}\right\|_\infty\\=&\left\|2\sin\frac{\epsilon}{2}\left(\cos(\phi_A^i+\frac{\epsilon}{2})\sigma_x-\sin(\phi_A^i+\frac{\epsilon}{2})\sigma_z\right)\otimes 2\left(Q\left(\phi_B^j\right)-\begin{pmatrix}1&0\\0&0\end{pmatrix}\right)\right\|_\infty\\=&\left\|2\sin\left(\frac{\epsilon}{2}\right)\left[\left(\cos\left(\phi_A^i\right)\cos\left(\frac{\epsilon}{2}\right)-\sin\left(\phi_A^i\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_x-\left(\sin\left(\phi_A^i\right)\cos\left(\frac{\epsilon}{2}\right)+\cos\left(\phi_A^i\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_z\right]\right.\\&\qquad\left.\otimes 2\left(Q\left(\phi_B^j\right)-\begin{pmatrix}1&0\\0&0\end{pmatrix}\right)\right\|_\infty\end{aligned} \tag{F.51}$$

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 $x\ll 1, \cos x\approx 1$ and $\sin x\approx x$. Also using $\cos a+\sin b\cdot x\approx\cos a$ and $\sin a+\cos b\cdot x\approx\sin a$ for $x\ll 1$, one can have

$$\begin{aligned}&\sin\left(\frac{\epsilon}{2}\right)\approx\frac{\epsilon}{2}\\&\left(\cos\left(\phi_A^i\right)\cos\left(\frac{\epsilon}{2}\right)-\sin\left(\phi_A^i\right)\sin\left(\frac{\epsilon}{2}\right)\right)\approx\cos\left(\phi_A^i\right)\\&\left(\sin\left(\phi_A^i\right)\cos\left(\frac{\epsilon}{2}\right)+\cos\left(\phi_A^i\right)\sin\left(\frac{\epsilon}{2}\right)\right)\approx\sin\left(\phi_A^i\right)\end{aligned} \tag{F.52}$$

Now, combining the results of (F.51) and linear approximation for $\epsilon>0$ in (F.52) we get,

$$\begin{aligned}&\left\|CHSH(\phi_A^i,\phi_B^j)-CHSH(\phi_A^i+\epsilon,\phi_B^j)\right\|_\infty\\\approx&\left\|\epsilon\left(\cos\left(\phi_A^i\right)\sigma_x-\sin\left(\phi_A^i\right)\sigma_z\right)\otimes 2\left(Q\left(\phi_B^j\right)-\begin{pmatrix}1&0\\0&0\end{pmatrix}\right)\right\|_\infty\end{aligned} \tag{F.53}$$

Similarly, one can find the spectral norm when only ϕ_B^j is being perturbated by ϵ as,

$$
\begin{aligned}
&||CHSH(\phi_A^i, \phi_B^j) - CHSH(\phi_A^i, \phi_B^j + \epsilon)||_\infty \\
=&\left\| \begin{pmatrix} 2\cos^2(\frac{\phi_A^i}{2}) & 2\sin\left(\frac{\phi_A^i}{2}\right)\cos\left(\frac{\phi_A^i}{2}\right) \\ 2\sin\left(\frac{\phi_A^i}{2}\right)\cos\left(\frac{\phi_A^i}{2}\right) & -2\cos^2(\frac{\phi_A^i}{2}) \end{pmatrix} \otimes \begin{pmatrix} 2\sin\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & -2\cos\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \\ -2\cos\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & -2\sin\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \end{pmatrix} \right\|_\infty \\
=&\Bigg\| 2\left(Q\left(\phi_A^i\right) - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) \\
&\qquad \otimes -2\sin\left(\frac{\epsilon}{2}\right)\left[\left(\cos\left(\phi_B^j\right)\cos\left(\frac{\epsilon}{2}\right) - \sin\left(\phi_B^j\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_x - \left(\sin\left(\phi_B^j\right)\cos\left(\frac{\epsilon}{2}\right) + \cos\left(\phi_B^j\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_z\right] \Bigg\|_\infty \\
\approx&\left\| 2\left(Q\left(\phi_A^i\right) - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) \otimes \epsilon\left(\cos\left(\phi_B^j\right)\sigma_x - \sin\left(\phi_B^j\right)\sigma_z\right) \right\|_\infty
\end{aligned}
\tag{F.54}
$$

The spectral norm when both ϕ_A^i and ϕ_B^j perturbated by ϵ is found to be,

$$
\begin{aligned}
&||CHSH(\phi_A^i, \phi_B^j) - CHSH(\phi_A^i + \epsilon, \phi_B^j + \epsilon)||_\infty \\
=&\left\| \begin{pmatrix} 2\sin\left(\phi_A^i + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & -2\cos\left(\phi_A^i + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \\ -2\cos\left(\phi_A^i + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & -2\sin\left(\phi_A^i + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \end{pmatrix} \otimes \begin{pmatrix} 2\sin\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & -2\cos\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \\ -2\cos\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) & -2\sin\left(\phi_B^j + \frac{\epsilon}{2}\right)\sin\left(\frac{\epsilon}{2}\right) \end{pmatrix} \right\|_\infty \\
=&\Bigg\| -2\sin\left(\frac{\epsilon}{2}\right)\left[\left(\cos\left(\phi_A^i\right)\cos\left(\frac{\epsilon}{2}\right) - \sin\left(\phi_A^i\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_x - \left(\sin\left(\phi_A^i\right)\cos\left(\frac{\epsilon}{2}\right) + \cos\left(\phi_A^i\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_z\right] \\
&\qquad \otimes -2\sin\left(\frac{\epsilon}{2}\right)\left[\left(\cos\left(\phi_B^j\right)\cos\left(\frac{\epsilon}{2}\right) - \sin\left(\phi_B^j\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_x - \left(\sin\left(\phi_B^j\right)\cos\left(\frac{\epsilon}{2}\right) + \cos\left(\phi_B^j\right)\sin\left(\frac{\epsilon}{2}\right)\right)\sigma_z\right] \Bigg\|_\infty \\
\approx&\left\| \epsilon\left(\cos\left(\phi_A^i\right)\sigma_x - \sin\left(\phi_A^i\right)\sigma_z\right) \otimes \epsilon\left(\cos\left(\phi_B^j\right)\sigma_x - \sin\left(\phi_B^j\right)\sigma_z\right) \right\|_\infty
\end{aligned}
\tag{F.55}
$$

Hence, combining the results of (F.53), (F.54) and (F.55) into (F.48), we finally get,

$$
\begin{aligned}
\delta_p \approx \max\Bigg(&\left\| \epsilon\left(\cos\left(\phi_A^i\right)\sigma_x - \sin\left(\phi_A^i\right)\sigma_z\right) \otimes 2\left(Q\left(\phi_B^j\right) - \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right) \right\|_\infty, \\
&\left\| 2\left(Q\left(\phi_A^i\right) - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) \otimes \epsilon\left(\cos\left(\phi_B^j\right)\sigma_x - \sin\left(\phi_B^j\right)\sigma_z\right) \right\|_\infty, \\
&\left\| \epsilon\left(\cos\left(\phi_A^i\right)\sigma_x - \sin\left(\phi_A^i\right)\sigma_z\right) \otimes \epsilon\left(\cos\left(\phi_B^j\right)\sigma_x - \sin\left(\phi_B^j\right)\sigma_z\right) \right\|_\infty \Bigg)
\end{aligned}
\tag{F.56}
$$

One can provide an upper bound to the above relation using the property of submultiplicity of spectral norm[Wat18] as,

$$
\begin{aligned}
\delta_p \leq \max\Bigg(&\left\| \epsilon\left(\cos\left(\phi_A^i\right)\sigma_x - \sin\left(\phi_A^i\right)\sigma_z\right) \right\|_\infty \cdot \left\| 2\left(Q\left(\phi_B^j\right) - \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right) \right\|_\infty, \\
&\left\| 2\left(Q\left(\phi_A^i\right) - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) \right\|_\infty \cdot \left\| \epsilon\left(\cos\left(\phi_B^j\right)\sigma_x - \sin\left(\phi_B^j\right)\sigma_z\right) \right\|_\infty, \\
&\left\| \epsilon\left(\cos\left(\phi_A^i\right)\sigma_x - \sin\left(\phi_A^i\right)\sigma_z\right) \right\|_\infty \cdot \left\| \epsilon\left(\cos\left(\phi_B^j\right)\sigma_x - \sin\left(\phi_B^j\right)\sigma_z\right) \right\|_\infty \Bigg) \\
\leq \max&\left(\epsilon \cdot 2\sin\left(\frac{\phi_B^j}{2}\right), \quad 2\cos\left(\frac{\phi_A^i}{2}\right) \cdot \epsilon, \quad \epsilon^2 \right) \\
\leq & \quad 2\epsilon_0
\end{aligned}
\tag{F.57}
$$

The equality holds for the spectral norm of Kronecker products, therefore, this upper bound is exact. Now in the interval I, $\sin\left(\frac{\phi_B^j}{2}\right)$ and $\cos\left(\frac{\phi_A^i}{2}\right)$ functions 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^i_A is fixed at 0 but ϕ^j_B is increased by ϵ_0. Thus every state ρ^{ij}_{AB} that would attain CHSH value S_{ij} at $\phi^i_A = 0, \phi^j_B = \phi^j_{B_l}$ $\left(\phi^i_A = 0, \phi^j_B = \phi^j_B \pm \epsilon_0\right)$ would attain CHSH value $S_{ij} - 2\epsilon_0$ at $\phi^i_A = 0, \phi^j_B = \phi^j_B \pm \epsilon_0$ $\left(\phi^i_A = 0, \phi^j_B = \phi^j_{B_l}\right)$. Thus, we finally have,

$$\begin{aligned} &\mathrm{Tr}\big(\rho\, CHSH(0,\phi^j_B)\big) = S_{ij}, \quad \mathrm{Tr}\big(\rho\, CHSH(0,\phi^j_B+\epsilon_0)\big) = S_{ij}-2\epsilon_0, \quad \mathrm{Tr}\big(\rho\, CHSH(0,\phi^j_B-\epsilon_0)\big) = S_{ij}-2\epsilon_0, \\ \text{or}\quad &\mathrm{Tr}\big(\rho\, CHSH(0,\phi^j_B)\big) = S_{ij}-2\epsilon_0, \quad \mathrm{Tr}\big(\rho\, CHSH(0,\phi^j_B+\epsilon_0)\big) = S_{ij}, \quad \mathrm{Tr}\big(\rho\, CHSH(0,\phi^j_B-\epsilon_0)\big) = S_{ij}, \\ &\forall \rho \in M^{\phi^j_B}_{ij,\epsilon}. \end{aligned} \tag{F.58}$$

where

$$\begin{aligned} &M^{\phi^j_B}_{ij,\epsilon} \\ &= \underbrace{\left\{\rho \in M^{\phi^i_A \cup \phi^j_B}_{ij,\epsilon} \;\middle|\; \mathrm{Tr}\big(\rho\, CHSH(0,\phi^j_B)\big) = S_{ij}\right\}}_{\text{Case A}} \cup \underbrace{\left\{\rho \in M^{\phi^i_A \cup \phi^j_B}_{ij,\epsilon} \;\middle|\; \exists s \in \{+1,-1\} : \mathrm{Tr}\big(\rho\, CHSH(0,\phi^j_B + s\epsilon_0)\big) = S_{ij}\right\}}_{\text{Case B}}. \end{aligned} \tag{F.59}$$

