LUND UNIVERSITY

FACULTY OF SCIENCE

Detection Loophole in Bell Tests: Post-Selection with Outcome-Dependent Efficiencies

Aria Ghanbari

Thesis submitted for the degree of Bachelor of Science
Project duration: 5 months

Supervised by Armin Tavakoli

Department of Physics
Division of Mathematical Physics
May 2026

Abstract

Bell tests provide a suitable framework for investigating whether the correlations predicted by quantum mechanics can be explained by local hidden-variable (LHV) theories. These tests are susceptible to loopholes, among which the detection loophole has attracted considerable attention. This thesis investigates the detection loophole in Bell tests, with a particular focus on the role of post-selection in the presence of outcome-dependent detection efficiencies. The Bell violations obtained from both quantum-mechanical and local hidden-variable models are quantified in post-selected scenarios. The detection efficiencies associated with outcomes 0 and 1 are denoted by η_0 and η_1, respectively. We encode the asymmetry in the detection efficiencies using the parameter $R = \frac{\eta_0 - \eta_1}{\eta_0 + \eta_1}$. We show that post-selected LHV models can produce apparent violations of the Clauser–Horne–Shimony–Holt (CHSH) Bell inequality for $R \neq 0$. For sufficiently large asymmetry ($R \gtrsim 0.67$), the maximal CHSH violation obtained from the post-selected LHV model exceeds that of a two-qubit quantum model. This quantity follows the numerically observed relation $\mathcal{W}_{\text{CHSH}} \approx 2R + 2$, while the two-qubit quantum model reproduces an approximately quadratic dependence on R. This suggests that the behavior of the LHV model is directly associated with the asymmetry between the detection efficiencies, whereas the quantum behavior depends on a joint contribution of both efficiencies. Later, a post-selected two-ququart quantum model was optimized, which scores a CHSH violation consistently higher than or equal to that of a post-selected LHV model. The behavior of both the two-qubit quantum model and the LHV model was also studied for the I_{3322} Bell inequality. This work also explores the relation between a post-selected LHV model and a measurement-dependent hidden-variable (MDHV) model. The analysis includes a numerical calculation of the measurement independence parameter M as a function of R, suggesting a numerically observed relation of $M \approx R$.

Keywords: detection loophole, post-selection, detector inefficiencies, Bell inequalities, CHSH, local hidden variable models, ququarts, qubits, local, nonlocal

Acknowledgements

I would like to first thank my supervisor, Armin Tavakoli. Since the beginning of this work, he has consistently prioritized my interests and helped me identify the concepts I enjoy most while learning them effectively. He welcomed me into his group, introduced me to this thesis topic, and gave me freedom to explore and learn independently, while remaining patient with my questions and providing effective suggestions.

I am also grateful to my classmate and friend Amro, who previously completed his thesis with Armin. He was always willing to discuss my thesis and answer my questions. He helped me to better recognize the direction of my work. He also carefully reviewed parts of my thesis, which I greatly appreciate. I am thankful to my friend Olof, with whom I had conversations about Bell inequalities. Trying to explain these concepts during our discussions often revealed gaps in my understanding and deepened my knowledge.

I want to express my deepest gratitude to my aunt and my mother. My aunt supported me throughout my studies, both financially and emotionally. She always stood by me through every challenge I faced. Without her support, I would not have been able to continue my studies, and I owe her more than words can express. My mother, despite the distance between us, always provided me with emotional support and encouragement, and has always believed in me, which strengthened my determination throughout this journey.

Lastly, I would like to thank Armin's group and the people in my whole office for welcoming me into their academic community, especially Oskar, with whom I had many constructive and enjoyable conversations throughout my work.

ChatGPT (OpenAI) was used as a support tool to refine grammar, provide linguistic suggestions, and polish the presentation of scientific work.

Abbreviations

LHV: Local hidden variable model
POVM: Positive operator-valued measure
CHSH: Clauser–Horne–Shimony–Holt inequality
MDHV: Measurement-dependent hidden-variable model

Contents

1 Introduction

Since the dawn of quantum mechanics, the interpretation of quantum phenomena has remained a central challenge in physics. Although its mathematical formalism has been remarkably successful, the conceptual foundations of the theory are yet to be properly understood, as they often conflict with classical intuitions.

One of the most important contexts in which the contrast between classical models and quantum mechanics becomes evident is provided by Bell inequalities. These inequalities offer a rigorous framework for the study of quantum correlations, and they allow for a clear comparison between the predictions of quantum theory and those of classical theories. As quantum mechanics continued to raise questions about classical intuitions, many physicists aimed to develop a more complete theory that resolves these conceptual problems. Among them, Albert Einstein questioned whether quantum mechanics respects the principle of locality and realism. [1]

In 1935, Albert Einstein, with two other physicists, Boris Podolsky and Nathan Rosen, published a paper titled "Can Quantum-mechanical Description of Physical Reality Be Considered Complete" [1]. In this paper, they raised a paradox, now known as the EPR Paradox, providing a quantum-mechanical scenario in which two particles are entangled, and measuring one particle allows prediction of the outcome of measurements on the other, even when the particles are spatially separated. They concluded that quantum mechanics cannot respect both locality and realism simultaneously. As a result, they proposed that quantum mechanics might be incomplete, and that a more complete description can be a local hidden variable (LHV) theory, in which measurement outcomes are determined by additional, unobserved parameters while preserving locality.

A breakthrough in this area was made by John Stewart Bell. In 1967, he introduced the Bell inequalities, building upon the EPR Paradox. Bell then showed that quantum physics predicts correlations that violate these inequalities [2], showing that no local hidden variable theory can reproduce all quantum predictions. Although quantum mechanics does not allow faster-than-light communication, it violates the assumption of local realism. For this reason, quantum mechanics is often described as "nonlocal". This refers to the failure of local realism, rather than a violation of relativistic locality. Quantum nonlocality has also been successfully tested in experiments, where many laboratory experiments [3, 4, 5] suggest that nature rules out local realism.

Bell inequalities in their simplest case build on a scenario in which an entangled quantum state is shared between two parties, often called Alice and Bob. Then, Alice and Bob each perform their independent measurements, denoted x and y, respectively, on the shared quantum state. These measurements lead to outcomes a for Alice, and b for Bob. The inequality can then be constructed using a suitable combination of the probabilities of observing each outcome. Although quantum predictions violate these inequalities, they respect an upper bound called Tsirelson's bound.

Measurements in Bell inequality experiments are subject to inefficiencies, meaning some particles go undetected. If the detected particles do not constitute a fair sample of all emitted particles, any statistical analysis, in particular Bell tests, could be regarded as unreliable. This issue is known as the detection loophole. When only the detected events are included in the analysis (a procedure known as post-selection), it has been shown that an LHV model can reproduce apparent violations of Bell inequalities in the post-selected data [6]. In this thesis, we further investigate the post-selected data under different assumptions. An earlier work in this area is the paper "Bell and EPR experiments with signaling data" [7], where it is argued that such inefficiencies can lead to an induced inflation of the observed correlations relative to Tsirelson's bound. In addition, [8] discusses how outcome and measurement dependent efficiencies can arise in photonic Bell experiments and how they may affect the observed violation of the Bell inequality.

Initially, we assumed that the statistics of the underlying ensemble are governed by quantum mechanics.

[1]In this context, locality is understood as no information can propagate faster than the speed of light in vacuum, while realism implies that the physical properties exist independently of measurements.

We then examined how much the maximal Bell parameter in the post-selected analysis can be shifted by introducing outcome-dependent efficiencies, η_0 and η_1. This is done by maximizing the Bell parameter over all possible measurements for a two-qubit system (and later extending the work to higher-dimensional systems). We subsequently switched to an LHV model to determine the amount of Bell inequality violation observed after post-selection. We performed these analyses for both the CHSH and I_{3322} Bell inequalities.

The CHSH violation obtained from a two-qubit quantum model cannot consistently surpass that of an LHV model in the presence of inefficiencies. Therefore, the work is extended to a two-ququart quantum model. The CHSH violation achieved from this model does not fall below that of an LHV model. Additionally, an LHV model with post-selection behaves similarly to a measurement-dependent hidden-variable (MDHV) model [2]. Therefore, we investigated the numerical relationship between measurement independence and detection efficiencies.

The technological relevance of this work is also significant. Quantum information plays a central role in a range of growing technologies, such as quantum computing and quantum communication [9]. In this thesis, we develop quantum information theory beyond idealized assumptions for measurement systems. By incorporating these experimental inefficiencies, our models provide a more realistic framework for understanding and interpreting quantum experiments. This study touches on one of the key assumptions of device-independent protocols [10], where Bell violations are used to certify nonlocality, in the absence of significant experimental loopholes.

2 Theoretical Framework

Quantum mechanics provides us with a rigorous mathematical framework to treat physical systems. This framework revolves around a vector denoted as state vector $|\psi\rangle$, describing an isolated system. In a broader context, this vector maps complex numbers to different system configurations.

2.1 Wavefunction and Density Operator

Every isolated system in quantum mechanics has an associated Hilbert space, where the normalized state vector $|\psi\rangle$ lives in. Hilbert space is a complex vector space equipped with an inner product. Let $\{|u_i\rangle\}$ be an orthonormal basis of the Hilbert space $\mathcal{H}$, then any state $|\psi\rangle \in \mathcal{H}$ can be expressed as

$$|\psi\rangle = \sum_i c_i |u_i\rangle \text{ such that } \quad \langle\psi|\psi\rangle = 1, \tag{2.1}$$

where $c_i \in \mathbb{C}$. Although the state vector is a powerful tool for dealing with isolated systems, we require a more general formalism to extend the theory beyond isolated systems. This is done with the density operator formalism ρ. For a system with a known quantum state $|\psi\rangle$, or otherwise known as a pure state, the density operator is $\rho = |\psi\rangle\langle\psi|$. However, in a more realistic situation, the quantum state is not completely known, and instead we know that the system is in one of a number of states $|\psi_i\rangle$ with the probability of $p_i \geq 0$, the density operator is then given by [9]

$$\rho \equiv \sum_i p_i\, |\psi_i\rangle\langle\psi_i|, \quad \text{where} \quad \sum_i p_i = 1 \text{ and } p_i \geq 0 \text{ for all i.} \tag{2.2}$$

[2] A measurement-dependent hidden-variable model is a model in which the hidden variable distribution depends on the measurement settings, or equivalently, the measurement settings are statistically correlated with the hidden variables.

We now show that the density operator satisfies $\mathrm{Tr}(\rho) = 1$ and is positive semi-definite, $\rho \geq 0$. To prove the first property, we write

$$\mathrm{Tr}(\rho) = \mathrm{Tr}(\sum_i p_i |\psi_i\rangle\langle\psi_i|) = \sum_i p_i \mathrm{Tr}(|\psi_i\rangle\langle\psi_i|) = \sum_i p_i \langle\psi_i|\psi_i\rangle = \sum_i p_i = 1. \tag{2.3}$$

In order to prove that the density operator is positive semi-definite, we verify that it is hermitian and that $\langle\phi|\rho|\phi\rangle > 0$ for every $|\phi\rangle \in \mathcal{H} \setminus \{0\}$. Hermiticity follows immediately

$$\rho^\dagger = (\sum_i p_i |\psi_i\rangle\langle\psi_i|)^\dagger = \sum_i p_i (|\psi_i\rangle\langle\psi_i|)^\dagger = \sum_i p_i |\psi_i\rangle\langle\psi_i| = \rho. \tag{2.4}$$

Moreover,

$$\langle\phi|\rho|\phi\rangle = \langle\phi| \sum_i p_i |\psi_i\rangle\langle\psi_i||\phi\rangle = \sum_i p_i \langle\phi|\psi_i\rangle\langle\psi_i|\phi\rangle = \sum_i p_i ||\langle\phi|\psi_i\rangle||^2 \geq 0. \tag{2.5}$$

The discussion can be extended to a system composed of n independent and uncorrelated subsystems, denoted by index i, each with density operator ρ_i. In this scenario, the composite density operator is given by the tensor product between individual density operators ($\rho = \rho_1 \otimes \rho_2 \otimes \ldots \otimes \rho_n$). Luckily, the mathematical formalism is well-suited to reconstructing the individual density operators, i.e., the reduced density operators. For a bipartite system given ρ^{AB}, the reduced density operators can be written as

$$\rho^A = \mathrm{Tr}_\mathrm{B}(\rho^{AB}), \quad \text{and} \quad \rho^B = \mathrm{Tr}_\mathrm{A}(\rho^{AB}). \tag{2.6}$$

2.2 Quantum Measurements

Measurements play a central role in the construction of quantum mechanics. A general quantum mechanical measurement is described by a set of $\{M_a\}$ measurement operators, where a corresponds to the measurement outcome that can occur in the experiment. Then the probability of obtaining outcome a can be expressed as

$$p(a) = \langle\psi|M_a^\dagger M_a|\psi\rangle. \tag{2.7}$$

Furthermore, the measurement operators satisfy the completeness relation $\sum_a M_a^\dagger M_a = I$. We can define a set of operators $\{E_a\}$ such that:

$$E_a = M_a^\dagger M_a. \tag{2.8}$$

Then the probability $p(a)$ becomes

$$p(a) = \langle\psi|E_a|\psi\rangle. \tag{2.9}$$

The set $\{E_a\}$ is called POVM (Positive Operator-Valued Measure). As the name suggests, its elements satisfy the following conditions

- $\sum_a E_a = I$ (completeness relation),
- $E_a \geq 0$ (positivity condition).

Projective measurements are a special type of measurement operators that are mutually orthogonal, satisfying $\Pi_a \Pi_b = \delta_{ab}\Pi_a$. These operators project the state vector onto an orthogonal basis $\{|a\rangle\}$, and can be written as $\Pi_a = |a\rangle\langle a|$. It can be shown that the projective operators are Hermitian, i.e., $\Pi_a^\dagger = \Pi_a$. In this set-up, the projective measurement is represented by an observable A, which, via the spectral decomposition [11], can be written as

$$A = \sum_a a\Pi_a = \sum_a a|a\rangle\langle a|. \tag{2.10}$$

The measurement outcomes correspond to the eigenvalues of A, which is why A is called an observable. One of the important set of observables is the Pauli operators $\sigma_X, \sigma_Y, \sigma_Z$ (and the identity I), given by

$$\sigma_X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.11}$$

2.3 Qubits and Bloch Sphere

A qubit (quantum bit) is one of the simplest quantum systems, consisting of two basis states $|0\rangle$ and $|1\rangle$. Its Hilbert space is therefore two-dimensional, and a general qubit state is a superposition of these basis states

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad \text{where} \quad \alpha, \beta \in \mathbb{C} \ \text{ and } \ \alpha^2 + \beta^2 = 1. \tag{2.12}$$

Consequently, the state is fully described by two complex numbers α, β. Since global phases are physically irrelevant [11], the state can alternatively be parametrized using spherical coordinates θ and ϕ as

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle, \quad \text{where} \quad \theta \in [0, \pi] \text{ and } \phi \in [0, 2\pi). \tag{2.13}$$

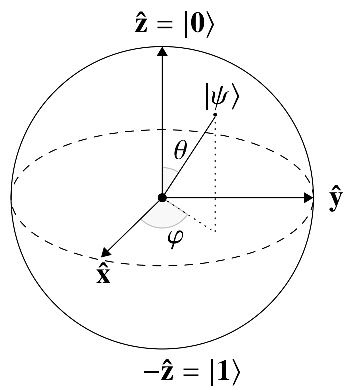


Figure 2.1: Bloch sphere tik Z code adapted from TEX Stack Exchange [12].

This structure provides the convenience of representing the quantum state on a unit sphere, known as the Bloch sphere, using the polar and azimuthal angles θ and ϕ, as shown in **Figure 2.1**. It can be shown that the coordinate of the quantum state $(\sin(\theta)\cos(\phi), \sin(\theta)\sin(\phi), \cos(\theta))$ corresponds to the expectation values of the Pauli operators $\sigma_X, \sigma_Y, \sigma_Z$, respectively [11]. In other words, if the quantum state $|\psi\rangle$ is represented by the unit vector $\hat{s}_\psi$ on the Bloch sphere, we have

$$\hat{s}_\psi = (\sin(\theta)\cos(\phi), \sin(\theta)\sin(\phi), \cos(\theta)) = (\langle\sigma_X\rangle, \langle\sigma_Y\rangle, \langle\sigma_Z\rangle). \tag{2.14}$$

It is straightforward to show that any pair of orthogonal projectors $\{\Pi_\pm\}$ can be represented on the Bloch sphere with antipodal vectors [11]. Defining $\boldsymbol{\sigma} = (\sigma_X, \sigma_Y, \sigma_Z)$, any density state ρ takes the form of

$$\rho = \frac{1}{2}\left(I + \mathbf{s}\cdot\boldsymbol{\sigma}\right), \tag{2.15}$$

where $\mathbf{s} = |s|(\sin(\theta)\cos(\phi), \sin(\theta)\sin(\phi), \cos(\theta))$ is called the Bloch vector and $||\mathbf{s}|| \leq 1$. A vector of length

1 corresponds to pure states (the surface of the sphere), while mixed states correspond to $||\vec{s}|| < 1$ (points inside the sphere). The density matrix formalism can be appropriately extended to a two-bipartite quantum system.

We consider a bipartite system [1] consisting of two parties, Alice and Bob, with associated Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Then the Hilbert space for the joint system is given by $\mathcal{H}_A \otimes \mathcal{H}_B$. Thus, the joint density matrix is a linear operator on this space, i.e.

$$\rho : \mathcal{H}_A \otimes \mathcal{H}_B \to \mathcal{H}_A \otimes \mathcal{H}_B, \tag{2.16}$$

where for two qubits $\mathcal{H}_A \otimes \mathcal{H}_B \cong \mathbb{C}^4$. Following this, a general two-qubit density operator can be expressed as [13]

$$\rho = \frac{1}{4}\left(I \otimes I + \mathbf{r} \cdot \boldsymbol{\sigma} \otimes I + I \otimes \mathbf{s} \cdot \boldsymbol{\sigma} + \sum_{i,j=1}^{3} T_{ij}\, \sigma_i \otimes \sigma_j \right). \tag{2.17}$$

Here T_{ij} are the components of the correlation matrix, and $\mathbf{r}$ and $\mathbf{s}$ are the reduced Bloch vectors for subsystems A and B, respectively, defined by

$$T_{ij} = \mathrm{tr}\left[\rho\,(\sigma_i \otimes \sigma_j)\right], \qquad \mathbf{r} = \mathrm{Tr}_B[\rho\,(\boldsymbol{\sigma} \otimes I)], \quad \mathbf{s} = \mathrm{Tr}_A[\rho\,(I \otimes \boldsymbol{\sigma})]. \tag{2.18}$$

This thesis mostly focuses on scenarios in which two parties, Alice and Bob, perform measurements labeled by $x, y \in \{0, 1\}$ and obtain outcomes $a, b \in \{0, 1\}$. The corresponding measurement operators are defined by projective operators $\{A_{a|x}\}$ and $\{B_{b|y}\}$. The Bloch sphere representations are

$$A_{a|x} = \frac{1}{2}\Big(I + (-1)^a\, \mathbf{a}_x \cdot \boldsymbol{\sigma}\Big), \qquad B_{b|y} = \frac{1}{2}\Big(I + (-1)^b\, \mathbf{b}_y \cdot \boldsymbol{\sigma}\Big). \tag{2.19}$$

Here, $\mathbf{a}_x, \mathbf{b}_y \in \mathbb{R}^3$ are unit vectors in the Bloch sphere that specify the direction of the measurement. Using the expression for $\{A_{a|x}\}$ and $\{B_{b|y}\}$ in addition to the general two-qubit density operator expression, one can derive the following relation for the joint probability $P(a, b|x, y)$

$$P(a, b|x, y) = \mathrm{tr}\big((A_{a|x} \otimes B_{b|y})\,\rho\big) = \frac{1}{4}\Big(1 + (-1)^a\, \mathbf{a}_x \cdot \mathbf{r} + (-1)^b\, \mathbf{b}_y \cdot \mathbf{s} + (-1)^{a+b}\, \mathbf{a}_x^{\mathrm{T}}\, T\, \mathbf{b}_y\Big). \tag{2.20}$$

This expression holds for any arbitrary two-qubit state ρ. We now consider a special case of states without correlations between the subsystems, called product states, which can be written as $\rho_A \otimes \rho_B$. A separable bipartite state is then defined as a convex combination of product states, i.e.

$$\rho = \sum_{\lambda} q_\lambda \rho_A^\lambda \otimes \rho_B^\lambda, \tag{2.21}$$

where $q_\lambda \geq 0$ and $\sum_\lambda q_\lambda = 1$. Any bipartite state that can not be written in this form is called entangled. We discuss more about bipartite entangled states in the next section.

2.4 Entanglement

Entanglement and Bell nonlocality are closely related, and they underlie many quantum advantages in quantum computing [13]. Entanglement describes a system in which the information contained in its individual subsystems cannot fully account for the total information of the composed system [14]. Schrodinger (1935a), in his paper entitled ”The present situation in quantum mechanics” wrote

[1] Two-party system

> "Maximal knowledge of a total system does not necessarily include total knowledge of all its parts, not even when these are fully separated from each other and at the moment are not influencing each other at all"[14]

Let us consider an electron-positron pair, produced from a spinless light. The total spin of the system must sum up to zero (angular momentum conservation). Therefore, if the positron is found to be spin up, then the electron must have spin down, and vice versa. In other words, the wavefunction can be expressed with

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle_e |\downarrow\rangle_p - |\downarrow\rangle_e |\uparrow\rangle_p\right). \tag{2.22}$$

This state is a superposition of $|\uparrow\rangle_e|\downarrow\rangle_p$ and $|\downarrow\rangle_e|\uparrow\rangle_p$. In general, consider a quantum system composed of N subsystems, each associated with its own Hilbert space $\mathcal{H}_n$, such that the total Hilbert space is given by $(\bigotimes_{n=1}^{N}\mathcal{H}_n = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_N)$. A pure state is said to be entangled if it can not be written as a product state,

$$|\psi\rangle \neq |\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_N\rangle = \bigotimes_{n=1}^{N} |\psi_n\rangle. \tag{2.23}$$

By utilizing Schmidt Decomposition, any bipartite d-dimensional pure state can, up to local unitary transformations [2], be written in the following form [13]

$$|\psi\rangle = \sum_{k=1}^{d} c_k |k\rangle|k\rangle, \qquad c_k \in \mathbb{R}^+, \tag{2.24}$$

where, $\{|k\rangle\}$ denotes the bi-orthogonal basis states and c_k are the non-negative coefficients sorted from the largest c_1 to the smallest c_d. This implies that the state is entangled if and only if $c_2 \neq 0$. A maximally entangled state is a pure state for which all the Schmidt coefficients are equal. In this thesis, we focus on two qubits and, in a special case, on two ququarts (d = 4). A maximally entangled two-qubit state can be represented as

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \tag{2.25}$$

A more general two-qubit state is a partially entangled state, which is defined as

$$|\psi\rangle = \cos\theta\,|00\rangle + \sin\theta\,|11\rangle, \quad 0 < \theta < \frac{\pi}{4}. \tag{2.26}$$

Using the **Equation 2.18** the correlation matrix for a partially entangled state is given by

$$T = \mathrm{diag}(\sin 2\theta, -\sin 2\theta, 1). \tag{2.27}$$

Moreover, according to **Equation 2.18**, the reduced bloch vectors are given as

$$\mathbf{r} = \mathbf{s} = \begin{pmatrix} 0 & 0 & \cos 2\theta \end{pmatrix}^{\dagger}. \tag{2.28}$$

2.5 Bell Inequalities and Local Hidden Variable Models

The EPR paper ([1]) introduced two conditions for a physical theory: initially, is the theory correct, and secondly, is the theory complete. They argued that if the assumption of locality and realism holds, then quantum

[2] That is, for any bipartite pure state, there exists local unitary operators U and V such that $(U \otimes V)|\psi\rangle = \sum_{k=1}^{d} c_k|k\rangle|k\rangle$. These unitary transformations do not affect the entanglement properties of the state

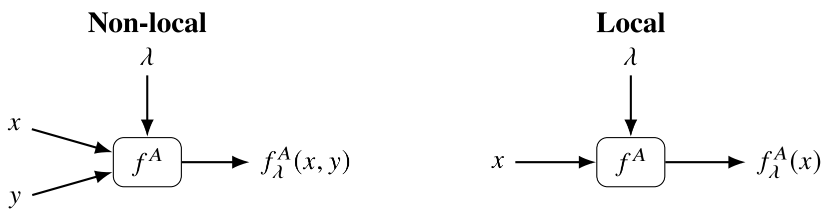
mechanics is incomplete. In 1964, John Bell introduced an inequality that draws the distinction between the quantum theory and a theory that respects locality and realism at the same time, such as local hidden variable (LHV) model [2]. Let us consider an entangled pair of photons. These photons are sent to Alice and Bob, where they perform polarization measurements $x, y \in \{0, 1\}$ respectively. We denote the outcome of these measurements by $a, b \in \{0, 1\}$. Based on these assumptions, the expectation value of ab in the measurement setting x, y, denoted by E_{xy} can be expressed as $\langle ab\rangle_{xy} = \langle A_x B_y\rangle = E_{xy} = \sum_{a,b\in\{0,1\}}(-1)^{a+b}P(a, b|x, y)$ [11]. The simplest form of Bell inequalities is the CHSH inequality [15]:

$$S = \langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle = E_{00} + E_{01} + E_{10} - E_{11} = \sum_{x,y\,\in\,\{0,1\}} (-1)^{xy}\, E_{xy} \leq 2. \tag{2.29}$$