Thus, for each segment I_B and I_A being centralised around discrete $\phi^j_{B_l}$ and $\phi^i_{A_k}$ for l^{th} and k^{th} segment, respectively, performing the optimisation with relaxed constraint value $S - 2\epsilon_0$ would take into account the dependency on ϕ^i_A and ϕ^j_B through the maximum deviation in the CHSH value.

5. Relating the change in ϕ^i_A to the optimization problem

The objective function in (F.19) is a function of the density operator describing the joint state between Alice and Bob and Alice's angle ϕ^i_A. To analyse the dependency of the solution of the optimisation problem on ϕ^i_A, we reinterpret the Frobenius norm terms in (F.19) using dual norms. Dual norm is an indispensable tool in quantum information theory and semi-definite programming because they bridge the limits of quantum operations, e.g., distinguishability, entanglement, and convex optimisation via conic duality, enabling tractable solutions having 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^i_A governs alignment conditions between the states ρ^{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^i_A.

Theorem 4: 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 Hilbert spaces $\mathcal{X}, \mathcal{Y}$. The Frobenius norm is given as[Wat18],

$$\|A\|_F = \sqrt{\mathrm{Tr}\big(A^\dagger A\big)} = \sqrt{\langle A,\, A\rangle}, \tag{F.60}$$

Now, from the definition of the dual of a norm,

$$\|A\|_{F,*} = \{\sup\, \langle Y,\, A\rangle : \|Y\|_F \leq 1\} \tag{F.61}$$

where $Y \in L(\mathcal{Y}, X)$ Now from Cauchy–Schwarz inequality for $\langle\cdot,\cdot\rangle_F$

$$\langle Y,\, A\rangle_F \;\leq\; \|Y\|_F\, \|A\|_F \;\leq\; \|A\|_F \tag{F.62}$$

Thus from (F.61) and (F.62),

$$\|A\|_{F,*} \leq \|A\|_F \tag{F.63}$$

The equality will hold only when Y is aligned with A. Let $Y = A/\|A\|_F$, then $\|Y\|_F = 1$ and

$$\langle Y,\, A\rangle_F = \frac{1}{\|A\|_F}\, \langle A,\, A\rangle_F = \frac{\|A\|_F^2}{\|A\|_F} = \|A\|_F. \tag{F.64}$$

Therefore the supremum is attained when Y aligned with A and we have,

$$\|A\|_{F,*} = \|A\|_F \tag{F.65}$$

□

Now using the definition of channel Λ_0 and Λ_1 in (B.2), one can expand the modified optimisation problem in (F.19) as,

$$\begin{aligned}
n^*(S_{ij}) = \inf \quad & \lambda \mathrm{Tr}\Big(\left[\rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]\right]^* \left[\rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]\right] \Big) \\
& + (1-\lambda)\mathrm{Tr}\left(\left[\rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]\right]^* \left[\rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]\right]\right) \\
& + \frac{\mu}{2}\mathrm{Tr}\left(\left[\rho^{ij}_{AB}\right]^* \left[\rho^{ij}_{AB}\right]\right) \\
= \inf \quad & \lambda \mathrm{Tr}\Big(\left[\rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]\right]^2 \Big) + (1-\lambda)\mathrm{Tr}\left(\left[\rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]\right]^2\right) \\
& + \frac{\mu}{2}\mathrm{Tr}\left(\left[\rho^{ij}_{AB}\right]^2\right) \\
= \inf \quad & \lambda \mathrm{Tr}\Big(\Big[\rho^{ij}_{AB} - \Big(Q(0)\otimes \mathbb{I}\, \rho^{ij}_{AB}\, Q(0)\otimes\mathbb{I} + \{\mathbb{I} - Q(0)\otimes\mathbb{I}\}\, \rho^{ij}_{AB}\, \{\mathbb{I} - Q(0)\otimes\mathbb{I}\}\Big)\Big]^2 \Big) \\
& + (1-\lambda)\mathrm{Tr}\Big(\Big[\rho^{ij}_{AB} - \Big(Q(\phi^i_A)\otimes\mathbb{I}\, \rho^{ij}_{AB}\, Q(\phi^i_A)\otimes\mathbb{I} \\
& \qquad\qquad + \{\mathbb{I} - Q(\phi^i_A)\otimes\mathbb{I}\}\, \rho^{ij}_{AB}\, \{\mathbb{I} - Q(\phi^i_A)\otimes\mathbb{I}\}\Big)\Big]^2\Big) \\
& + \frac{\mu}{2}\mathrm{Tr}\left(\left[\rho^{ij}_{AB}\right]^2\right) \\
& \mathrm{Tr}\left(\rho^{ij}_{AB}\, CHSH(\phi^i_A, \phi^j_B)\right) = S_{ij} \\
\text{s.t.} \quad & \phi^i_A, \phi^j_B \in [0, \pi/2], \\
& \rho^{ij}_{AB} \succeq 0 \\
& \mathrm{Tr}(\rho^{ij}_{AB}) = 1
\end{aligned} \tag{F.66}$$

We can also express the above optimization problem in terms of commutator and anti-commutator between the density matrix and the projectors. The product of any density operator ρ and a projector can be decomposed using commutator and the anti-commutator through the following Lemma.

Lemma 10: $P\rho = \frac{1}{2}\{\rho, P\} - \frac{1}{2}[\rho, P]$ for $\rho \in D(\mathcal{X})$ for some Hilbert space $\mathcal{X}$ and $P \in Pos(\mathcal{X})$.

Proof: Expanding the right-hand using the definition of anti-commutator and commutator, we get,

$$\begin{aligned}
& \frac{1}{2}\{\rho, P\} - \frac{1}{2}[\rho, P] \\
=& \frac{1}{2}\left(\rho P + P\rho\right) - \frac{1}{2}[\rho P - P\rho] \\
=& \frac{1}{2}(P\rho) + \frac{1}{2}(P\rho) \\
=& P\rho
\end{aligned} \tag{F.67}$$

□

Now using the *Lemma 10* in (F.66), one can have,

$$\begin{aligned} n^*(S_{ij}) = \inf \quad & \lambda \mathrm{Tr}\Bigg(\Bigg[\rho^{ij}_{AB} - \Big(\frac{1}{2}\{\rho^{ij}_{AB}, Q(0)\otimes \mathbb{I}\} - \frac{1}{2}[\rho^{ij}_{AB}, Q(0)\otimes\mathbb{I}]\Big) Q(0)\otimes\mathbb{I} \\ & \qquad - \Big(\frac{1}{2}\{\rho^{ij}_{AB}, (\mathbb{I} - Q(0)\otimes\mathbb{I})\} - \frac{1}{2}[\rho^{ij}_{AB}, (\mathbb{I}-Q(0)\otimes\mathbb{I})]\Big)\{\mathbb{I} - Q(0)\otimes\mathbb{I}\}\Bigg]^2\Bigg) \\ & + (1-\lambda)\mathrm{Tr}\Bigg(\Bigg[\rho^{ij}_{AB} - \Big(\frac{1}{2}\{\rho^{ij}_{AB}, Q(\phi^i_A)\otimes \mathbb{I}\} - \frac{1}{2}[\rho^{ij}_{AB}, Q(\phi^i_A)\otimes\mathbb{I}]\Big) Q(\phi^i_A)\otimes\mathbb{I} \\ & \qquad - \Big(\frac{1}{2}\{\rho^{ij}_{AB}, (\mathbb{I} - Q(\phi^i_A)\otimes\mathbb{I})\} - \frac{1}{2}[\rho^{ij}_{AB}, (\mathbb{I}-Q(\phi^i_A)\otimes\mathbb{I})]\Big)\{\mathbb{I} - Q(\phi^i_A)\otimes\mathbb{I}\}\Bigg]^2\Bigg) \\ & + \frac{\mu}{2}\mathrm{Tr}\left(\left[\rho^{ij}_{AB}\right]^2\right) \\ \text{s.t.} \quad & \mathrm{Tr}\left(\rho^{ij}_{AB}\, CHSH(\phi^i_A, \phi^j_B)\right) = S_{ij} \\ & \phi^i_A, \phi^j_B \in [0, \pi/2], \\ & \rho^{ij}_{AB} \succeq 0 \\ & \mathrm{Tr}(\rho^{ij}_{AB}) = 1 \end{aligned} \tag{F.68}$$

a. Perturbing the parameter ϕ^i_A in $Q(\phi^i_A)$

The projector $Q(\cdot)$ is of dimension 2 × 2 and orthogonal onto the one-dimensional subspace spanned by the unit vectors $v(\theta) = \begin{pmatrix}\cos\frac{\theta}{2} \\ \sin\frac{\theta}{2}\end{pmatrix}$. Hence, we can write the projectors as $Q(\theta) = v(\theta)\, v(\theta)^T$. If $Q(\theta)$ is a smooth function, the first-order Taylor expansion of this function about θ under a small increment $\epsilon \ll 1$ yields,

$$Q(\theta + \epsilon) \;=\; Q(\theta) + \epsilon\, Q'(\theta) + \mathcal{O}(\epsilon^2), \tag{F.69}$$

where $Q'(\theta)$ is the first order derivative of the projector Q. 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.

$$Q(\theta + \epsilon) \approx Q(\theta) + \epsilon Q'(\theta) \tag{F.70}$$

The first order derivative of the vector $v(\theta)$ is given as,

$$v'(\theta) = \frac{d}{d\theta}\begin{pmatrix}\cos(\theta/2) \\ \sin(\theta/2)\end{pmatrix} = \begin{pmatrix}-\frac{1}{2}\sin(\theta/2) \\ \frac{1}{2}\cos(\theta/2)\end{pmatrix}. \tag{F.71}$$

Now from the definition of $Q(\theta)$ and matrix product rule one can have,

$$Q'(\theta) \;=\; \frac{d}{d\theta}(v(\theta)v(\theta)^T) = v'(\theta)\, v(\theta)^T + v(\theta)\, \left(v'(\theta)\right)^T \tag{F.72}$$

and finally,

$$Q'(\theta) \;=\; v(\theta)'v(\theta)^T + v(\theta)(v(\theta)')^T = \begin{pmatrix} -\cos(\theta/2)\sin(\theta/2) & \frac{1}{2}(\cos^2(\theta/2) - \sin^2(\theta/2)) \\ \frac{1}{2}(\cos^2(\theta/2) - \sin^2(\theta/2)) & \cos(\theta/2)\sin(\theta/2) \end{pmatrix}. \tag{F.73}$$

Using the identities $\sin\theta = 2\sin(\theta/2)\cos(\theta/2)$ and $\cos\theta = \cos^2(\theta/2) - \sin^2(\theta/2)$, the above equation simplifies to the following compact form,

$$Q'(\theta) \;=\; \frac{1}{2}\begin{pmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{pmatrix} \tag{F.74}$$

Hence, the first-order Taylor approximation of $Q(\theta)$ comes out to be,

$$\begin{aligned} Q(\theta+\epsilon) &\approx Q(\theta) + \epsilon\, Q'(\theta) \\ &\approx \begin{pmatrix} \cos^2\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) \\ \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) & \sin^2\left(\frac{\theta}{2}\right) \end{pmatrix} + \frac{\epsilon}{2}\begin{pmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{pmatrix} \\ &\approx \begin{pmatrix} \cos^2\left(\frac{\theta}{2}\right) - \frac{\epsilon\sin(\theta)}{2} & \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) + \frac{\epsilon\cos(\theta)}{2} \\ \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) + \frac{\epsilon\cos(\theta)}{2} & \sin^2\left(\frac{\theta}{2}\right) + \frac{\epsilon\sin(\theta)}{2} \end{pmatrix} \end{aligned} \tag{F.75}$$

After establishing an approximate value of $Q(\phi+\epsilon)$ one can proceed by expressing $(Q(\phi^i_A+\epsilon)\otimes\mathbb{I})\rho^{ij}_{AB}$ in first order linear approximation as,

$$(Q(\phi^i_A+\epsilon)\otimes\mathbb{I})\rho^{ij}_{AB} = (Q(\phi^i_A)\otimes\mathbb{I})\rho^{ij}_{AB} + \epsilon(Q'(\phi^i_A)\otimes\mathbb{I})\rho^{ij}_{AB} + \mathcal{O}(\epsilon^2) \tag{F.76}$$

and terms of anti-commuting and commuting operators from the result of *Lemma 10* and (F.69) as,

$$(Q(\phi^i_A+\epsilon)\otimes\mathbb{I})\rho^{ij}_{AB} = \frac{1}{2}\{\rho^{ij}_{AB}, Q(\phi^i_A+\epsilon)\otimes\mathbb{I}\} - \frac{1}{2}[\rho^{ij}_{AB}, Q(\phi^i_A+\epsilon)\otimes\mathbb{I}] \tag{F.77}$$

We expand the commuting and the anti-commuting term individually as,

$$\begin{aligned} [\rho^{ij}_{AB}, Q(\phi^i_A+\epsilon)\otimes\mathbb{I}] &= \rho^{ij}_{AB}\, Q(\phi^i_A+\epsilon)\otimes\mathbb{I} - Q(\phi^i_A+\epsilon)\otimes\mathbb{I}\,\rho^{ij}_{AB} \\ &= \rho^{ij}_{AB}\left(Q(\phi^i_A)\otimes\mathbb{I} + \epsilon(Q'(\phi^i_A)\otimes\mathbb{I}) + \mathcal{O}(\epsilon^2)\right) \\ &\quad - \left(Q(\phi^i_A)\otimes\mathbb{I} + \epsilon(Q'(\phi^i_A)\otimes\mathbb{I}) + \mathcal{O}(\epsilon^2)\right)\rho^{ij}_{AB} \\ &= \left(\rho^{ij}_{AB}\, Q(\phi^i_A)\otimes\mathbb{I} - Q(\phi^i_A)\otimes\mathbb{I}\,\rho^{ij}_{AB}\right) \\ &\quad + \epsilon\left(\rho^{ij}_{AB}\, Q'(\phi^i_A)\otimes\mathbb{I} - Q'(\phi^i_A)\otimes\mathbb{I}\,\rho^{ij}_{AB}\right) + \mathcal{O}\left(\epsilon^2\right) \\ &= [\rho^{ij}_{AB}, Q(\phi^i_A)\otimes\mathbb{I}] + \epsilon[\rho^{ij}_{AB}, Q'(\phi^i_A)\otimes\mathbb{I}] + \mathcal{O}(\epsilon^2) \end{aligned} \tag{F.78}$$

Similarly, for other terms we get,

$$\begin{aligned} &\{\rho^{ij}_{AB}, Q(\phi^i_A+\epsilon)\otimes\mathbb{I}\} = \{\rho^{ij}_{AB}, Q(\phi^i_A)\otimes\mathbb{I}\} + \epsilon\{\rho^{ij}_{AB}, Q'(\phi^i_A)\otimes\mathbb{I}\} + \mathcal{O}(\epsilon^2) \\ &[\rho^{ij}_{AB}, (\mathbb{I} - Q(\phi^i_A+\epsilon)\otimes\mathbb{I})] = [\rho^{ij}_{AB}, (\mathbb{I} - Q(\phi^i_A)\otimes\mathbb{I})] + \epsilon[\rho^{ij}_{AB}, (\mathbb{I} - Q'(\phi^i_A)\otimes\mathbb{I})] + \mathcal{O}(\epsilon^2) \\ &\{\rho^{ij}_{AB}, (\mathbb{I} - Q(\phi^i_A+\epsilon)\otimes\mathbb{I})\} = \{\rho^{ij}_{AB}, (\mathbb{I} - Q(\phi^i_A)\otimes\mathbb{I})\} + \epsilon\{\rho^{ij}_{AB}, (\mathbb{I} - Q'(\phi^i_A)\otimes\mathbb{I})\} + \mathcal{O}(\epsilon^2) \end{aligned} \tag{F.79}$$

Now, let us consider the following function,

$$h : (\lambda, \phi^i_A, \rho^{ij}_{AB}) \mapsto \mathbb{R} \tag{F.80}$$

being defined as follows,

$$\begin{aligned} h(\lambda, \phi^i_A, \rho^{ij}_{AB}) = (1-\lambda)\mathrm{Tr}\Bigg(\Bigg[\rho^{ij}_{AB} - \left(\frac{1}{2}\{\rho^{ij}_{AB}, Q(\phi^i_A)\otimes\mathbb{I}\} - \frac{1}{2}[\rho^{ij}_{AB}, Q(\phi^i_A)\otimes\mathbb{I}]\right) Q(\phi^i_A)\otimes\mathbb{I} \\ - \left(\frac{1}{2}\{\rho^{ij}_{AB}, (\mathbb{I} - Q(\phi^i_A)\otimes\mathbb{I})\} - \frac{1}{2}[\rho^{ij}_{AB}, (\mathbb{I} - Q(\phi^i_A)\otimes\mathbb{I})]\right)\{\mathbb{I} - Q(\phi^i_A)\otimes\mathbb{I}\}\Bigg]^2\Bigg) \end{aligned} \tag{F.81}$$

Now, one can bound the small perturbation ϵ in ϕ^i_A using triangular inequality as,