A bipartite Bell scenario is denoted by $(m_A, m_B; d_A, d_B)$, where m_i and d_i denote the number of settings and outcomes for party i, respectively. We investigate this equation for LHV and quantum models. A Hidden Variable model can be represented by a hidden variable λ, which is passed to two photons (two parties), and the statistics are produced based on this variable. This variable is stochastically distributed according to q_λ [3], therefore the statistics of the measurements is described by

$$P(a, b|x, y) = \sum_\lambda q_\lambda\, P_\lambda(a, b|x, y), \qquad q_\lambda \geq 0 \quad \text{and} \quad \sum_\lambda q_\lambda = 1. \tag{2.30}$$

In our model, λ defines a mapping f_λ, used by each party to determine the outcome. This mapping can be either local [4] or nonlocal:



In other words a local process is defined as a process for which [13]

$$P_\lambda(a, b|x, y) = P_\lambda(a|x)P_\lambda(b|y). \tag{2.31}$$

We denote the statistics $P_\lambda(a|x)$ and $p_\lambda(b|y)$ for a local hidden variable model by $D^A_\lambda(a|x)$ and $D^B_\lambda(b|y)$ which is then given by

$$D^A_\lambda(a|x) = \delta(f^A_\lambda(x) - a), \qquad D^B_\lambda(b|y) = \delta(f^B_\lambda(y) - b). \tag{2.32}$$

For simplicity of notation, we sometimes drop the superscript A and B when no confusion arises. Nonlocality is a necessary feature of quantum models that violate the CHSH inequality **Equation 2.29**; however, it is not sufficient. Although quantum models can exhibit nonlocality, they still satisfy a constraint known as the no-signaling condition. This condition states that the marginal probability distribution of one party is independent from the other party's choice of measurement. Marginal probabilities are obtained by summing over the other party's outcome

$$P(a|x, y) = \sum_b P(a, b|x, y)\ \forall a, x, y \quad \text{and} \quad P(b|x, y) = \sum_a P(a, b|x, y)\ \forall b, x, y. \tag{2.33}$$

[3] For example λ_1 might appear more frequently than λ_2

[4] In this context, "local" means the outcome of measurement $x(y)$ only depends on the hidden variable λ and the measurement setting $x(y)$, whereas "non-local" refers to the scenario in which the outcome of one party may depend on the distant measurement setting and/or outcome. Note that here we are referring to individual outcomes, not the statistical correlations

Subsequently, we define the signaling terms $\alpha^A_{a,x}$ and $\alpha^B_{b,y}$, which vanish in the no-signalling regime and are given by

$$\alpha^A_{a,x} \equiv P(a|x,0) - P(a|x,1) = 0 \; \forall a,x \quad \text{and} \quad \alpha^B_{b,y} \equiv P(b|0,y) - P(b|1,y) = 0 \; \forall b,y. \tag{2.34}$$

Having established the key assumptions of locality, realism, and no-signaling, we now introduce the following important statements:

- Any bipartite LHV model satisfies the CHSH inequality (**Equation 2.29**), the proof is provided in the **subsection 2.5.2**.
- All pure entangled bipartite states violate the CHSH inequality (Gisin's theorem [16]) (see **appendix A**).
- All separable states (**Equation 2.21**) produce local behavior (see **appendix B**).

2.5.1 Conceptual perspective on Bell Inequalities

> "One should no more rack one's brain about the problem of whether something one cannot know anything about exists all the same, than about the ancient question of how many angels are able to sit on the point of a needle..[5] "

One can approach the Bell inequalities through both mathematical and conceptual perspectives. Let us provide a conceptual example motivated by the work presented in [18]. Consider an independent setup that generates two sets of three quantum coins, each ordered by the colors red, green, and blue (see **Figure 2.2**). The first set is sent to an observer called Alice, and the second set to an observer called Bob. Alice and Bob are then allowed to examine the state of only one coin from their respective set. They may choose either a red, green, or blue coin. Moreover, we assume that Alice's and Bob's measurements are spacelike separated events[6]. The choices are random and independent. Surprisingly, the final measurement statistics obey two rules:

- (a) If Alice and Bob choose the same color, then they always observe the same outcome, Head (H) or Tail (T).
- (b) If Alice and Bob choose different colors, they observe the same outcome in one quarter of the measurements.

Now consider that the outcome of each coin is predetermined, and we associate a three-letter representation with each set. For example, (HTT) means that the red coin will show Head, while the green and blue coins will show Tail if measured. Rule (a) dictates that the two sets of coins must follow the same letter representation. However, it can be shown that no assignment of predetermined letter representations satisfying rule (a) can also satisfy rule (b) (see **appendix C**). This implies that

$$P_{\text{LHV}}(\text{same outcome} \mid \text{different colors}) \geq \frac{1}{3}, \qquad P_{\text{QM}}(\text{same outcome} \mid \text{different colors}) = \frac{1}{4}. \tag{2.35}$$

Here, LHV denotes any local predetermined assignment method, and QM denotes the results of the current setup. Therefore, these rules exclude any local pre-preparation of the coins. But at the same time, the two coin sets have no way of signaling to each other in order to determine which measurement Alice or Bob has chosen and accordingly adjust their statistical outcome. One possible explanation can be given by considering

[5] Attributed to O. Stern by W. Pauli in a letter to M. Born, [17], p. 223.

[6] Two events are spacelike separated if they are sufficiently far apart in space, but sufficiently close in time, that no signal traveling at or below the speed of light could propagate from one event to the other, according to Special Relativity.

quantum-mechanical analogue of this scenario. One might say that the coins are described by a joint state vector, and once this state vector is measured by Alice/Bob on one set, it collapses accordingly for the other set measured by Bob/Alice. On one hand, this may be interpreted as a violation of locality; on the other hand, Alice and Bob can not use their measurement choices to communicate with each other. Thus, the measurements influence one another while still preventing information transfer between the observers.

This device is an analogy to Bohm's version [19] of the Einstein-Podolsky-Rosen experiment. The quantum coin generator, corresponds to a source generating two spin-$\frac{1}{2}$ particles in the singlet state. In our case, each particle corresponds to a set of three coins. Alice and Bob each possess a detector with Stern-Gerlach magnets. These magnets can measure the spin along three directions: one perpendicular to the plane and two at 120° angles with respect to the plane's vertical. Each direction corresponds to a choice of color in our example. Furthermore, the measured spins are either aligned with or opposite to the magnetic field, corresponding to the coins showing Head or Tail, respectively.

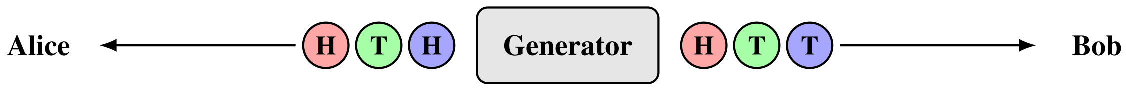


Figure 2.2: Quantum coin generator sending coin triples to Alice and Bob.

2.5.2 All LHV Models Satisfy the CHSH Inequality

Here we prove that any Local Hidden Variable (LHV) model satisfies the CHSH inequality (**Equation 2.29**), given by

$$S = \sum_{x,y} (-1)^{xy} E_{xy} = \sum_{x,y} (-1)^{xy} \sum_{a,b} (-1)^{a+b} P(a, b|x, y) \leq 2. \tag{2.36}$$

An LHV model is described by combining the measurement statistics of a hidden-variable model given in **Equation 2.30** with the definition of a local process (**Equation 2.31**). Therefore, the CHSH parameter can be written as

$$S = \sum_{x,y} (-1)^{xy} \sum_{a,b} (-1)^{a+b} \sum_{\lambda} q_\lambda P_\lambda(a, b|x, y) \tag{2.37}$$

$$= \sum_{\lambda} q_\lambda \sum_{x,y} (-1)^{xy} \sum_{a,b} (-1)^{a+b} D^A_\lambda(a|x) D^B_\lambda(b|y). \tag{2.38}$$

The last holds because q_λ is independent of the measurement settings and outcomes. Rearranging the CHSH parameter gives

$$S = \sum_{\lambda} q_\lambda \sum_{x,y} (-1)^{xy} \Big(\sum_{a} (-1)^a D^A_\lambda(a|x)\Big) \Big(\sum_{b} (-1)^b D^B_\lambda(b|y)\Big). \tag{2.39}$$

Consequently, we define the following deterministic response functions

$$A^\lambda_x = \sum_{a} (-1)^a D^A_\lambda(a|x) \in -1, 1, \quad \text{and} \quad B^\lambda_y = \sum_{b} (-1)^b D^B_\lambda(b|y) \in -1, 1. \tag{2.40}$$

Then, the CHSH expression becomes

$$S = \sum_{\lambda} q_\lambda (A^\lambda_0 B^\lambda_0 + A^\lambda_1 B^\lambda_0 + A^\lambda_0 B^\lambda_1 - A^\lambda_1 B^\lambda_1) \leq \max_{\lambda} (A^\lambda_0 B^\lambda_0 + A^\lambda_1 B^\lambda_0 + A^\lambda_0 B^\lambda_1 - A^\lambda_1 B^\lambda_1). \tag{2.41}$$

The last inequality follows from the fact that $q_\lambda \geq 0$ and $\int q_\lambda d\lambda = 1$, implying that S is a convex combination of the terms in the sum. Finally, it is evident that for any assignment of $-1, 1$ to A_x^λ and B_y^λ the expression is bounded by

$$A_0^\lambda B_0^\lambda + A_1^\lambda B_0^\lambda + A_0^\lambda B_1^\lambda - A_1^\lambda B_1^\lambda = A_0^\lambda(B_0^\lambda + B_1^\lambda) + A_1^\lambda(B_0^\lambda - B_1^\lambda) \leq 2. \tag{2.42}$$

Therefore, we established that any LHV model satisfies

$$S_{\text{LHV}} \leq 2. \tag{2.43}$$

2.5.3 Maximum CHSH Violation in Quantum Mechanics

Here we provide a derivation of the maximal CHSH violation achievable within quantum mechanics; for a less operator-heavy alternative derivation, see **appendix D**. Following a similar treatment to that presented in [13], the CHSH parameter given in **Equation 2.29** can be alternatively described using the operator $\mathbf{S}$ as

$$S = \langle\psi|\mathbf{S}|\psi\rangle = \mathrm{Tr}\big[\mathbf{S}\rho\big], \quad \text{where} \quad \mathbf{S} = A_0 \otimes B_0 + A_0 \otimes B_1 + A_1 \otimes B_0 - A_1 \otimes B_1. \tag{2.44}$$

We assume that the observables are projective with the eigenvalues ± 1. Then $A_x^2 = \mathbb{I}_A, B_y^2 = \mathbb{I}_B$, and $\|A_x\|_\infty = \|B_y\|_\infty = 1$. For simplicity we rewrite $\mathbf{S}$ as

$$\mathbf{S} = A_0 \otimes (B_0 + B_1) + A_1 \otimes (B_0 - B_1). \tag{2.45}$$

To facilitate the estimation of the operator norm, we consider the square of the CHSH operator $\mathbf{S}$

$$\begin{aligned} \mathbf{S}^2 = &A_0^2 \otimes (B_0 + B_1)^2 + A_1^2 \otimes (B_0 - B_1)^2 \\ &+ A_0A_1 \otimes (B_0 + B_1)(B_0 - B_1) \\ &+ A_1A_0 \otimes (B_0 - B_1)(B_0 + B_1). \end{aligned} \tag{2.46}$$

Simplifying further and using the properties of projective observables gives

$$\mathbf{S}^2 = 4\,\mathbb{I}_A \otimes \mathbb{I}_B - [A_0, A_1] \otimes [B_0, B_1]. \tag{2.47}$$

Using the triangle inequality, we obtain

$$\|\mathbf{S}^2\|_\infty \leq \|4I\|_\infty + \|[A_0, A_1] \otimes [B_0, B_1]\|_\infty. \leq 4 + 2 \cdot 2 = 8. \tag{2.48}$$

The last inequality follows from the standard properties of the operator norm. For example,

$$\begin{aligned} \|[A_0, A_1]\|_\infty &= \|A_0A_1 - A_1A_0\|_\infty \\ &\leq \|A_0A_1\|_\infty + \|A_1A_0\|_\infty \\ &\leq \|A_0\|_\infty\|A_1\|_\infty + \|A_1\|_\infty\|A_0\|_\infty \\ &= 2. \end{aligned} \tag{2.49}$$

Finally, it follows that $|S| = |\langle\psi|\mathbf{S}|\psi\rangle| \leq \|\mathbf{S}\|_\infty = \sqrt{\|\mathbf{S}^2\|_\infty} \leq 2\sqrt{2}$. It is important to note that this upper bound is derived purely from the properties of quantum mechanical operators, in particular the fact that not all operators commute in quantum mechanics. The LHV bound is recovered when the operators commute in **Equation 2.48**, giving $\|\mathbf{S}^2\|_\infty \leq 4$. So far, we have only obtained an upper bound for the CHSH parameter; however, it remains to show that this bound is achievable. This can be easily done by providing a quantum state and measurement settings that saturate the bound, such as those given in [11]:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \qquad A_0 = \sigma_z, \quad A_1 = \sigma_x, \qquad B_0 = \frac{\sigma_z + \sigma_x}{\sqrt{2}}, \quad B_1 = \frac{\sigma_z - \sigma_x}{\sqrt{2}}. \tag{2.50}$$

It can be verified that these choices achieve Tsirelson's bound $S = 2\sqrt{2}$. One can show that the maximum violation of the CHSH inequality achieved by a partially entangled state is given by $2\sqrt{1+\sin^2(2\theta)}$, where θ is the entanglement parameter of such a state. (for proof see **appendix E** where we also introduce the Horodecki criterion [20]).