$$\begin{aligned}
&\Big|h(\lambda,\phi^i_A,\rho^{ij}_{AB})-h(\lambda,\phi^i_A+\epsilon,\rho^{ij}_{AB})\Big|\\
\leq&\Big|h(\lambda,\phi^i_A,\rho^{ij}_{AB})\Big|+\Big|h(\lambda,\phi^i_A+\epsilon,\rho^{ij}_{AB})\Big|\\
\leq&(1-\lambda)\mathrm{Tr}\Bigg(\Bigg[\rho^{ij}_{AB}-\Bigg(\frac{1}{2}\{\rho^{ij}_{AB},Q(\phi^i_A)\otimes\mathbb{I}\}-\frac{1}{2}[\rho^{ij}_{AB},Q(\phi^i_A)\otimes\mathbb{I}]\Bigg)\,Q(\phi^i_A)\otimes\mathbb{I}\\
&\qquad-\Bigg(\frac{1}{2}\{\rho^{ij}_{AB},(\mathbb{I}-Q(\phi^i_A)\otimes\mathbb{I})\}-\frac{1}{2}[\rho^{ij}_{AB},(\mathbb{I}-Q(\phi^i_A)\otimes\mathbb{I})]\Bigg)\,\{\mathbb{I}-Q(\phi^i_A)\otimes\mathbb{I}\}\Bigg]^2\Bigg)\\
&+(1-\lambda)\mathrm{Tr}\Bigg(\Bigg[\rho^{ij}_{AB}-\Bigg(\frac{1}{2}\{\rho^{ij}_{AB},Q(\phi^i_A+\epsilon_0)\otimes\mathbb{I}\}-\frac{1}{2}[\rho^{ij}_{AB},Q(\phi^i_A+\epsilon_0)\otimes\mathbb{I}]\Bigg)\,Q(\phi^i_A+\epsilon_0)\otimes\mathbb{I}\\
&\qquad-\Bigg(\frac{1}{2}\{\rho^{ij}_{AB},(\mathbb{I}-Q(\phi^i_A+\epsilon_0)\otimes\mathbb{I})\}-\frac{1}{2}[\rho^{ij}_{AB},(\mathbb{I}-Q(\phi^i_A+\epsilon_0)\otimes\mathbb{I})]\Bigg)\,\{\mathbb{I}-Q(\phi^i_A+\epsilon_0)\otimes\mathbb{I}\}\Bigg]^2\Bigg)\\
\leq&(1-\lambda)\Big[\mathrm{Tr}\Bigg(\Bigg[\rho^{ij}_{AB}-\Bigg(\frac{1}{2}\{\rho^{ij}_{AB},Q(\phi^i_A)\otimes\mathbb{I}\}-\frac{1}{2}[\rho^{ij}_{AB},Q(\phi^i_A)\otimes\mathbb{I}]\Bigg)\,Q(\phi^i_A)\otimes\mathbb{I}\\
&\qquad-\Bigg(\frac{1}{2}\{\rho^{ij}_{AB},(\mathbb{I}-Q(\phi^i_A)\otimes\mathbb{I})\}-\frac{1}{2}[\rho^{ij}_{AB},(\mathbb{I}-Q(\phi^i_A)\otimes\mathbb{I})]\Bigg)\,(\mathbb{I}-Q(\phi^i_A)\otimes\mathbb{I})\Bigg]^2\\
&\qquad+\Bigg[\rho^{ij}_{AB}-\Bigg(\frac{1}{2}\{\rho^{ij}_{AB},Q(\phi^i_A+\epsilon_0)\otimes\mathbb{I}\}-\frac{1}{2}[\rho^{ij}_{AB},Q(\phi^i_A+\epsilon_0)\otimes\mathbb{I}]\Bigg)\,Q(\phi^i_A+\epsilon_0)\otimes\mathbb{I}\\
&\qquad-\Bigg(\frac{1}{2}\{\rho^{ij}_{AB},(\mathbb{I}-Q(\phi^i_A+\epsilon_0)\otimes\mathbb{I})\}-\frac{1}{2}[\rho^{ij}_{AB},(\mathbb{I}-Q(\phi^i_A+\epsilon_0)\otimes\mathbb{I})]\Bigg)\,(\mathbb{I}-Q(\phi^i_A+\epsilon_0)\otimes\mathbb{I})\Bigg]^2\Bigg)\,\Big]
\end{aligned}\tag{F.82}$$

To simplify the calculation, we simplify the perturbing and non-perturbing terms separately as,

$$\begin{aligned}
\mathrm{A}&=\Bigg[\rho^{ij}_{AB}-\Bigg(\frac{1}{2}\{\rho^{ij}_{AB},P_0\}-\frac{1}{2}[\rho^{ij}_{AB},P_0]\Bigg)\,P_0-\Bigg(\frac{1}{2}\{\rho^{ij}_{AB},(\mathbb{I}-P_0)\}-\frac{1}{2}[\rho^{ij}_{AB},(\mathbb{I}-P_0)]\Bigg)\,(\mathbb{I}-P_0)\Bigg]^2\\
&=\Bigg[\rho^{ij}_{AB}-\frac{1}{2}\{\rho^{ij}_{AB},P_0\}P_0+\frac{1}{2}[\rho^{ij}_{AB},P_0]P_0-\frac{1}{2}\{\rho^{ij}_{AB},(\mathbb{I}-P_0)\}(\mathbb{I}-P_0)+\frac{1}{2}[\rho^{ij}_{AB},(\mathbb{I}-P_0)](\mathbb{I}-P_0)\Bigg]^2\\
&=\Big[\rho^{ij}_{AB}-(P_0)\,\rho^{ij}_{AB}\,(P_0)-(\mathbb{I}-P_0)\,\rho^{ij}_{AB}\,(\mathbb{I}-P_0)\Big]^2\\
&=\Big[\rho^{ij}_{AB}\,P_0+P_0\,\rho^{ij}_{AB}-2P_0\,\rho^{ij}_{AB}\,P_0\Big]^2\\
&=\Big[(\mathbb{I}-P_0)\,\rho^{ij}_{AB}\,(P_0)+(P_0)\,\rho^{ij}_{AB}(\mathbb{I}-P_0)\Big]^2\\
&=\Big((\mathbb{I}-P_0)\,\rho^{ij}_{AB}\,Q(\phi^i_A)\otimes\mathbb{I}\,\rho^{ij}_{AB}(\mathbb{I}-P_0)+P_0\,\rho^{ij}_{AB}(\mathbb{I}-P_0)\,\rho^{ij}_{AB}\,P_0\Big)\\
&=A_0+A_1
\end{aligned}\tag{F.83}$$

and

$$\begin{aligned}
\mathrm{B} = &\left[\rho^{ij}_{AB} - \left(\frac{1}{2}\left\{\rho^{ij}_{AB}, Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\right\} - \frac{1}{2}\left[\rho^{ij}_{AB}, Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\right]\right) Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\right.\\
&\left. - \left(\frac{1}{2}\left\{\rho^{ij}_{AB}, \mathbb{I} - Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\right\} - \frac{1}{2}\left[\rho^{ij}_{AB}, \mathbb{I} - Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\right]\right) (\mathbb{I} - Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I})\right]^2\\
= &\left[\{\rho^{ij}_{AB}, Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\} - 2Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\rho^{ij}_{AB}Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\right]^2\\
= &\left[\left(\{\rho^{ij}_{AB}, Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\}\right)^2 + \left(-2Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\rho^{ij}_{AB}Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\right)^2\right.\\
&-\{\rho^{ij}_{AB}, Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\}2Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\rho^{ij}_{AB}Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\\
&\left.-2Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\rho^{ij}_{AB}Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\{\rho^{ij}_{AB}, Q(\phi^i_A + \epsilon_0) \otimes \mathbb{I}\}\right]\\
= &\left[\left(\{\rho^{ij}_{AB}, P_0\} + \epsilon_0\{\rho^{ij}_{AB}, P_1\} + \mathcal{O}(\epsilon_0^2)\right)^2 + \left(4\xi\rho^{ij}_{AB}\xi\rho^{ij}_{AB}\xi\right) - \left(\{\rho^{ij}_{AB}, \xi\}2\xi\rho^{ij}_{AB}\xi\right) - \left(2\xi\rho^{ij}_{AB}\xi\{\rho^{ij}_{AB}, \xi\}\right)\right]\\
= &\left[\left(\{\rho^{ij}_{AB}, P_0\}^2 + \epsilon_0\left(\{\rho^{ij}_{AB}, P_0\}\{\rho^{ij}_{AB}, P_1\} + \{\rho^{ij}_{AB}, P_1\}\{\rho^{ij}_{AB}, P_0\}\right) + \epsilon_0^2\{\rho^{ij}_{AB}, P_1\}^2\right) + \left(4\xi\rho^{ij}_{AB}\xi\rho^{ij}_{AB}\xi\right)\right.\\
&\left. - \left(\{\rho^{ij}_{AB}, \xi\}2\xi\rho^{ij}_{AB}\xi\right) - \left(2\xi\rho^{ij}_{AB}\xi\{\rho^{ij}_{AB}, \xi\}\right)\right] + \mathcal{O}(\epsilon_0^3)\\
= &\left[\left(\{\rho^{ij}_{AB}, P_0\}^2 + \epsilon_0\left(\{\rho^{ij}_{AB}, P_0\}\{\rho^{ij}_{AB}, P_1\} + \{\rho^{ij}_{AB}, P_1\}\{\rho^{ij}_{AB}, P_0\}\right) + \epsilon_0^2\{\rho^{ij}_{AB}, P_1\}^2\right)\right.\\
&+4\left(P_0\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_0 + \epsilon_0(P_1\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_0 + P_0\rho^{ij}_{AB}P_1\rho^{ij}_{AB}P_0 + P_0\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_1)\right.\\
&\left.\left.+\epsilon_0^2(P_0\rho^{ij}_{AB}P_1\rho^{ij}_{AB}P_1 + P_1\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_1 + P_1\rho^{ij}_{AB}P_1\rho^{ij}_{AB}P_0 + P_1\rho^{ij}_{AB}P_1\rho^{ij}_{AB}P_1)\right) - C_1 - C_2\right] + \mathcal{O}(\epsilon_0^3)\\
= &B_0 + \epsilon_0 B_1 + \epsilon_0^2 B_2 + \mathcal{O}(\epsilon_0^3)
\end{aligned} \tag{F.84}$$

$$\begin{aligned}
B_0 = &\{\rho^{ij}_{AB}, P_0\}^2 + 4P_0\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_0 - 2\{\rho, P_0\}P_0\rho P_0 - 2P_0\rho P_0\{\rho, P_0\}\\
B_1 = &\{\rho^{ij}_{AB}, P_0\}\{\rho^{ij}_{AB}, P_1\} + \{\rho^{ij}_{AB}, P_1\}\{\rho^{ij}_{AB}, P_0\} + 4(P_1\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_0 + P_0\rho^{ij}_{AB}P_1\rho^{ij}_{AB}P_0 + P_0\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_1)\\
&- 2(\{\rho, P_0\}(P_0\rho P_1 + P_1\rho P_0) + \{\rho, P_1\}P_0\rho P_0) - 2((P_0\rho P_1 + P_1\rho P_0)\{\rho, P_0\} + P_0\rho P_0\{\rho, P_1\})\\
B_2 = &\{\rho^{ij}_{AB}, P_1\}^2 + 4(P_0\rho^{ij}_{AB}P_1\rho^{ij}_{AB}P_1 + P_1\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_1 + P_1\rho^{ij}_{AB}P_1\rho^{ij}_{AB}P_0 + P_1\rho^{ij}_{AB}P_1\rho^{ij}_{AB}P_1)\\
&- 2\left(\{\rho, P_0\}P_1\rho P_1 + \{\rho, P_1\}(P_0\rho P_1 + P_1\rho P_0)\right) - 2\left(P_1\rho P_1\{\rho, P_0\} + (P_0\rho P_1 + P_1\rho P_0)\{\rho, P_1\}\right)
\end{aligned}$$

where $\xi = (P_0 + \epsilon_0 P_1 + \mathcal{O}(\epsilon_0^2)), P_0 = Q(\phi^i_A) \otimes \mathbb{I}, P_1 = Q'(\phi^i_A) \otimes \mathbb{I}$

$$C_1 = -2\left[\{\rho, P_0\}P_0\rho P_0 + \epsilon_0\left(\{\rho, P_0\}(P_0\rho P_1 + P_1\rho P_0) + \{\rho, P_1\}P_0\rho P_0\right) + \epsilon_0^2\left(\{\rho, P_0\}P_1\rho P_1 + \{\rho, P_1\}(P_0\rho P_1 + P_1\rho P_0)\right)\right] + \mathcal{O}(\epsilon_0^2)$$

$$C_2 = -2\left[P_0\rho P_0\{\rho, P_0\} + \epsilon_0\left((P_0\rho P_1 + P_1\rho P_0)\{\rho, P_0\} + P_0\rho P_0\{\rho, P_1\}\right) + \epsilon_0^2\left(P_1\rho P_1\{\rho, P_0\} + (P_0\rho P_1 + P_1\rho P_0)\{\rho, P_1\}\right)\right] + \mathcal{O}(\epsilon_0^2)$$