2.5.4 I_{3322} Bell Inequality

Another well-studied Bell inequality is the I_{3322} inequality. In this scenario, both Alice and Bob have three possible inputs, $x, y \in \{1, 2, 3\}$, corresponding to three different measurement settings for each party, as opposed to two settings in the CHSH scenario. However, the number of outputs remains two, as in the CHSH case, giving a $(3, 3; 2, 2)$ configuration.[7] This inequality can be expressed as

$$\mathcal{W}_{\mathrm{I}3322} = -2P_B(0|1) - P_B(0|2) - P_A(0|1) + \sum_{x,y} c_{xy} P(0, 0|x, y) \leq 0 \tag{2.51}$$

where

$$c = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}. \tag{2.52}$$

The quantum bound of I_{3322} parameter is $\mathcal{W}_{\mathrm{I}3322} \approx 0.25$ [21].

2.6 Detection Loophole and Post-selection

In Bell tests, as in most physical experiments, detectors are imperfect. Non-zero detector inefficiency implies that some events remain undetected. Consequently, the detected events may not constitute a fair sample of the total ensemble, giving rise to a detection loophole. This loophole was first pointed out in [22], where it was shown that a LHV model can reproduce the quantum-mechanical correlations in the detected subset of events. Restricting analysis to this subset is referred to as post-selection. One can follow two main strategies when dealing with detector inefficiencies: either (i) assign an outcome to undetected events and include all runs in the analysis, or (ii) post-select on the successful events. We begin with the first case.

Let us consider the (2,2;2,2) scenario, where both Alice and Bob have detection efficiency η. Suppose that Alice and Bob aim to reach Tsirelson's bound by sharing a maximally entangled state and performing measurements, which would saturate the Tsirelson's bound under perfect detection. However, in this case, they reach $S = 2\sqrt{2}$ with a probability of η^2 corresponding to a successful detection on both sides. There is no contribution to the Bell parameter when only one party registers a detection, since CHSH correlations require joint outcomes. In such a case, the registered detection is a reduced state of a maximally entangled state, namely, a maximally mixed state, producing completely random local outcomes. With a probability of $(1 - \eta)^2$, both parties fail to register an outcome. In these runs, parties can only assign a predetermined outcome, which can be described by a LHV model and therefore bounded by $S = 2$. Therefore, the resulting CHSH parameter is

$$S = 2\sqrt{2}\,\eta^2 + 2(1 - \eta)^2. \tag{2.53}$$

[7]This notation refers only to the number of measurement settings and outcomes per party and does not constrain the dimensionality of the underlying quantum systems.

A violation of the CHSH inequality happens when $\eta > \eta_{\text{critical}}$ with $\eta_{\text{critical}} = 82.8\%$ [6].

Although sharing a maximally entangled state results in a maximal violation of CHSH inequality under ideal conditions, they are not the most optimal choice to minimize the critical efficiency η_{critical}. In 1993, Philippe Eberhard showed that the critical efficiency required for a loophole free Bell experiment [8] can be reduced below 82.8%, by using nonmaximally entangled states, provided that the signal-over-noise is very high [6]. He further, included that the critical efficiency can be as low as 66.7%.

A more complicated situation is when Alice's and Bob's detection efficiencies η_A, η_B, are not necessarily equal. Brunner et al. [24] investigated this asymmetric scenario. The Bell expression can be written as a combination of different sectors: joint detection, single detections, and no-detection events, weighted respectively by $\eta_A\eta_B$, $(1-\eta_A)\eta_B$, $(1-\eta_B)\eta_A$, and $(1-\eta_A)(1-\eta_B)$. By sharing a maximally entangled state, following a similar treatment as for the symmetric case, the Bell parameter in CHSH becomes

$$S = 2\sqrt{2}\,\eta_A\eta_B + 2(1-\eta_A)(1-\eta_B). \tag{2.54}$$

A relevant scenario studied in this paper corresponds to an atom-photon system, where the atom measurement can be performed with near unit efficiency ($\eta_A \approx 1$). This leads to a critical efficiency of $\eta_B^{\text{crit}} \approx 70.7\%$. However, it was shown that this efficiency can be reduced using a partially entangled state. In particular, a weakly entangled state gives optimal values of $\eta_B^{\text{critical}} \approx 50\%$ for CHSH and $\eta_B^{\text{critical}} \approx 43\%$ for I_{3322}.

The asymmetric scenario was also investigated by Branciard in [25], where, in contrast to the previous treatment, the focus was on post-selected data. In his treatment, a no-detection event is regarded as a new outcome, denoted by $\emptyset$. Consequently, the (2,2;2,2) scenario is extended to a (2,2;3,3) scenario, corresponding to 2 measurement settings and 3 outcomes for both Alice and Bob [9]. According to his work, to observe non-locality, we need to satify the following condition on efficiencies

$$\eta_A + \eta_b < 3\eta_A\eta_B. \tag{2.55}$$

For the case $\eta_A = \eta_B = \eta$, **Equation 2.55** results in $\eta > 2/3$ corresponding to the critical efficiency found by Eberhard [6].

Let us further investigate the post-selection scenario by introducing outcome-dependent efficiencies. In photonic Bell experiments, such efficiencies may depend on the measurement outcome or the setting due to imperfections in the optical setup. Two examples of mechanisms leading to this dependence are discussed in [8]. First the measurement setting is usually implemented using wave plates and polarizers. These optical elements can produce small beam displacements that affect coupling efficiencies when Single-Mode Fibers (SMFs) are used for photon collection. Second, fluctuations in the laser pump power can lead to variations in the detected count rate, if the measurements are performed in sequential runs. This, in turn, violates the fair sampling assumption.

We assume that the underlying model produces outcomes with the probabilities $P(a, b|x, y)$, but since we don't observe all the outcomes, the post-selected probability is then $P'(a, b|x, y)$. We can then define outcome and measurement dependent efficiency $\eta_{a|x}$ as

$$\eta_{a|x} = \frac{n_x(a)}{n_x(a) + n_x(\emptyset)}, \tag{2.56}$$

Where $n_x(a)$ is the number of photons detected with outcome a, and $n_x(\emptyset)$ is the number of photons that were not detected (and analogously for $\eta_{b|y}$). To include no-detection events as possible outcomes, we define an extended set of outcomes denoted by $\tilde{a}(\tilde{b})$, including both the standard outcomes $a(b)$ and no-detection

[8] Using the Clauser-Horne (CH) inequality [23].

[9] For further discussion of Bell inequalities in (2,2;3,3) scenario, see [26]

events $\emptyset$. In this scenario, the POVMs corresponding to the measurements are given by

$$A_{\tilde{a}|x} = \begin{cases} \eta_{a|x}\, A_{a|x}, & \tilde{a} = a, \\ \mathbb{I} - \sum_a \eta_{a|x}\, A_{a|x}, & \tilde{a} = \emptyset, \end{cases} \qquad B_{\tilde{b}|y} = \begin{cases} \eta_{b|y}\, B_{b|y}, & \tilde{b} = b, \\ \mathbb{I} - \sum_b \eta_{b|y}\, B_{b|y}, & \tilde{b} = \emptyset. \end{cases} \tag{2.57}$$

Therefore, we can define the total probability of both Alice and Bob obtaining a successful event with [7]

$$N_{x,y} = \sum_{a,b\in\{0,1\}} \eta_{a|x}\, \eta_{b|y}\, P(a,b|x,y).$$

We denote quantities in the post-selected scenario, accounting for detector inefficiencies with a prime. Therefore, the post-selected probabilities are defined as

$$P'(a,b|x,y) = \frac{\eta_{a|x}\, \eta_{b|y}}{N_{x,y}}\, P(a,b|x,y) = \frac{\eta_{a|x}\, \eta_{b|y}}{N_{x,y}}\, \mathrm{tr}\big[(A_{a|x} \otimes B_{b|y})\, \rho\big], \tag{2.58}$$

where the $\eta_{a|x}$ and $\eta_{b|y}$ are the detection efficiencies for Alice and Bob, respectively. In the case of $\eta_{a|x} = \eta_{b|y} = \eta$, it is clear that no-signaling principle **Equation 2.34** holds. However, outside this regime, there is no guarantee that the no-signaling principle will not be violated.

An example of violation of the no-signaling principle arising from post-selection with outcome-dependent efficiency is given in [8]. In this scenario, Bob's detector has a unit efficiency, while Alice's detector has efficiency $\eta < 1$ for outcome 1 and unit efficiency for outcome 0. Alice's measurement settings are given by $\sigma_x(\sigma_z)$ for A_0 and A_1, respectively, while Bob's measurement setting B_1 is $\frac{\sigma_z - \sigma_z}{\sqrt{2}}$. For this setup, the no-signaling term becomes (**Equation 2.34**) $\alpha^B_{+1,y} = \sqrt{\frac{1}{2}\,\frac{1-\eta}{1+\eta}} \neq 0$ [8]. This demonstrates a violation of no-signaling.

3 Model: Post-selection in Bell Inequality Scenarios

In this section, we investigate the post-selection effect on Bell inequalities and introduce the main models used in this thesis. We denote the Bell parameter in the post-selected scenario by $\mathcal{W}$ instead of S.

3.1 Post-selection in CHSH Inequality

To construct the CHSH inequality, we use the post-selected probabilities as defined in **Equation 2.58**. Thus, the corresponding correlation function becomes

$$E'_{x,y} = \sum_{a,b\in\{0,1\}} (-1)^{a+b}\, P'(a,b|x,y) \tag{3.1}$$

$$= \sum_{a,b\in\{0,1\}} (-1)^{a+b}\, \frac{\eta_{a|x}\eta_{b|y}\, P(a,b|x,y)}{N_{x,y}}. \tag{3.2}$$

For simplicity we choose: $\eta^A_{0|x} = \eta^B_{0|y} = \eta_0$ and $\eta^A_{1|x} = \eta^B_{1|y} = \eta_1$. Using the **Equation 2.20** the CHSH parameter becomes

$$\mathcal{W}_{\text{CHSH}} = \sum_{x,y} (-1)^{xy} E'_{x,y} = \sum_{x,y} (-1)^{xy} \sum_{a,b} (-1)^{a+b} \frac{\eta_a \eta_b \, P(a,b|x,y)}{N_{x,y}} \tag{3.3}$$

$$= \sum_{x,y} (-1)^{xy} \frac{\sum_{a,b} (-1)^{a+b} \eta_a \eta_b \, P(a,b|x,y)}{\sum_{a,b} \eta_a \eta_b \, P(a,b|x,y)}. \tag{3.4}$$

The assumption of underlying model, whether quantum or LHV, determines the form of $P(a,b|x,y)$ in **Equation 3.7**. In a quantum mechanical scenario, the joint probabilities are given by $P(a,b|x,y) = \mathrm{Tr}\big[(A_{a|x} \otimes B_{b|y})\,\rho\big]$, where ρ denotes the quantum system of the bipartite state.