Inserting the explicit forms of A and B in (F.82) yields,

$$\begin{aligned}
&\left|h(\lambda, \phi^i_A, \rho^{ij}_{AB}) - h(\lambda, \phi^i_A + \epsilon_0, \rho^{ij}_{AB})\right|\\
&\leq (1-\lambda)\left[\mathrm{Tr}(A + B)\right]\\
&\leq (1-\lambda)\left[\mathrm{Tr}(A_0) + \mathrm{Tr}(A_1) + \mathrm{Tr}(B_0) + \epsilon_0\mathrm{Tr}(B_1) + \epsilon_0^2\,\mathrm{Tr}(B_2)\right]
\end{aligned} \tag{F.85}$$

Now, upon solving each of the term separately we have,

$$\begin{aligned}
\mathrm{Tr}(A_0) &= \mathrm{Tr}((\mathbb{I} - Q(\phi^i_A) \otimes \mathbb{I})\,\rho^{ij}_{AB}\,Q(\phi^i_A) \otimes \mathbb{I}\,\rho^{ij}_{AB}(\mathbb{I} - Q(\phi^i_A) \otimes \mathbb{I}))\\
&= \mathrm{Tr}((\mathbb{I} - Q(\phi^i_A) \otimes \mathbb{I})\,\rho^{ij}_{AB}\,(Q(\phi^i_A) \otimes \mathbb{I})\,\rho^{ij}_{AB})\\
&= \mathrm{Tr}(\,\rho^{ij}_{AB}\,(\mathbb{I} - Q(\phi^i_A) \otimes \mathbb{I})\,\rho^{ij}_{AB}\,(Q(\phi^i_A) \otimes \mathbb{I}))\\
&\leq \mathrm{Tr}(\,\rho^{ij}_{AB}\,(\mathbb{I} - Q(\phi^i_A) \otimes \mathbb{I})\,\rho^{ij}_{AB})
\end{aligned} \tag{F.86}$$

The inequality in the above equation arises from the Loewner order for two hermitian $\mathbb{I}$ and projector $P = (Q(\phi_A^i) \otimes \mathbb{I})$ as,

$$\begin{aligned}
&(\mathbb{I} - P) \succeq 0 \\
&P \preceq \mathbb{I} \\
&YP \preceq Y\mathbb{I} \quad (\text{for Y} = \rho(\mathbb{I} - P)\rho \succeq 0) \\
&YP \preceq Y \\
&Y - YP \succeq 0 \\
&\text{Tr}(Y - YP) \geq 0 \\
&\text{Tr}(Y) - \text{Tr}(YP) \geq 0 \\
&\text{Tr}(Y) \geq \text{Tr}(YP) \\
&\text{Tr}(\rho(\mathbb{I} - P)\rho) \geq \text{Tr}(\rho(\mathbb{I} - P)\rho P)
\end{aligned} \tag{F.87}$$

for $\rho = \rho_{AB}^{ij}$.
Coming back to the proof for $\text{Tr}(A_0)$

$$\begin{aligned}
\text{Tr}(A_0) &\leq \text{Tr}(\ \rho_{AB}^{ij}\ (\mathbb{I} - Q(\phi_A^i) \otimes \mathbb{I})\ \rho_{AB}^{ij}) \\
&\leq \text{Tr}(\ \rho_{AB}^{ij}\ (\mathbb{I} - Q(\phi_A^i) \otimes \mathbb{I})) \\
&\leq \text{Tr}(\ \rho_{AB}^{ij}) - \text{Tr}(\ \rho_{AB}^{ij}(Q(\phi_A^i) \otimes \mathbb{I}) \\
&\leq 1 - \text{Tr}(\ \rho_{AB}^{ij}(Q(\phi_A^i) \otimes \mathbb{I}) \\
&\leq 1 \quad (\ \text{as Tr}(\ \rho_{AB}^{ij}(Q(\phi_A^i) \otimes \mathbb{I}) \geq 0\ )
\end{aligned} \tag{F.88}$$

Similarly,

$$\begin{aligned}
\text{Tr}(A_1) &= \text{Tr}((Q(\phi_A^i) \otimes \mathbb{I})\ \rho_{AB}^{ij}(\mathbb{I} - Q(\phi_A^i) \otimes \mathbb{I})\ \rho_{AB}^{ij}\ (Q(\phi_A^i) \otimes \mathbb{I})) \\
&= \text{Tr}((Q(\phi_A^i) \otimes \mathbb{I})\ \rho_{AB}^{ij}\ (\mathbb{I} - Q(\phi_A^i) \otimes \mathbb{I})\ \rho_{AB}^{ij}) \\
&\leq 1
\end{aligned} \tag{F.89}$$

Now we can proceed towards computing other trace in a similar fashion.

$$\begin{aligned}
\text{Tr}(B_0) &= \text{Tr}\left(\{\rho_{AB}^{ij}, P_0\}^2 + 4P_0\rho_{AB}^{ij}P_0\rho_{AB}^{ij}P_0 - 2\{\rho, P_0\}P_0\rho P_0 - 2P_0\rho P_0\{\rho, P_0\}\right) \\
&= \text{Tr}\left((\rho_{AB}^{ij}\ P_0 + P_0\ \rho_{AB}^{ij})^2 + 4P_0\rho_{AB}^{ij}P_0\rho_{AB}^{ij}P_0 - 2\{\rho, P_0\}P_0\rho P_0 - 2P_0\rho P_0\{\rho, P_0\}\right) \\
&= \text{Tr}\left((\rho_{AB}^{ij}\ P_0)^2 + (P_0\ \rho_{AB}^{ij})^2 + \rho_{AB}^{ij}\ P_0P_0\ \rho_{AB}^{ij} + P_0\ \rho_{AB}^{ij}\rho_{AB}^{ij}\ P_0 + 4P_0\rho_{AB}^{ij}P_0\rho_{AB}^{ij}P_0 - 2\{\rho, P_0\}P_0\rho P_0 - 2P_0\rho P_0\{\rho, P_0\}\right) \\
&= \Big(\text{Tr}(\rho_{AB}^{ij}\ P_0\ \rho_{AB}^{ij}\ P_0) + \text{Tr}(P_0\ \rho_{AB}^{ij}P_0\ \rho_{AB}^{ij}) + \text{Tr}(\rho_{AB}^{ij}\ P_0\ \rho_{AB}^{ij}) + \text{Tr}(P_0\ \rho_{AB}^{ij}\rho_{AB}^{ij}\ P_0) + \text{Tr}(4P_0\rho_{AB}^{ij}P_0\rho_{AB}^{ij}P_0) \\
&\qquad -2\text{Tr}((\rho\ P_0\ \rho P_0) - 2\text{Tr}(P_0\ \rho\ P_0\rho P_0) - 2\text{Tr}(P_0\rho P_0\ \rho\ P_0) - 2\text{Tr}(P_0\rho P_0\ \rho))) \\
&= \Big(\text{Tr}(\ P_0\rho_{AB}^{ij}\ P_0\ \rho_{AB}^{ij}) + \text{Tr}(P_0\ \rho_{AB}^{ij}P_0\ \rho_{AB}^{ij}) + \text{Tr}(\rho_{AB}^{ij}\ P_0\ \rho_{AB}^{ij}) + \text{Tr}(P_0\ \rho_{AB}^{ij}\rho_{AB}^{ij}) + 4\text{Tr}(P_0\rho_{AB}^{ij}P_0\rho_{AB}^{ij}) \\
&\qquad -2\text{Tr}((P_0\rho\ P_0\ \rho) - 2\text{Tr}(P_0\ \rho\ P_0\rho) - 2\text{Tr}(P_0\rho P_0\ \rho\ ) - 2\text{Tr}(P_0\rho P_0\ \rho))) \\
&= \Big(\text{Tr}(P_0\ \rho_{AB}^{ij}\rho_{AB}^{ij})) + \text{Tr}(P_0\ \rho_{AB}^{ij}\rho_{AB}^{ij}) - 2\text{Tr}(P_0\rho P_0\ \rho))\Big) \\
&= 2(\text{Tr}(P_0\ (\rho_{AB}^{ij})^2) - \text{Tr}((P_0\rho_{AB}^{ij})^2))
\end{aligned} \tag{F.90}$$

Using a similar form of Loewner order for two hermetian $\mathbb{I}$ and projector P it can be proven that $0 \leq \text{Tr}(P_0\ (\rho_{AB}^{ij})^2) \leq$

1 and $\mathrm{Tr}(P_0\ (\rho^{ij}_{AB})^2) \geq \mathrm{Tr}((P_0\rho^{ij}_{AB})^2)$.

$$
\begin{aligned}
&P \preceq \mathbb{I}\\
&(\rho^{ij}_{AB})^2\ P \preceq (\rho^{ij}_{AB})^2 \qquad (Y = (\rho^{ij}_{AB})^2 \succeq 0)\\
&(\rho^{ij}_{AB})^2 - (\rho^{ij}_{AB})^2\ P \succeq 0\\
&\mathrm{Tr}((\rho^{ij}_{AB})^2) - \mathrm{Tr}((\rho^{ij}_{AB})^2\ P) \geq 0\\
&1 - \mathrm{Tr}((\rho^{ij}_{AB})^2\ P) \geq 0\\
&0 \leq \mathrm{Tr}((\rho^{ij}_{AB})^2\ P) \leq 1 \quad (\text{since } \rho^{ij}_{AB} \succeq 0 \text{ and } P \succeq 0, \text{ lower bound by } 0)
\end{aligned}
\tag{F.91}
$$

and ,

$$
\begin{aligned}
&P \preceq \mathbb{I}\\
&(\rho^{ij}_{AB}\ P\ \rho^{ij}_{AB})P \preceq (\rho^{ij}_{AB}\ P\ \rho^{ij}_{AB}) \qquad (Y = \rho^{ij}_{AB}\ P\ \rho^{ij}_{AB} \succeq 0)\\
&\mathrm{Tr}(\rho^{ij}_{AB}\ P\ \rho^{ij}_{AB}) - \mathrm{Tr}((\rho^{ij}_{AB}\ P)^2) \geq 0\\
&\mathrm{Tr}(\rho^{ij}_{AB}\ P\ \rho^{ij}_{AB}) \geq \mathrm{Tr}((\rho^{ij}_{AB}\ P)^2)
\end{aligned}
$$

Thus,

$$\mathrm{Tr}(B_0) \leq 2 \tag{F.92}$$

$\mathrm{Tr}(B_1)$ can also be computed in a similar fashion.

$$
\begin{aligned}
\mathrm{Tr}(B_1) &= \mathrm{Tr}(\{\rho^{ij}_{AB}, P_0\}\{\rho^{ij}_{AB}, P_1\} + \{\rho^{ij}_{AB}, P_1\}\{\rho^{ij}_{AB}, P_0\} + 4(P_1\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_0 + P_0\rho^{ij}_{AB}P_1\rho^{ij}_{AB}P_0 + P_0\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_1)\\
&\quad - 2(\{\rho, P_0\}(P_0\rho P_1 + P_1\rho P_0) + \{\rho, P_1\}P_0\rho P_0) - 2((P_0\rho P_1 + P_1\rho P_0)\{\rho, P_0\} + P_0\rho P_0\{\rho, P_1\}))\\
&= \mathrm{Tr}\left((\rho^{ij}_{AB}\ P_0 +\ P_0\ \rho^{ij}_{AB})(\rho^{ij}_{AB}\ P_1 +\ P_1\ \rho^{ij}_{AB})\right) + \mathrm{Tr}\left((\rho^{ij}_{AB}\ P_1 +\ P_1\ \rho^{ij}_{AB})(\rho^{ij}_{AB}\ P_0 +\ P_0\ \rho^{ij}_{AB})\right)\\
&\qquad + 4\ \mathrm{Tr}\left(P_1\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_0 + P_0\rho^{ij}_{AB}P_1\rho^{ij}_{AB}P_0 + P_0\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_1\right) - 2\mathrm{Tr}\left(\{\rho, P_0\}(P_0\rho P_1 + P_1\rho P_0) + \{\rho, P_1\}P_0\rho P_0\right)\\
&\qquad - 2\mathrm{Tr}\left((P_0\rho P_1 + P_1\rho P_0)\{\rho, P_0\} + P_0\rho P_0\{\rho, P_1\}\right)\\
&= \mathrm{Tr}\left(\rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_1\right) + \mathrm{Tr}\left(\rho^{ij}_{AB}\ P_0\ P_1\ \rho^{ij}_{AB}\right) + \mathrm{Tr}\left(P_0\ \rho^{ij}_{AB}\ \rho^{ij}_{AB}\ P_1\right) + \mathrm{Tr}\left(P_0\ \rho^{ij}_{AB}\ P_1\ \rho^{ij}_{AB}\right)\\
&\qquad + \mathrm{Tr}\left(\rho^{ij}_{AB}\ P_1\ \rho^{ij}_{AB}\ P_0\right) + \mathrm{Tr}\left(\rho^{ij}_{AB}\ P_1\ P_0\ \rho^{ij}_{AB}\right) + \mathrm{Tr}\left(P_1\ \rho^{ij}_{AB}\ \rho^{ij}_{AB}\ P_0\right) + \mathrm{Tr}\left(P_1\ \rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\right)\\
&\qquad + 4\mathrm{Tr}\left(P_1\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_0\right) + 4\mathrm{Tr}\left(P_0\rho^{ij}_{AB}P_1\rho^{ij}_{AB}P_0\right) + 4\mathrm{Tr}\left(P_0\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_1\right)\\
&\qquad - 2\mathrm{Tr}\left(\rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_1\right) - 2\mathrm{Tr}\left(\rho^{ij}_{AB}\ P_0\ P_1\ \rho^{ij}_{AB}\ P_0\right) - 2\mathrm{Tr}\left(P_0\ \rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_1\right) - 2\mathrm{Tr}\left(P_0\ \rho^{ij}_{AB}\ P_1\ \rho^{ij}_{AB}\right)\\
&\qquad - 2\mathrm{Tr}\left(\rho^{ij}_{AB}\ P_1\ P_0\ \rho^{ij}_{AB}\ P_0\right) - 2\mathrm{Tr}\left(P_1\ \rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_0\right)\\
&\qquad - 2\mathrm{Tr}\left(P_0\ \rho^{ij}_{AB}\ P_1\ \rho^{ij}_{AB}\right) - 2\mathrm{Tr}\left(P_0\ \rho^{ij}_{AB}\ P_1\ P_0\ \rho^{ij}_{AB}\right) - 2\mathrm{Tr}\left(P_1\ \rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_0\right) - 2\mathrm{Tr}\left(P_1\ \rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\right)\\
&\qquad - 2\mathrm{Tr}\left(P_0\ \rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_1\right) - 2\mathrm{Tr}\left(P_0\ \rho^{ij}_{AB}\ P_0\ P_1\ \rho^{ij}_{AB}\right)\\
&= \mathrm{Tr}\left(\rho^{ij}_{AB}\ P_0\ P_1\ \rho^{ij}_{AB}\right) + \mathrm{Tr}\left(P_0\ \rho^{ij}_{AB}\ \rho^{ij}_{AB}\ P_1\right) + \mathrm{Tr}\left(\rho^{ij}_{AB}\ P_1\ P_0\ \rho^{ij}_{AB}\right) + \mathrm{Tr}\left(P_1\ \rho^{ij}_{AB}\ \rho^{ij}_{AB}\ P_0\right) + 4\mathrm{Tr}\left(P_0\rho^{ij}_{AB}P_1\rho^{ij}_{AB}P_0\right)\\
&\qquad - 2\mathrm{Tr}\left(\rho^{ij}_{AB}\ P_0\ P_1\ \rho^{ij}_{AB}\ P_0\right) - 2\mathrm{Tr}\left(\rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_1\right) - 2\mathrm{Tr}\left(\rho^{ij}_{AB}\ P_1\ P_0\ \rho^{ij}_{AB}\ P_0\right)\\
&\qquad - 2\mathrm{Tr}\left(\rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_1\right) - 2\mathrm{Tr}\left(P_0\ \rho^{ij}_{AB}\ P_1\ P_0\ \rho^{ij}_{AB}\right) - 2\mathrm{Tr}\left(P_0\ \rho^{ij}_{AB}\ P_0\ P_1\ \rho^{ij}_{AB}\right)\\
&= \mathrm{Tr}\left(\ \rho^{ij}_{AB}\ \rho^{ij}_{AB}\ P_0\ P_1\right) + \mathrm{Tr}\left(P_0\ \rho^{ij}_{AB}\ \rho^{ij}_{AB}\ P_1\right) + \mathrm{Tr}\left(P_0\ \rho^{ij}_{AB}\ \rho^{ij}_{AB}\ P_1\right) + \mathrm{Tr}\left(\ \rho^{ij}_{AB}\ \rho^{ij}_{AB}\ P_0\ P_1\right) + 4\mathrm{Tr}\left(\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_1\right)\\
&\qquad - 2\mathrm{Tr}\left(\ \rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_0\ P_1\right) - 2\mathrm{Tr}\left(\rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_1\right) - 2\mathrm{Tr}\left(\ P_0\ \rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_1\right)\\
&\qquad - 2\mathrm{Tr}\left(\rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_1\right) - 2\mathrm{Tr}\left(\ P_0\ \rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_1\right) - 2\mathrm{Tr}\left(\ \rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_0\ P_1\right)\\
&= 2\ \mathrm{Tr}\left(\ \rho^{ij}_{AB}\ \rho^{ij}_{AB}\ P_0\ P_1\right) + 2\ \mathrm{Tr}\left(P_0\ \rho^{ij}_{AB}\ \rho^{ij}_{AB}\ P_1\right) - 4\mathrm{Tr}\left(\ \rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_0\ P_1\right) - 4\mathrm{Tr}\left(\ P_0\ \rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB}\ P_1\right)
\end{aligned}
\tag{F.93}
$$

The four trace term can't be solve as the earlier three cases. $P_1 = Q'(\phi_A^i) \times \mathbb{I} = Q'(\theta) = \frac{1}{2}\begin{pmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{pmatrix}$ is a Hermitian operator but not a Positive semi-definite operator(eigenvalues are $\pm\frac{1}{2}$).

$$\begin{aligned} &(\mathbb{I} - P_1) \succeq 0 \quad \left(\text{eigenvalues are } \frac{1}{2} \text{ and } \frac{3}{2} \right) \\ &\mathbb{I} \succeq P_1 \\ &\rho_{AB}^{ij}\, \rho_{AB}^{ij}\, P_0 \not\succeq \rho_{AB}^{ij}\, \rho_{AB}^{ij}\, P_0\, P_1 \quad \left(Y = \rho_{AB}^{ij}\, \rho_{AB}^{ij}\, P_0 \text{ does not commute with } P_1 \right) \end{aligned} \tag{F.94}$$

Since $\rho_{AB}^{ij}\, \rho_{AB}^{ij}\, P_0$ and P_1 does not commute we can't use the above partial ordering but since P_1 is hermitian we can have spectral double sided spectral bound.

$$\lambda_{min}(H)\, \mathbb{I} \preceq H \preceq \lambda_{max}(H)\, \mathbb{I} \tag{F.95}$$

Equivalently we can write,

$$\lambda_{min}(H)\, \mathrm{Tr}(Y) \leq \mathrm{Tr}(YH) \leq \lambda_{max}(H)\, \mathrm{Tr}(Y) \tag{F.96}$$

for $Y \succeq 0$. Alternatively, $(\mathbb{I} - P_1)$ is a Hermitian so we can the following spectral bound.