3.2 Bell Bound for Bipartite Product States under Post-selection

Here, we analytically prove that a bipartite product state, even in the presence of detection efficiency, cannot violate the CHSH inequality. By substituting a product state $\rho_A \otimes \rho_B$ in the formula for the CHSH parameter **Equation 3.7**, one gets

$$\mathcal{W}_{\text{CHSH}} = \sum_{x,y} (-1)^{xy} \frac{\sum_{a,b} (-1)^{a+b} \eta_a \eta_b \, \mathrm{Tr}\big[(A_{a|x} \otimes B_{b|y})\,(\rho_A \otimes \rho_B)\big]}{\sum_{a,b} \eta_a \eta_b \, \mathrm{Tr}\big[(A_{a|x} \otimes B_{b|y})\,(\rho_A \otimes \rho_B)\big]} \tag{3.5}$$

$$= \sum_{x,y} (-1)^{xy} \frac{\sum_{a,b} (-1)^{a+b} \eta_a \eta_b \, \mathrm{Tr}\big[A_{a|x} \rho_A\big] \mathrm{Tr}\big[B_{b|y}\, \rho_B\big]}{\sum_{a,b} \eta_a \eta_b \, \mathrm{Tr}\big[A_{a|x} \rho_A\big] \mathrm{Tr}\big[B_{b|y}\, \rho_B\big]} \tag{3.6}$$

$$= \sum_{x,y} (-1)^{xy} \frac{\left(\sum_a (-1)^a \eta_a \, \mathrm{Tr}\big[A_{a|x} \rho_A\big]\right) \left(\sum_b (-1)^b \eta_b \, \mathrm{Tr}\big[B_{b|y} \rho_B\big]\right)}{\left(\sum_a \eta_a \, \mathrm{Tr}\big[A_{a|x} \rho_A\big]\right) \left(\sum_b \eta_b \, \mathrm{Tr}\big[B_{b|y} \rho_B\big]\right)}. \tag{3.7}$$

Let us define

$$\alpha_x = \frac{\sum_a (-1)^a \eta_a \, \mathrm{Tr}\big[A_{a|x} \rho_A\big]}{\sum_a \eta_a \, \mathrm{Tr}\big[A_{a|x} \rho_A\big]}, \quad \text{and} \quad \beta_y = \frac{\sum_b (-1)^b \eta_b \, \mathrm{Tr}\big[B_{b|y} \rho_B\big]}{\sum_b \eta_b \, \mathrm{Tr}\big[B_{b|y} \rho_B\big]}. \tag{3.8}$$

First we prove that $\alpha_x, \beta_y \in [-1, 1]$. This step assumes the completeness of binary measurement $A_{0|x} + A_{1|x} = \mathbb{I}$. Let us consider α_x

$$\alpha_x = \frac{\eta_0 \mathrm{Tr}\big[A_{0|x} \rho_A\big] - \eta_1 \mathrm{Tr}\big[A_{1|x} \rho_A\big]}{\eta_0 \mathrm{Tr}\big[A_{0|x} \rho_A\big] + \eta_1 \mathrm{Tr}\big[A_{1|x} \rho_A\big]} \tag{3.9}$$

$$= \frac{\eta_0 \mathrm{Tr}\big[A_{0|x} \rho_A\big] - \eta_1 (1 - \mathrm{Tr}\big[A_{0|x} \rho_A\big])}{\eta_0 \mathrm{Tr}\big[A_{0|x} \rho_A\big] + \eta_1 (1 - \mathrm{Tr}\big[A_{0|x} \rho_A\big])} \tag{3.10}$$

$$= \frac{\mathrm{Tr}\big[A_{0|x} \rho_A\big] (\eta_0 + \eta_1) - \eta_1}{\mathrm{Tr}\big[A_{0|x} \rho_A\big] (\eta_0 - \eta_1) + \eta_1} = \alpha_x(\mathrm{Tr}\big[A_{0|x} \rho_A\big]). \tag{3.11}$$

Where $\mathrm{Tr}\big[A_{0|x} \rho_A\big] \in [0, 1]$ We can show that the function $\alpha_x(\mathrm{Tr}\big[A_{0|x} \rho_A\big])$ is monotonic:

$$\mathrm{sgn}\Big(\frac{\mathrm{d}\alpha_{\mathrm{x}}(\mathrm{Tr}\big[\mathrm{A}_{0|\mathrm{x}} \rho_{\mathrm{A}}\big])}{\mathrm{d}\mathrm{Tr}\big[A_{0|x} \rho_A\big]}\Big) = \mathrm{sgn}\Big(\frac{2\eta_1(\eta_0 - \eta_1)}{(\mathrm{Tr}\big[A_{0|x} \rho_A\big] (\eta_0 - \eta_1) + \eta_1)^2}\Big) = \mathrm{sgn}(\eta_0 - \eta_1) = \text{constant}. \tag{3.12}$$

Therefore, the upper and lower bounds correspond to the endpoint values $\alpha_x(0) = -1$ and $\alpha_x(1) = 1$, so

$\alpha_x \in [-1, 1]$. Similar argument holds for β_y. Subsequently the CHSH parameter can be written as

$$\mathcal{W}_{\text{CHSH}} = \sum_{x,y} (-1)^{xy} \alpha_x \beta_y \leq 2 \quad \text{for} \quad \alpha_x, \beta_y \in [-1, 1]. \tag{3.13}$$

The last step follows from the fact that $\mathcal{W}_{\text{CHSH}}$ is linear in α_x and β_y, so its maximum over the domain $[-1, 1]$ is attained at the extremal points $\alpha_x, \beta_y \in \{-1, 1\}$, where the standard CHSH bound equals to 2. Therefore, inefficiencies do not affect the Bell bound for a product state.

3.3 Maximally and Partially Entangled States with Post-selection

For a bipartite entanlged state, the $\text{tr}\left[(A_{a|x} \otimes B_{b|y})\,\rho\right]$ can be substitued using **Equation 2.20**, then the CHSH parameter becomes

$$\mathcal{W}_{\text{CHSH}} = \sum_{x,y} (-1)^{xy} \frac{\sum_{a,b} \eta_a \eta_b \left[1 + (-1)^b \mathbf{a}_x^{\mathrm{T}} \mathbf{r} + (-1)^a \mathbf{b}_y^{\mathrm{T}} \mathbf{s} + \mathbf{a}_x^T T \mathbf{b}_y\right]}{\sum_{a,b} \eta_a \eta_b \left[1 + \mathbf{a}_x^{\mathrm{T}} \mathbf{r} + \mathbf{b}_y^{\mathrm{T}} \mathbf{s} + \mathbf{a}_x^T T \mathbf{b}_y\right]} \tag{3.14}$$

$$= \sum_{x,y} (-1)^{xy} \frac{(\eta_0 - \eta_1)^2 + (\eta_0^2 - \eta_1^2)\,(\mathbf{a}_x^{\mathrm{T}} \mathbf{r} + \mathbf{b}_y^{\mathrm{T}} \mathbf{s}) + (\eta_0 + \eta_1)^2\, \mathbf{a}_x^{\mathrm{T}}\, T\, \mathbf{b}_y}{(\eta_0 + \eta_1)^2 + (\eta_0^2 - \eta_1^2)\,(\mathbf{a}_x^{\mathrm{T}} \mathbf{r} + \mathbf{b}_y^{\mathrm{T}} \mathbf{s}) + (\eta_0 - \eta_1)^2\, \mathbf{a}_x^{\mathrm{T}}\, T\, \mathbf{b}_y}. \tag{3.15}$$

We see only the difference and twice the average of efficiencies η_0 and η_1 appear in CHSH parameter formula, therefore we define

$$\Delta = \frac{\eta_0 - \eta_1}{2}, \quad \eta = \frac{\eta_0 + \eta_1}{2}. \tag{3.16}$$

An important observation is that CHSH parameter $\mathcal{W}_{\text{CHSH}}$ can be expressed as a function of a single dimensionless parameter, the efficiency asymmetry $R = \frac{\Delta}{\eta}$, which encapsulates the relevant information of the system without any loss of generality. This allows us to construct the function $\mathcal{W}_{\text{CHSH}}(R)$, which brings almost all of the analysis in the single variable space; additionally, it leaves room for an extra parameter if necessary (later we see for a partially entangled state we need a parameter θ) to make a 3D plot. Therefore, the CHSH parameter can be written as

$$\mathcal{W}_{\text{CHSH}} = \sum_{x,y} (-1)^{xy} \frac{\Delta^2 + \eta\Delta\,(\mathbf{a}_x^{\mathrm{T}} \mathbf{r} + \mathbf{b}_y^{\mathrm{T}} \mathbf{s}) + \eta^2\, \mathbf{a}_x^{\mathrm{T}}\, T\, \mathbf{b}_y}{\eta^2 + \eta\Delta\,(\mathbf{a}_x^{\mathrm{T}} \mathbf{r} + \mathbf{b}_y^{\mathrm{T}} \mathbf{s}) + \Delta^2\, \mathbf{a}_x^{\mathrm{T}}\, T\, \mathbf{b}_y} \tag{3.17}$$

$$= \sum_{x,y} (-1)^{xy} \frac{R^2 + R\,(\mathbf{a}_x^{\mathrm{T}} \mathbf{r} + \mathbf{b}_y^{\mathrm{T}} \mathbf{s}) + \mathbf{a}_x^{\mathrm{T}}\, T\, \mathbf{b}_y}{1 + R\,(\mathbf{a}_x^{\mathrm{T}} \mathbf{r} + \mathbf{b}_y^{\mathrm{T}} \mathbf{s}) + R^2\, \mathbf{a}_x^{\mathrm{T}}\, T\, \mathbf{b}_y}. \tag{3.18}$$

Assuming a maximally entangled state $r = s = \mathbf{0}$ this expression simplifies to

$$\mathcal{W}_{\text{CHSH}} = \sum_{x,y} (-1)^{xy} \frac{R^2 + \mathbf{a}_x^{\mathrm{T}}\, T\, \mathbf{b}_y}{1 + R^2\, \mathbf{a}_x^{\mathrm{T}}\, T\, \mathbf{b}_y}. \tag{3.19}$$

For a partially entangled state, the reduced Bloch vectors and correlation matrix are given from **Equation 2.28 and Equation 2.27**.

We define a measure $\Delta(R, \theta)$ representing how much the maximal CHSH violation changes when we introduce inefficiencies

$$\Delta(R, \theta) = \max(\mathcal{W}_{\text{CHSH}}(R, \theta)) - 2\sqrt{1 + \sin^2(2\theta)}. \tag{3.20}$$

3.4 Local Hidden Variable Models under Post-selection

For a classical local hidden variable model the joint probability is given by

$$P(a,b|x,y) = \sum_\lambda q_\lambda D_\lambda(a|x) D_\lambda(b|y). \tag{3.21}$$

By substituting back in the **Equation 3.7** we get:

$$\mathcal{W}_{\text{CHSH}} = \sum_{x,y} (-1)^{xy} \frac{\sum_\lambda q_\lambda (\sum_a (-1)^a \eta_a D_\lambda(a|x))(\sum_b (-1)^b \eta_b D_\lambda(b|y))}{\sum_\lambda q_\lambda (\sum_a \eta_a D_\lambda(a|x))(\sum_b \eta_b D_\lambda(b|y))} = \sum_{x,y} (-1)^{xy} \frac{\sum_\lambda q_\lambda A_x^\lambda B_y^\lambda}{\sum_\lambda q_\lambda A_x'^\lambda B_y'^\lambda}, \tag{3.22}$$

$$\begin{aligned} A_x^\lambda &= \sum_a (-1)^a \eta_a D_\lambda(a|x) \in \{\eta_0, -\eta_1\}, \quad A_x'^\lambda = \sum_a \eta_a D_\lambda(a|x) \in \{\eta_0, \eta_1\}, \\ B_y^\lambda &= \sum_b (-1)^b \eta_b D_\lambda(b|y) \in \{\eta_0, -\eta_1\}, \quad B_y'^\lambda = \sum_b \eta_b D_\lambda(b|y) \in \{\eta_0, \eta_1\}. \end{aligned} \tag{3.23}$$

Expressed in terms of R, these quantities simplify to

$$A_x^\lambda, B_y^\lambda \in \{R+1, -(1-R)\}, \quad A_x'^\lambda, B_y'^\lambda \in \{R+1, 1-R\} \tag{3.24}$$

One can notice the relationship for $\mathcal{W}_{\text{CHSH}}$ can be reshaped as

$$\mathcal{W}_{\text{CHSH}} = \sum_{x,y} (-1)^{xy} \sum_\lambda \frac{q_\lambda A_x'^\lambda B_y'}{\sum_\lambda q_\lambda A_x'^\lambda B_y'^\lambda} \cdot \frac{A_x^\lambda}{A_x'^\lambda} \cdot \frac{B_y^\lambda}{B_y'^\lambda} = \sum_\lambda q_\lambda'(x,y) \alpha_x \beta_y, \tag{3.25}$$

where we defined $q_\lambda'(x,y) = \frac{q_\lambda A_x'^\lambda B_y'^\lambda}{\sum_\mu q_\mu A_x'^\mu B_y'^\mu}$, $\alpha_x = \frac{A_x^\lambda}{A_x'^\lambda}$, and $\beta_y = \frac{B_y^\lambda}{B_y'^\lambda}$. It is clear that $\alpha_x, \beta_y \in \{-1, 1\}$. Therefore, this shows that a Local hidden variable model in the presence of detector inefficiencies can be modeled as a measurement-dependent hidden-variable (MDHV) model with the hidden variable distribution $q_\lambda'(x,y)$ [22]. Therefore, by giving up the measurement independence assumption [1], the violation of the CHSH inequality achieved by inefficient measurements can be reproduced using fully efficient measurements. Thus, the question arises about how much one needs to relax this assumption. Following the work in [27], one can introduce the measure M for this particular purpose

$$M := \frac{1}{2} \sup_{x,x',y,y'} \sum_\lambda \left| q_\lambda'(x,y) - q_\lambda'(x',y') \right|. \tag{3.26}$$

M quantifies the measurement independence [2], where $M = 0$ corresponds to completely independent measurement settings, and $M = 1$ means no free will in choosing measurement settings. According to the work done in [27], the maximal CHSH violation achieved by a measurement-dependent hidden variable model is given by

$$\mathcal{W}_{\text{CHSH}} \le min(2 + 6M, 4). \tag{3.27}$$

[1]The measurement settings are chosen independently of any variable in the system.