$$\begin{aligned} &\lambda_{min}((\mathbb{I} - P_1))\, \mathrm{Tr}(Y) \leq \mathrm{Tr}(Y(\mathbb{I} - P_1)) \leq \lambda_{max}((\mathbb{I} - P_1))\, \mathrm{Tr}(Y) \\ &\frac{1}{2}\mathrm{Tr}(Y) \leq \mathrm{Tr}(Y) - \mathrm{Tr}(YP_1) \leq \frac{3}{2}\mathrm{Tr}(Y) \end{aligned} \tag{F.97}$$

The above relation requires Y to be a positive semi definite matrix. This possess a serious flaw. As per the given setting this requires the complete knowledge of the density matrix and the projector operator which compromises the generality of the proof. Alternatively we can can be upper bounded using the Holder inequality [Wat18] as,

$$\begin{aligned} &\left| \mathrm{Tr}\left(\rho_{AB}^{ij}\, \rho_{AB}^{ij}\, P_0\, P_1 \right) \right| \\ &\leq \left\| \rho_{AB}^{ij}\, \rho_{AB}^{ij}\, P_0 \right\|_1 \left\| P_1 \right\|_\infty \\ &\leq \frac{1}{2} \left\| (\rho_{AB}^{ij})^2 \right\|_1 \left\| P_0 \right\|_\infty \\ &\leq \frac{1}{2} \mathrm{Tr}(\rho_{AB}^{ij}) \\ &\leq \frac{1}{2} \end{aligned} \tag{F.98}$$

Similarly,

$$\begin{aligned} &\left| \mathrm{Tr}\left(P_0\, \rho_{AB}^{ij}\, \rho_{AB}^{ij}\, P_1 \right) \right| \\ &\leq \left\| P_0\, \rho_{AB}^{ij}\, \rho_{AB}^{ij} \right\|_1 \left\| P_1 \right\|_\infty \\ &\leq \frac{1}{2} \end{aligned} \tag{F.99}$$

Also,

$$\begin{aligned} &\left| \mathrm{Tr}\left(\rho_{AB}^{ij}\, P_0\, \rho_{AB}^{ij}\, P_0\, P_1 \right) \right| \\ &\leq \|\rho_{AB}^{ij} P_0 \rho_{AB}^{ij}\|_1 \, \|P_0 P_1\|_\infty \\ &\leq \|\rho_{AB}^{ij}\, P_0\|_2 \, \|\rho_{AB}^{ij}\|_2 \, \|P_0 P_1\|_\infty \\ &\leq \|\rho_{AB}^{ij}\|_2 \, \|\rho_{AB}^{ij}\|_2 \, \|P_0\|_\infty \, \|P_1\|_\infty \\ &= \|\rho_{AB}^{ij}\|_2^2 \, \|P_0\|_\infty \, \|P_1\|_\infty = \mathrm{Tr}\big((\rho_{AB}^{ij})^2\big) \cdot 1 \cdot \frac{1}{2} \\ &\leq \frac{1}{2}. \end{aligned} \tag{F.100}$$

Finally for $\mathrm{Tr}\big(P_0\rho^{ij}_{AB}P_0\rho^{ij}_{AB}\,P_1\big)$

$$\begin{aligned}
&\Big|\mathrm{Tr}\big(P_0\rho^{ij}_{AB}P_0\rho^{ij}_{AB}\,P_1\big)\Big| \\
&\le \|P_0\rho^{ij}_{AB}P_0\rho^{ij}_{AB}\|_1\,\|P_1\|_\infty \\
&\le \|P_0\rho^{ij}_{AB}\|_2\,\|P_0\rho^{ij}_{AB}\|_2\,\|P_1\|_\infty \\
&\le \|\rho^{ij}_{AB}\|_2^2\,\|P_0\|_\infty^2\,\|P_1\|_\infty \\
&\le \mathrm{Tr}\big((\rho^{ij}_{AB})^2\big)\cdot 1\cdot\frac{1}{2} \\
&\le \frac{1}{2}.
\end{aligned} \tag{F.101}$$

Coming back to evaluating $\mathrm{Tr}(B_1)$

$$\mathrm{Tr}(B_1) = 2x + 2y - 4u - 4v, \quad |x|,|y|,|u|,|v| \le \tfrac{1}{2}, \tag{F.102}$$

the worst-case bounds are

$$\max\mathrm{Tr}(B_1) = 2\Big(\tfrac{1}{2}\Big) + 2\Big(\tfrac{1}{2}\Big) - 4\Big(-\tfrac{1}{2}\Big) - 4\Big(-\tfrac{1}{2}\Big) = 6, \tag{F.103}$$

$$\min\mathrm{Tr}(B_1) = 2\Big(-\tfrac{1}{2}\Big) + 2\Big(-\tfrac{1}{2}\Big) - 4\Big(\tfrac{1}{2}\Big) - 4\Big(\tfrac{1}{2}\Big) = -6. \tag{F.104}$$

So,

$$|\mathrm{Tr}(B_1)| \le 6. \tag{F.105}$$

A tighter bound exists but for that an additional assumption is needed.

$$\mathrm{supp}(\rho) \subseteq \mathrm{ran}(P_0) \tag{F.106}$$

Support of ρ is the span of all eigenvectors with non-zero eigenvalues. Range of the projector P_0 is span of all nonzero columns of P_0. Following two relation can directly be followed from the given assumption.

$$\begin{aligned}
&\rho|\psi\rangle = \lambda|\psi\rangle \quad \forall \quad (\lambda > 0) \qquad \Longrightarrow\ \psi \in \mathrm{ran}(P_0) \\
&P_0|\psi\rangle = |\psi\rangle \quad \forall|\psi\rangle \in \mathrm{supp}(\rho) \quad (P_0 \text{ act as identity operator to } \psi)
\end{aligned} \tag{F.107}$$

For any vector $|\phi\rangle$, decompose $\rho|\phi\rangle$:

$$\begin{aligned}
\rho|\phi\rangle &= \sum_{\lambda_k>0}\lambda_k\langle\psi_k|\phi\rangle|\psi_k\rangle + \sum_{\lambda_k=0} 0\cdot\langle\psi_k|\phi\rangle|\psi_k\rangle \\
&= \sum_k \lambda_k\langle\psi_k|\phi\rangle|\psi_k\rangle, \quad \lambda_k > 0, \quad |\psi_k\rangle \in \mathrm{supp}(\rho) \\
P(\rho|\phi\rangle) &= P\left(\sum_k \lambda_k\langle\psi_k|\phi\rangle|\psi_k\rangle\right) \\
&= \sum_k \lambda_k\langle\psi_k|\phi\rangle P|\psi_k\rangle \\
&= \sum_k \lambda_k\langle\psi_k|\phi\rangle|\psi_k\rangle \\
&= \rho|\phi\rangle \quad \forall|\phi\rangle \\
P\rho &= \rho
\end{aligned} \tag{F.108}$$

$$\begin{aligned}
\rho\dagger &= (P\rho)\dagger \\
&= \rho^\dagger P^\dagger \\
&= \rho P \\
&= \rho
\end{aligned} \tag{F.109}$$

From the last two equation,

$$\rho = P\rho = \rho P = P\rho P \tag{F.110}$$

Now using the above relation $\mathrm{Tr}(B_1)$ bound can be tighten using the third term $\mathrm{Tr}(\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_0P_1)$.

$$\begin{aligned}
&\mathrm{Tr}(\rho^{ij}_{AB}P_0\rho^{ij}_{AB}P_0P_1)\\
&= \mathrm{Tr}((P_0\rho^{ij}_{AB}P_0)P_0(P_0\rho^{ij}_{AB}P_0)P_1) \quad (\text{substitute } \rho^{ij}_{AB} = P_0\rho^{ij}_{AB}P_0)\\
&= \mathrm{Tr}(P_0(\rho^{ij}_{AB}P_0\rho^{ij}_{AB})P_0P_1)\\
&= \mathrm{Tr}((\rho^{ij}_{AB}P_0\rho^{ij}_{AB})P_0P_1P_0)\\
&= \mathrm{Tr}((\rho^{ij}_{AB}P_0\rho^{ij}_{AB})(P_0P_1P_0)).
\end{aligned} \tag{F.111}$$

Now, $P_0 = Q(\theta)$, $P_1 = Q'(\theta)$ we have,

$$\begin{aligned}
&P_0^2 = P_0\\
&P_0P_0 = P_0\\
&P_1P_0 + P_0P_1 = P_1\\
&P_0P_1P_0 + P_0P_0P_1 = P_0P_1\\
&P_0P_1P_0 + P_0P_1 = P_0P_1\\
&P_0P_1P_0 = 0
\end{aligned} \tag{F.112}$$

thus,

$$\mathrm{Tr}(B_1) = 2x + 2y - 4v, \quad |x|, |y|, |v| \leq \tfrac{1}{2} \quad \Rightarrow \quad |\mathrm{Tr}(B_1)| \leq 4. \tag{F.113}$$

Since $\epsilon_0 > 0$ is a very small quantity, $\epsilon_0^2 << \epsilon_0$ we have $B = B_0 + \epsilon_0 B_1 + \mathcal{O}(\epsilon_0^2)$. Now, from (F.85),

$$\begin{aligned}
&\left|h(\lambda, \phi^i_A, \rho^{ij}_{AB}) - h(\lambda, \phi^i_A + \epsilon_0, \rho^{ij}_{AB})\right|\\
&\leq (1-\lambda)\,[\mathrm{Tr}(A+B)]\\
&\leq (1-\lambda)\,[\mathrm{Tr}(A_0) + \mathrm{Tr}(A_1) + \mathrm{Tr}(B_0) + \epsilon_0\mathrm{Tr}(B_1)]\\
&\leq (1-\lambda)[1 + 1 + 2 + \epsilon_0 4]\\
&\leq (1-\lambda)[4 + 4\epsilon_0 + \mathcal{O}(\epsilon_0^2)]
\end{aligned} \tag{F.114}$$

The above quantity is $\mathcal{O}(1)$ to make it $\mathcal{O}(\epsilon)$ we can use the result $\rho = P\rho = \rho P = P\rho P$.