[2]We note that the quantity M defined in [27] differs by a factor of two from our definition

3.5 Optimization for Higher-Dimensional Quantum States

To generate higher-dimensional quantum systems, we parametrized the density matrix ρ using the Cholesky decomposition [28]. Specifically, any positive semidefinite matrix can be written as $LL^\dagger$, and we construct the density operator as

$$\rho = \frac{LL^\dagger}{\mathrm{Tr}\left[LL^\dagger\right]}. \tag{3.28}$$

Here L is any lower triangular complex matrix with a real elements on the diagonal. This expression automatically produces a positive semidefinite and hermitian matrix satisfying $\mathrm{Tr}\left[\rho\right] = 1$.

3.6 Post-selection in I_{3322} Bell Inequality

We replace all the probabilities in the **Equation 2.51** with their post-selected counterpart. In analogy to **section 3.1** we define:

$$P'_B(0|1) = \frac{\eta_0^B}{N_{y=1}^B} P_B(0|1), \qquad N_{y=1}^B = \sum_{b\in\{0,1\}} \eta_b^B P_B(b|1), \tag{3.29}$$

$$P'_B(0|2) = \frac{\eta_0^B}{N_{y=2}^B} P_B(0|2), \qquad N_{y=2}^B = \sum_{b\in\{0,1\}} \eta_b^B P_B(b|2), \tag{3.30}$$

$$P'_A(0|1) = \frac{\eta_0^A}{N_{x=1}^A} P_A(0|1), \qquad N_{x=1}^A = \sum_{a\in\{0,1\}} \eta_a^A P_A(a|1). \tag{3.31}$$

Additionally, $P'(a,b|x,y)$ is given in **Equation 2.58**. We consider the general two-qubit density operator (**Equation 2.17**) to derive an expression for the maximal I_{3322} violation. In order to derive the probabilities, the relationship given in **Equation 2.20** is used for $P(a,b|x,y)$, while $P_A(a|x) = \frac{1}{2}\left(1 + (-1)^a\, \mathbf{a}_x \cdot \mathbf{r}\right)$ and $P_B(b|y) = \frac{1}{2}\left(1 + (-1)^b\, \mathbf{b}_y \cdot \mathbf{s}\right)$. To implement an LHV model, we simply substitute $P(a|x) = \sum_\lambda q_\lambda D_\lambda(a|x)$ and $P(b|y) = \sum_\lambda q_\lambda D_\lambda(b|y)$, while using **Equation 3.21** for $P(a,b|x,y)$.

4 Methodology

The numerical optimization carried out in this thesis were performed using nonlinear programming (NLP). Nonlinear programming refers to the class of optimization processes in which the objective or at least one of the constraints depends nonlinearly on the variable. Generally, this problem can be formulated as finding the minimum of a real-valued function f, where the optimization variable is $x \in \mathbb{R}^n$. subject to equality and inequality constraints:

$$\begin{aligned} \min_x \quad & f(x) \\ \text{subject to} \quad & c_i(x) = 0, \quad i \in E, \\ & c_i(x) \geq 0, \quad i \in I. \end{aligned} \tag{4.1}$$

Here, each c_i is a real-valued constraint function of x, while E and I denote sets of indices corresponding to equality and inequality constraints, respectively. [29] To solve the optimization problems numerically, routines from SciPy [30] were used, specifically `L-BFGS`, `SLSQP`, and differential evolution. Due to their efficiency, the gradient-based methods (`L-BFGS`, `SLSQP`) were primarily used to obtain high-precision local optima. To ensure the results were not artifacts of local maxima, the solutions were validated using the global optimization method differential evolution.

To maximize the CHSH parameter over measurement settings for a maximally entangled two-qubit state, only the measurement vectors $\mathbf{a}_x$ and $\mathbf{b}_y$ were parametrized on the Bloch sphere and optimized numerically. For a partially entangled two-qubit state, an additional entanglement parameter θ was introduced and therefore included in the optimization. In this scenario, the corresponding Bloch vectors $\mathbf{r}, \mathbf{s}$ and correlation matrix T depend explicitly on θ. A similar approach was utilized for the optimizations involving I_{3322} Bell inequality.

To optimize over all two-qubit states, the density matrix ρ was parametrized using the Cholesky decomposition (see **section 3.5**). From this parametrization, the Bloch vectors $\mathbf{r}, \mathbf{s}$ and the correlation matrix T were computed numerically. This allowed us to benefit from the computational efficiency obtained by the analytically derived expression of the CHSH parameter in terms of $\mathbf{r}, \mathbf{s}$ and T.

The situation was different in the two-ququart optimization, where both the state and the projective measurements were varied over their full parameter spaces. In this case, the Bloch-vector formalism was no longer used. Instead, the measurement operators $A_{a|x}$ and $B_{b|y}$ were constructed explicitly through matrix parametrizations. This led to a substantially higher computational cost. For the LHV model, each hidden variable λ is described as a deterministic strategy specifying outputs for all measurment settings, namely

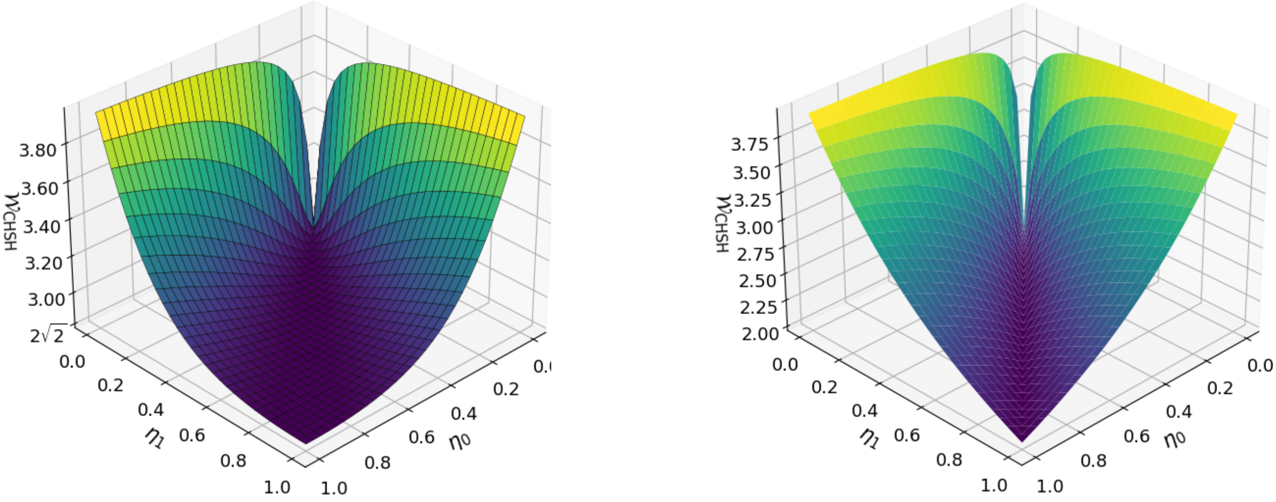
$$\lambda = (a_0, a_1, b_0, b_1). \tag{4.2}$$

5 Results and Discussion

5.1 Post-selection Effects on the CHSH Inequality

The effect of post-selection on the maximal CHSH violation was investigated through numerical optimization of **Equation 3.15**, where the optimization is performed over measurement settings for a two-qubit maximally entangled state. The maximal CHSH violation achieved by this state is represented as a function of outcome-based efficiencies η_0 and η_1 in **Figure 5.1a**. We observe that the maximal CHSH parameter remains at $2\sqrt{2}$ when $\eta_0 = \eta_1$, whereas for $\eta_0 \neq \eta_1$ it exceeds this value. When the maximization was also carried out over all two-qubit states using the **Equation 3.15** (instead of restricting to the maximally entangled state), the same plot was reproduced. This suggests that, within the set of two-qubit states, the maximal entangled state achieves the highest CHSH value for the considered post-selected scenario.

The **Equation 3.22** was used to calculate the maximal CHSH violation of a post-selected LHV model; the result is shown in **Figure 5.1b**. In this scenario, there are no apparent CHSH violations for $\eta_0 = \eta_1$. However, for $\eta_1 \neq \eta_1$, there exist LHV models that can produce CHSH violations exceeding the original Tsirelson's bound ($\mathcal{W}_{\text{CHSH}} = 2\sqrt{2}$).



(a) Maximally entangled state

(b) LHV model

Figure 5.1: Maximal CHSH parameter under post-selection as a function of outcome-dependent detection efficiencies η_0 and η_1. (a) maximally entangled quantum state. (b) LHV model.

We introduced the parameter R in **section 3.3**, which allows the maximal CHSH violation achieved for both the quantum model and the LHV model to be expressed as a function of a single variable. This facilitates a direct comparison between the two cases. For the two-qubit quantum model, the **Figure 5.1** is optimized, while for the LHV model, **Equation 3.22** is optimized using the parametrization given in **Equation 3.24**. The resulting behavior is displayed in **Figure 5.2**.

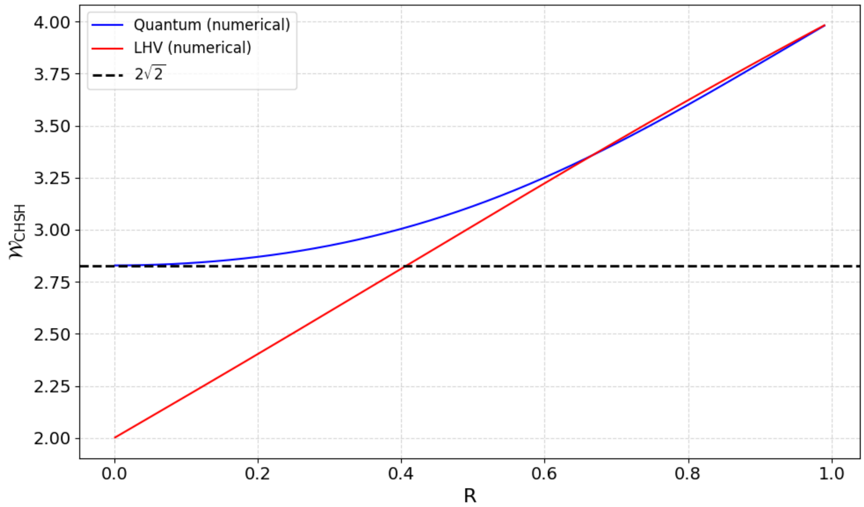


Figure 5.2: The maximal CHSH violation under post-selection as a function of $R = \frac{\eta_0 - \eta_1}{\eta_0 + \eta_1}$ for a two-qubit quantum model (blue) and an LHV model (red).

Fitting the data for both models gives:

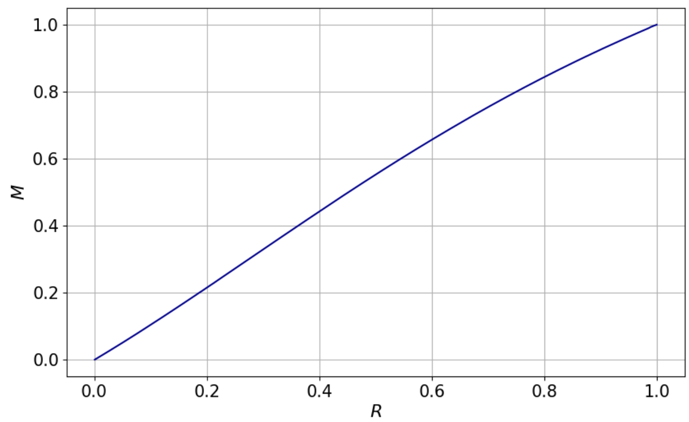
$$\text{LHV (linear):} \quad \mathcal{W}_{\mathrm{CHSH}} \approx 2.02R + 2.00, \qquad R^2_{\mathrm{fit}} = 0.99986, \tag{5.1}$$

$$\text{Quantum (quadratic):} \quad \mathcal{W}_{\mathrm{CHSH}} \approx R^2 - 0.08R + 2.83 \approx 1.3R^2 + 2\sqrt{2}, \qquad R^2_{\mathrm{fit}} = 0.99990. \tag{5.2}$$

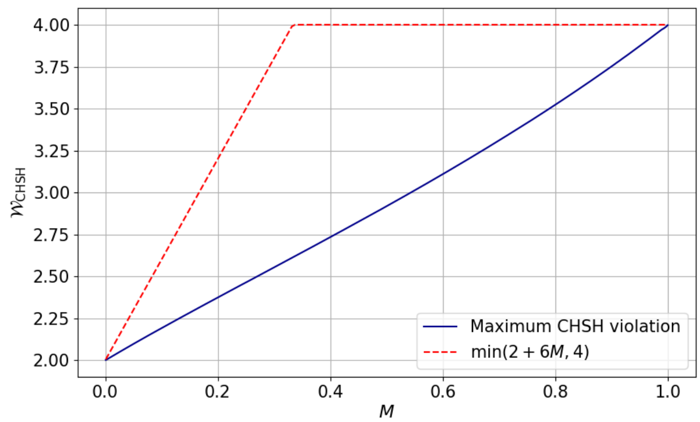
An essential observation is that the LHV model predicts a linear dependence on R, whereas the two-qubit quantum model follows an approximately quadratic trajectory. This reflects that the LHV model responds primarily to the imbalance between efficiencies encoded in R, while the quantum model shows a dependence that involves a joint contribution of both inefficiencies, leading to a nonlinear (quadratic) scaling in R.

The linear dependence on R also appears in the measurement-dependent hidden-variable (MDHV) representation of the LHV model in the presence of inefficiencies. As explained in **section 3.4**, a post-selected LHV model with inefficiencies can be mapped onto an MDHV model. Using the **Equation 3.26**, the measurement independence parameter M is plotted as a function of R in **Figure 5.3a**. In **Figure 5.3b**, the maximum violation achieved by the LHV model is plotted versus M, together with the reference line $\min(2 + 6M, 4)$. In both plots, an approximately linear dependence is observed. This implies that the measurement dependence M is generated proportionally to the detection asymmetry.

Although the inefficient LHV model can be represented by an MDHV model, this mapping is not the most optimal among such models for violating the CHSH inequality. Reproducing the original Tsirelson's bound through the measurement-dependent hidden variable mapping requires $M = 0.44$, corresponding to a loss of 44% of measurement independence. By contrast, an optimal MDHV model requires only 13% measurement independence loss to saturate the same bound.



(a) M as a function of R



(b) $\mathcal{W}_{\text{CHSH}}$ as a function of M

Figure 5.3: Relationship between detection asymmetry R, measurement independence M, and the maximal CHSH parameter in a post-selected LHV model mapped to an MDHV model. Panel (a) shows how M depends on R, while panel (b) illustrates the behavior of the maximal CHSH parameter ($\mathcal{W}_{\text{CHSH}}$) as a function of M.

As observed in **Figure 5.2** for the range $0.67 \lesssim R < 1$, the LHV model gives a larger violation of the CHSH inequality than the one attained by the two-qubit quantum model. This indicates that the two-qubit model is not sufficiently expressive to reproduce all the correlations generated by the LHV model in this regime, suggesting that the latter effectively requires a higher-dimensional representation.

To investigate this further, we examined the distribution of the hidden-variable probabilities q_λ. In this parameter region, only four values of q_λ are nonzero. For instance at $R = 0.8$, the corresponding values of q_λ are presented in **Table F.1** of **Appendix F**. This suggests that extending the quantum description to a higher-dimensional Hilbert space, specifically a bipartite system of two ququarts, may provide an advantage. Here, a ququart refers to a four-dimensional local Hilbert space, as opposed to a two-dimensional Hilbert space for a qubit.

Accordingly, in **Figure 5.4**, we plot the maximal CHSH parameter achieved by the two-ququart quantum model together with the one achievable by the LHV model. The resulting two-ququart curve closely resembles that of the two-qubit one, but it attains slightly larger values precisely in the interval where the two-qubit model failed to surpass the LHV bound. In that region, the two-ququart curve coincides with the LHV curve, and both follow an approximately linear dependence on R for $R \gtrsim 0.67$.

These results indicate that for sufficiently large detection efficiency bias [1] the LHV model and the two-ququart quantum model predict the same maximal violation of the CHSH inequality. Moreover, the linear dependence on R in the two-ququart case for large R suggests that, in the regime of strong asymmetry in the detection efficiencies, joint contributions become negligible and the behavior is governed primarily by the asymmetry itself, leading to an approximately linear scaling.

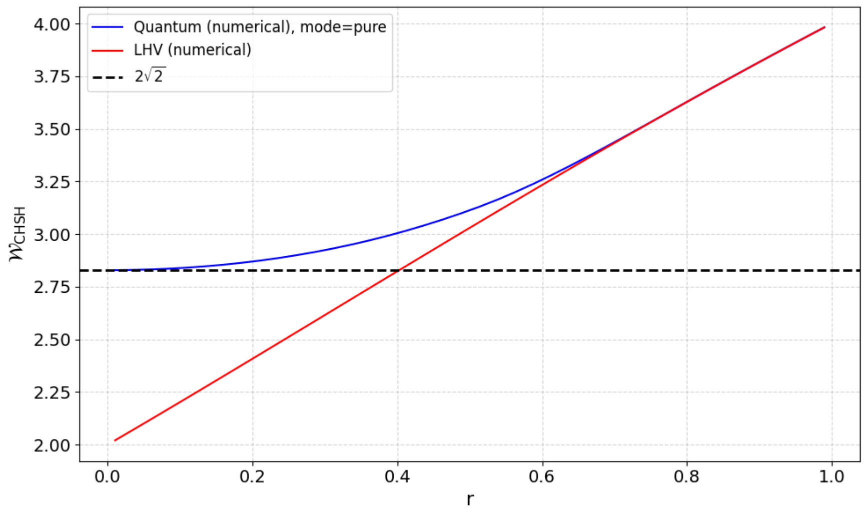


Figure 5.4: The maximal CHSH violation under post-selection is plotted as a function of the parameter $R = \frac{\eta_0 - \eta_1}{\eta_0 + \eta_1}$ for a two-ququart quantum model (blue) and local hidden variable models (red).

In the discussion of the critical efficiency required to perform a loophole-free Bell test (**section 2.6**), we observed that partially entangled two-qubit states can reduce the critical efficiency. Motivated by this, we investigate the maximal CHSH violation achieved by partially entangled states in the presence of post-selection. We substitute the reduced Bloch vectors and correlation matrix given in **Equation 2.28 and Equation 2.27** into and optimize the CHSH parameter over all measurement settings. The resulting plot is illustrated in **Figure 5.5**, where the maximal CHSH violation is presented as a function of asymmetry parameter R and the entanglement parameter θ. We observe that, for large asymmetry parameters R, the maximal CHSH violation achieved by weakly entangled states becomes highly sensitive to variations in the entanglement parameter θ. However, for all values of R, the maximally entangled state produces the largest CHSH violations.

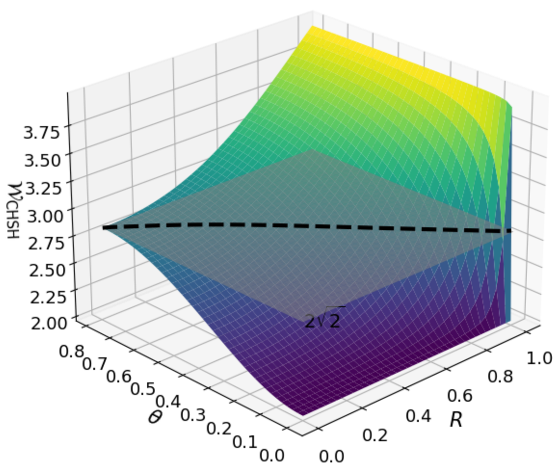
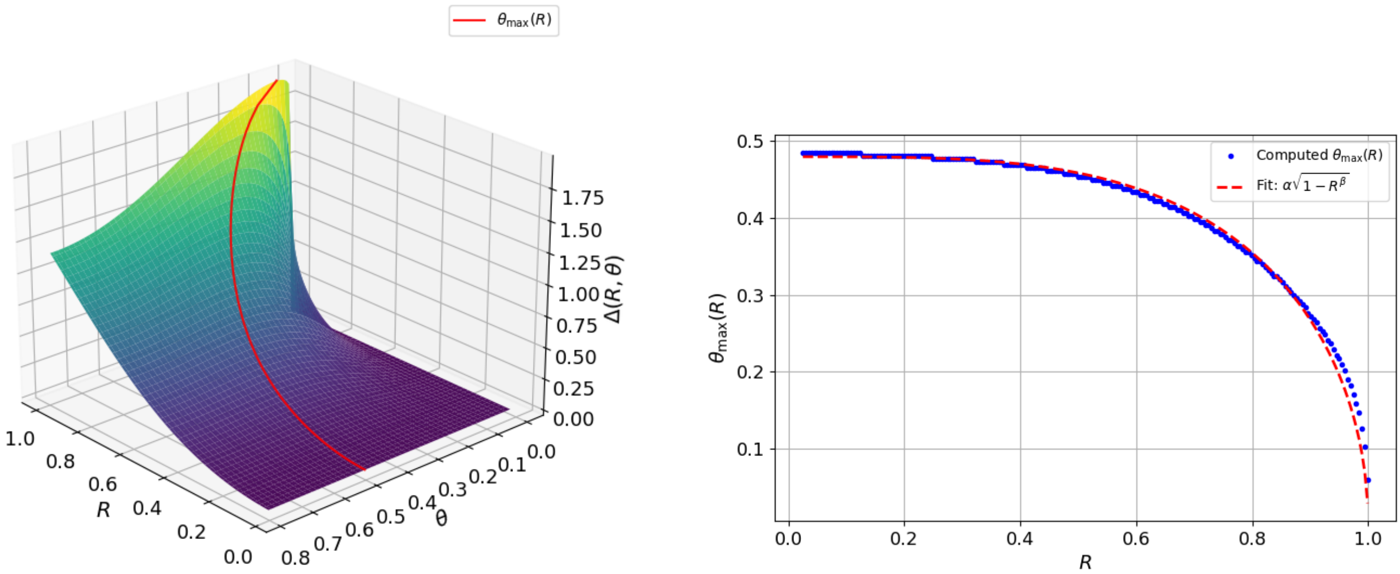


Figure 5.5: The maximal CHSH violation under post-selection is plotted as a function of R and θ for a partially entangled two-qubit state $|\psi\rangle = \cos\theta\,|00\rangle + \sin\theta\,|11\rangle$.

[1] When the difference in inefficiencies is comparable to their absolute values

A more informative measure is obtained by considering the difference between the maximal CHSH violation in the presence of post-selection from its original value, denoted by $\Delta(R, \theta)$ for each partially entangled two-qubit state. The plot of $\Delta(R, \theta)$ is shown in **Figure 5.6a**. The ridge in the plot defines θ_{max}, the value of θ that maximizes Δ for each fixed R. The resulting values of θ_{max} are plotted separately in **Figure 5.6**. For a large efficiency asymmetry ($R \longrightarrow 1$), the maximum of $\Delta(R, \theta)$ is achieved for weakly entangled states, approaching the product-state limit. Although this corresponds to a different operational setting, it again highlights the importance of weakly entangled states in post-selection scenarios, as also identified by Eberhard [6], where the minimum critical detection efficiency for a loophole-free Bell test is achieved using non-maximally entangled states in a specific limit approaching product states.



(a) The post-selection induced change in CHSH violation $\Delta(R, \theta)$ as a function of R and θ. The ridge in the plot corresponds to θ_{max} defined as the value of θ that maximizes Δ for each R.

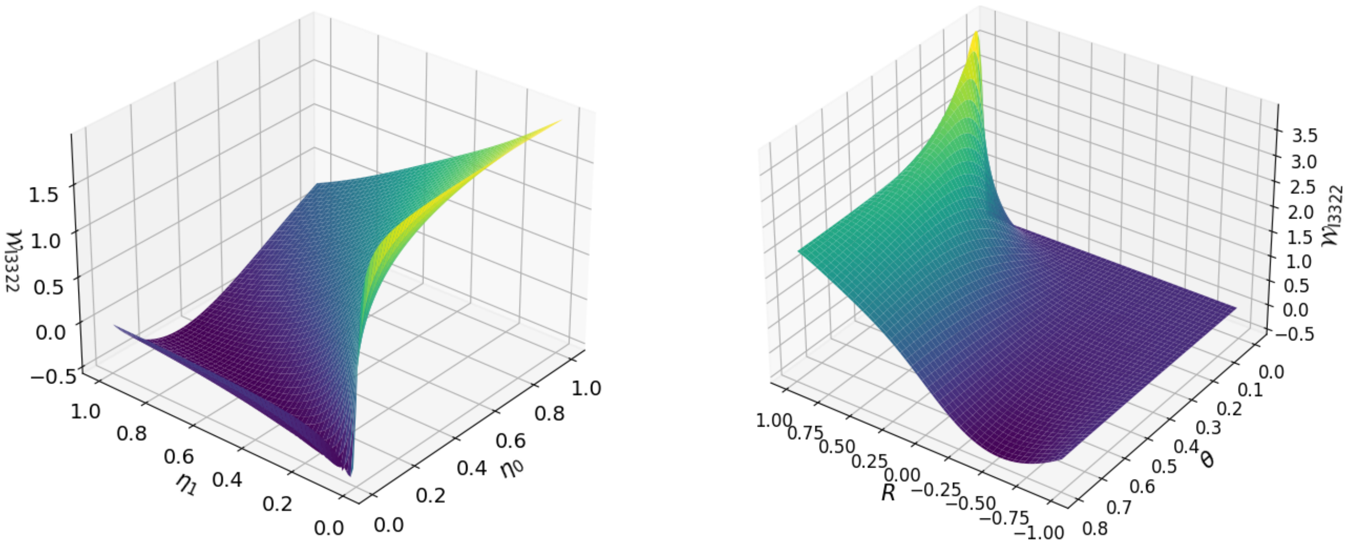
(b) Extracted optimal entanglement parameter θ_{max} as a function of R, with respect to its empirical fit ($\mathcal{R}_{\text{fit}}{}^2 = 0.991259,\ \alpha = 0.4802, \beta = 3.5190$).