$$\begin{aligned}
\mathrm{Tr}(A_0) &= \mathrm{Tr}(\ \rho^{ij}_{AB}\ (\mathbb{I} - P_0)\ \rho^{ij}_{AB})\\
&= \mathrm{Tr}(\ (\mathbb{I} - P_0)\ \rho^{ij}_{AB})\\
&= \mathrm{Tr}(\ (\mathbb{I} - P_0)P_0\ \rho^{ij}_{AB}\ P_0)\\
&= \mathrm{Tr}(P_0\ \rho^{ij}_{AB}\ P_0 - P_0^2\ \rho^{ij}_{AB}\ P_0)\\
&= 0
\end{aligned} \tag{F.115}$$

Similarly,

$$\begin{aligned}
\mathrm{Tr}(B_0) &= 2(\mathrm{Tr}(P_0\ \rho^{ij}_{AB}\ \rho^{ij}_{AB}) - \mathrm{Tr}((P_0\rho^{ij}_{AB}P_0\rho^{ij}_{AB})))\\
&= 2(\mathrm{Tr}(P_0\rho^{ij}_{AB}\ P_0\rho^{ij}_{AB}P_0) - \mathrm{Tr}((P_0\ \rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB})))\\
&= 2(\mathrm{Tr}(P_0\rho^{ij}_{AB}P_0\rho^{ij}_{AB}) - \mathrm{Tr}((P_0\ \rho^{ij}_{AB}\ P_0\ \rho^{ij}_{AB})))\\
&= 0
\end{aligned} \tag{F.116}$$

where $P_0 = Q(\phi^i_A) \otimes \mathbb{I}$ Thus finally,

$$\boxed{\begin{aligned}
&\left|h(\lambda, \phi^i_A, \rho^{ij}_{AB}) - h(\lambda, \phi^i_A + \epsilon_0, \rho^{ij}_{AB})\right|\\
&\leq (1-\lambda)[4\epsilon_0 + \mathcal{O}(\epsilon_0^2)]
\end{aligned}} \tag{F.117}$$

b. Modified optimization problem incorporating the change in Alice's angle

Now we are at a point where we can lower bound the maximal error in k^{th} segment centered around $\phi^i_{A_k}$. Using Eqs. (F.117) and (F.57) in the modified objective function given in (F.19), we get the modified optimization problem as,

$$\boxed{\begin{aligned} n^*(S_{ij}) \geq \inf \quad & \lambda \left\|\rho^{ij}_{AB} - \Lambda_0[\rho^{ij}_{AB}]\right\|^2_F + (1-\lambda)\left\|\rho^{ij}_{AB} - \Lambda_1[\rho^{ij}_{AB}]\right\|^2_F \\ & + \frac{\mu}{2}\|\rho^{ij}_{AB}\|^2_F \; - 2(1-\lambda)\epsilon_0 \\ \text{s.t.} \quad & \mathrm{Tr}\left(\rho^{ij}_{AB}\, CHSH(\phi^i_A, \phi^j_B)\right) = S_{ij} \; - 2\epsilon_0 \\ & \phi^i_A, \phi^j_B \in [0, \pi/2], \\ & \rho^{ij}_{AB} \succeq 0 \\ & \mathrm{Tr}(\rho^{ij}_{AB}) = 1 \end{aligned}} \tag{F.118}$$

By averaging multiple instances or scenarios in the optimisation 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 $\epsilon_0 = 4$ to $\epsilon'_0 = 2$. This reduction occurs because averaging tends to smooth out extreme values. Given that the optimal value lies within a symmetric interval of $\phi^i_{A_k} \pm \epsilon_0$, averaging over multiple such intervals can lead to a smaller overall range of uncertainty for the optimal value.

Appendix G: Creating a convex hull from all two qubit functions $C^*(S)$ for all $S \in (2, 2\sqrt{2}]$

In Eq. (C.22), it had been shown that the function $C^*(S)$ can be lower bounded as an integral of the function of $C^*(S)$ over two qubit blocks, $C^*_{\mathbb{C}^{4\times 4}}(S')$, over the range $S \in (2, 2\sqrt{2}]$, i.e., $C^*(S) \geq \int_{S'=2}^{2\sqrt{2}} \eta(dS') \cdot C^*_{\mathbb{C}^{4\times 4}}(S')$ such that $\eta([2, 2\sqrt{2}]) \leq 1, \quad \eta \geq 0, \quad$ and $\int_{S'=2}^{2\sqrt{2}} \eta(dS')S' = S$. Moreover, it has been showed that each such two qubit function $C^*_{\mathbb{C}^{4\times 4}}(S')$ can further be lower bounded through a strongly convex objective function $n^*(S_{ij})$, using a modified Pinsker's inequality[SGP+21] and the results from Eqs. (D.6) (F.19). Now given values of $C^*_{\mathbb{C}^{4\times 4}}(S)$ for all $S \in (2, 2\sqrt{2}]$, one needs a convex function, say, $\overline{C}(S)$[SGP+21] that would essentially give,

$$C^*_{\mathbb{C}^{4\times 4}}(S_{ij}) \geq \overline{C}(S_{ij}) \tag{G.1}$$

for ij^{th} block. Thus incorporating this into (C.22) as,

$$\begin{aligned} C^*(S) \geq & \int_{S'=2}^{2\sqrt{2}} \eta(dS') \cdot \overline{C}(S') \\ s.t \quad & \eta([2, 2\sqrt{2}]) \leq 1 \\ & \eta \geq 0 \\ & \int_{S'=2}^{2\sqrt{2}} \eta(dS')S' = S \end{aligned} \tag{G.2}$$

gives

$$C^*(S) \geq \overline{C}(S) \tag{G.3}$$

The final task is to prove the existence of one such function $\overline{C}(S)$ that lower bounds $C^*(S)$.

Theorem 6: $n^*(S)$ is a valid lower bound for $C^*(S)$.

Proof:

a. Strong Convexity of the Objective Function

The optimization problem defining $n^*(S)$ is given as,

$$\begin{aligned} n^*(S_{ij}) &= \inf\left\{\lambda \left\|\rho_{AB}^{ij} - \Lambda_0\left[\rho_{AB}^{ij}\right]\right\|_F^2 + (1-\lambda)\left\|\rho_{AB}^{ij} - \Lambda_1\left[\rho_{AB}^{ij}\right]\right\|_F^2 + \frac{\mu}{2}\left\|\rho_{AB}^{ij}\right\|_F^2\right\} \\ s.t.\ \mathrm{Tr}&\left(\rho_{AB}^{ij}\, CHSH(\phi_A^i, \phi_B^j)\right) = S_{ij}. \end{aligned} \tag{G.4}$$

Now, the given objective function involving the Frobenius norm $\|\cdot\|_F^2$ is strongly convex in ρ. The term $\frac{\mu}{2}\left\|\rho_{AB}^{ij}\right\|_F^2$ introduces μ-strong convexity, and the linear constraints of the optimization problem preserve the strong convexity on the feasible set. For fixed S_{ij}, the objective function is strongly convex in ρ. The parameterized problem's solution function $n^*(S_{ij})$ inherits convexity. Moreover, strong convexity implies quadratic dependence on the perturbations in S_{ij}. Thus, $n^*(S)$ is strongly convex in S for $S \in (2, 2\sqrt{2}]$

b. Application of Jensen's inequality

For the measure η with,

$$\int_2^{2\sqrt{2}} \eta(dS') = 1, \quad \eta \geq 0, \quad \int_2^{2\sqrt{2}} \eta(dS')S' = S \tag{G.5}$$

Now application of Jensen's inequality for strongly convex functions yields,

$$\int_2^{2\sqrt{2}} \eta(dS')n^*(S') \geq n^*\left(\int_2^{2\sqrt{2}} \eta(dS')S'\right) + \frac{\mu}{2}\,\mathrm{Var}_\eta(S') \tag{G.6}$$

Since $\mathrm{Var}_\eta(S') \geq 0$,

$$\int_2^{2\sqrt{2}} \eta(dS')n^*(S') \geq n^*(S). \tag{G.7}$$

Substituting $\overline{C}(S') = n^*(S')$ into the original inequality, we get,

$$C^*(S) \geq \int_2^{2\sqrt{2}} \eta(dS')n^*(S') \geq n^*(S) \tag{G.8}$$

Therefore, $\overline{C}(S) = n^*(S)$ is a valid convex lower bound. □

[ABG$^+$07] 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.

[AL70] Huzihiro Araki and Elliott H. Lieb. Entropy inequalities. *Communications In Mathematical Physics*, 18(2):160–170, June 1970.

[BCC$^+$10] 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.

[BCP$^+$14] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, Apr 2014.

[BS10] 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.

[BTN01] 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.

[BV04] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[CBTW17] Patrick J. Coles, Mario Berta, Marco Tomamichel, and Stephanie Wehner. Entropic uncertainty relations and their applications. *Rev. Mod. Phys.*, 89:015002, Feb 2017.

[CM11] Marcos Curty and Tobias Moroder. Heralded-qubit amplifiers for practical device-independent quantum key distribution. *Phys. Rev. A*, 84:010304, Jul 2011.

[CS12] Adán Cabello and Fabio Sciarrino. Loophole-free bell test based on local precertification of photon's presence. *Phys. Rev. X*, 2:021010, Jun 2012.

[GPS10] 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.

[Hol19] Alexander S. Holevo. *Quantum Systems, Channels, Information*. De Gruyter, Berlin, Boston, 2019.

[HST$^+$20] 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.

[KMS$^+$20] 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.

[LPT$^+$13] 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.

[Mam24] Z.Z. Mammeri. *Cryptography: Algorithms, Protocols, and Standards for Computer Security*. Wiley, 2024.

[MSMG$^+$16] 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.

[NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[PAB$^+$09] 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.

[PP08] K. B. Petersen and M. S. Pedersen. The matrix cookbook, October 2008. Version 20081110.

[Ras15] Ioan Rasa. Concavity of some entropies, 2015.

[Ren06] Renato Renner. Security of quantum key distribution, 2006.

[SGP$^+$21] 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.

[Sik12] Jamie William Jonathon Sikora. Analyzing quantum cryptographic protocols using optimization techniques. 2012.

[TBB08] B.S. Thomson, A.M. Bruckner, and J.B. Bruckner. *Elementary Real Analysis*. Number v. 1. www.classicalrealanalysis.com., 2008.

[Wat18] J. Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.

[Wil13] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2013.

[XZZP22] 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.