Figure 5.6: The post-selection enhancement $\Delta(R, \theta)$ and the corresponding optimal entanglement parameter $\theta_{\text{max}}(R)$, together with its empirical fit.

We approximate the data using the function: $\theta_{max}(R) = \alpha\sqrt{1 - R^\beta}$. The fit is represented in Figure 5.6b, giving the parameters $\alpha = 0.4802, \beta = 3.5190$. This leads to the approximation expression $\theta_{max}(R) \approx \frac{\sqrt{1-R^{3.5}}}{2}$. This result indicates that as the detection asymmetry increases, the entanglement that benefits most from the post-selection shifts continuously from strongly entangled to weakly entangled states.

5.2 Post-selection Effects on the I_{3322} Inequality

Here we simulate the effect of post-selection on the I_{3322} inequality using the model established in **section 3.6 and subsection 2.5.4**. For a maximally entangled two-qubit state, the maximal I_{3322} violation is plotted as a function of outcome-dependent efficiencies, η_0 and η_1, in **Figure 5.7a**. A notable feature of this plot is its asymmetry with respect to η_0 and η_1. Unlike the CHSH inequality, I_{3322} inequality (**Equation 2.51**) is asymmetric in its treatment of measurement outcomes. Motivated by this asymmetry, we extend the parameter range to $R \in [-1, 1]$ in order to capture all relevant information about the system. The maximal I_{3322} violation for partially entangled two-qubit states as a function of R and θ is represented in **Figure 5.7b**. We observe that for large efficiency asymmetry with $\eta_0 > \eta_1$, the maximal violation is largest in the weakly entangled regime. In contrast, for smaller asymmetries R or for $\eta_0 < \eta_1$, post-selection produces almost no apparent violation in the weakly entangled regime.



(a) Maximal I_{3322} violation for a maximally entangled two-qubit state as a function of η_0 and η_1.

(b) Maximal I_{3322} violation for a partially entangled two-qubit state as a function of the asymmetry parameter R and the entanglement parameter θ.

Figure 5.7: Maximal I_{3322} violation under post-selection for maximally and partially entangled two-qubit states.

To compare the behavior of different partially entangled two-qubit states in this scenario, we plot the maximal I_{3322} violation as a function of R in **Figure 5.8** for four representative values of θ.

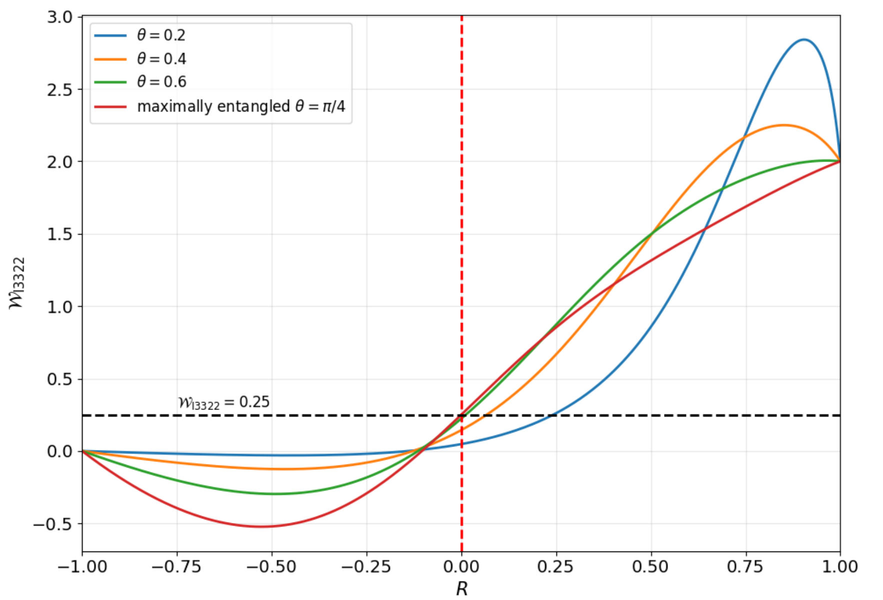


Figure 5.8: The maximal I_{3322} parameter is plotted as a function of the parameter $R = \frac{\Delta}{\eta}$ for two-qubit partially entangled state for different values of θ.

In **Figure 5.9**, the maximum I_{3322} violation achievable by an LHV model under post-selection is plotted as a function of η_0 and η_1. The line $\eta_0 = \eta_1$ draws a clear border in the plot. In the region $\eta_1 > \eta_0$, no violation is observed, i.e, $\mathcal{W}_{I3322} = 0$, whereas for $\eta_1 < \eta_0$, a clear violation of I_{3322} inequality emerges, even some values exceeding the quantum bound $\mathcal{W}_{I3322} \approx 0.25$ [21]. Therefore, there exists a regime of detection efficiencies $\eta_1 > \eta_0$ in which any post-selected LHV model fails to reproduce I_{3322} violations.

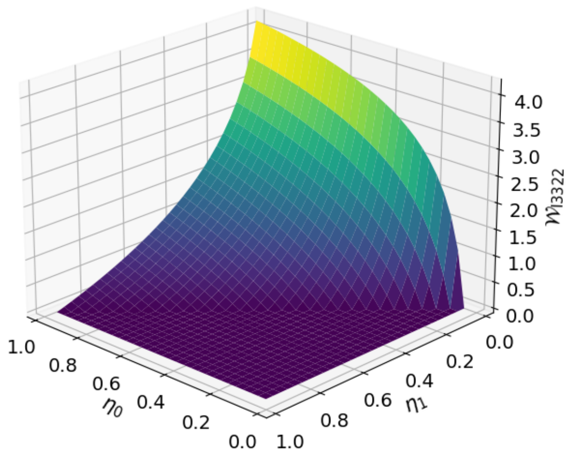


Figure 5.9: Maximal value of I_{3322} parameter as a function of detection efficiencies η_0 and η_1 , obtained using a LHV model under post-selection.

We have also found that for $R = 0.09$, the LHV model exceeds the standard quantum bound, corresponding to the efficiency ratio $\eta_0 = 1.2\eta_1$. In contrast, for the CHSH inequality, the post-selected LHV model saturates Tsirelson's bound when $R = 0.41$, corresponding to $\eta_0 = 2.4\,\eta_1$.

6 Conclusion

In this work, we explored the effect of introducing outcome-dependent efficiencies in Bell tests. The study investigated post-selected data under the assumption that the system follows either quantum mechanics or an LHV model. We showed that the behavior of both models can be effectively characterized by introducing the efficiency asymmetry parameter $R = \frac{\eta_0 - \eta_1}{\eta_0 + \eta_1}$. Numerical optimizations were carried out to quantify the amount of CHSH violation produced by an LHV model in the presence of inefficiencies. Remarkably, the violation achieved by the LHV model can even exceed that of the two-qubit quantum model for $R \gtrsim 0.67$. We also provided a derivation in **section 3.4** showing that an LHV model in a post-selection scenario resembles a Measurement-dependent hidden-variable model with perfect measurements, where the modified hidden-variable distribution $q'_\lambda(x, y)$ is given by

$$\boxed{q'_\lambda(x, y) = \frac{A'^{\lambda}_x B'^{\lambda}_y}{\sum_\mu q_\mu A'^{\mu}_x B'^{\mu}_y}\, q_\lambda.} \tag{6.1}$$

The measurement independence parameter M required for this post-selected LHV model to saturate the Tsirelson's bound was calculated to be 44%. In contrast, for an optimal Measurement-dependent hidden-variable model, only $M = 14\%$ is required to saturate Tsirelson's bound. The main results for the post-selected LHV model in the CHSH inequality are summarized below.

LHV Model

For post-selected data arising from asymmetric detection efficiencies $\eta_0 \neq \eta_1$, there exists an LHV model which reproduces an apparent CHSH violation. Therefore, from an experimental perspective, a reliable violation of the CHSH inequality requires either well-balanced efficiencies or a regime in which observed violations exceed the maximum achievable by a post-selected LHV model. In this scenario, the maximal CHSH parameter depends approximately linearly on R:

$$\mathcal{W}_{\text{CHSH}} \approx 2R + 2.$$

This shows that the effect of a post-selected LHV model on the maximal CHSH violation is primarily governed by the efficiency asymmetry R, rather than by nonlinear combinations of η_0 and η_1. The simulations further suggest that the measurement independence parameter M of the corresponding measurement-dependent hidden variable model follows the numerically observed relation

$$M \approx R.$$

Furthermore, we proved analytically that product states do not violate the CHSH inequality in the post-selected scenario, regardless of the values of the outcome-dependent efficiencies. Numerical calculations further showed that maximal CHSH violation for a two-qubit quantum model increases quadratically with R. This indicates the presence of nonlinear contributions of η_0 and η_1 in this case, in contrast to the post-selected LHV model. In the two-qubit quantum case, partially entangled states showed lower maximal CHSH values than the maximally entangled ones. However, the maximum difference between the ideal Bell value and the post-selected value, Δ, is not achieved by a maximally entangled state. Instead, a relation (**Figure 5.1**) was derived for the entanglement parameter $\theta(R)$ of a partially entangled state that maximizes this quantity. As observed, weakly entangled states show the largest enhancement of the CHSH violation induced by post-selection in the regime of large efficiency asymmetry R.

The two-qubit quantum model was found to be insufficient to reproduce CHSH violations as large as those produced by the post-selected LHV model for $R \gtrsim 0.67$. To address this, a two-ququart quantum model was optimized. Although this model follows a trajectory similar to the two-qubit case with respect to R, the slight improvement ensures that it remains at least as large as the LHV model for all values of R.

For the I_{3322} inequality, unlike the CHSH case, the maximal violation is not saturated by a maximally entangled state. The I_{3322} parameter was numerically maximized for several partially entangled states and plotted as a function of R. The LHV model does not produce any violation when $\eta_0 < \eta_1$, whereas violations occur when $\eta_0 > \eta_1$. Additionally, violations exceed the quantum bound ($I_{3322} = 0.25$) at $R = 0.09$ corresponding to $\eta_0 = 1.2\eta_1$.

This work can contribute to future studies of the detection loophole and post-selection effects in Bell experiments. The numerical framework developed here provides a systematic way to approach outcome-based asymmetric efficiencies and their effect on the observed non-local correlations. The results obtained here suggest several directions for further research. One direction would be to analytically characterize the observed dependence on the parameter R, in particular, the linear behavior found in the LHV model.

The analysis may also be relevant to investigating deviations from the Tsirelson's bound under imperfect measurement scenarios. Moreover, it can be potentially used in the studies of measurement dependence in hidden-variable theories. The framework developed in this work could help to clarify the relationship between detection loopholes, measurement dependence, and nonlocality.

Bibliography

[1] A. Einstein, B. Podolsky, and N. Rosen. "Can quantum-mechanical description of physical reality be considered complete?" In: *Physical Review* 47.10 (1935), pp. 777–780. DOI: 10.1103/PhysRev.47.777.

[2] J. S. Bell. "On the Einstein Podolsky Rosen paradox". In: *Phys. Phys. Fizika* 1.3 (1964), pp. 195–200. DOI: 10.1103/PhysicsPhysiqueFizika.1.195.

[3] A. Aspect, J. Dalibard, and G. Roger. "Experimental test of Bell's inequalities using time-varying analyzers". In: *Phys. Rev. Lett.* 49.0 (1982), pp. 1804–1807. DOI: 10.1103/PhysRevLett.49.1804.

[4] G. Weihs et al. "Violation of Bell's inequality under strict Einstein locality conditions". In: *Phys. Rev. Lett.* 81.0 (1998), pp. 5039–5043. DOI: 10.1103/PhysRevLett.81.5039.

[5] M. A. Rowe et al. "Experimental violation of a Bell's inequality with efficient detection". In: *Nature* 409 (2001), pp. 791–794.

[6] P. H. Eberhard. "Background level and counter efficiencies required for a loophole-free Einstein-Podolsky-Rosen experiment". In: *Phys. Rev. A* 47.2 (1993), pp. 747–750.

[7] L. Maquedano et al. "Bell and EPR experiments with signalling data". In: *arXiv preprint arXiv:2602.05507* (2026). arXiv: 2602.05507.

[8] M. Smania et al. "How to avoid (apparent) signalling in Bell tests". In: *Quantum* 9 (2025), p. 1760. DOI: 10.22331/q-2025-06-04-1760.

[9] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information: 10th anniversary edition*. Cambridge: Cambridge University Press, 2010. ISBN: 9780511976667. DOI: 10.1017/CBO9780511976667.

[10] J. Szangolies, H. Kampermann, and D. Bruß. "Device-independent bounds on detection efficiency". In: *Phys. Rev. Lett.* 118.0 (2017), p. 140502.

[11] A. Tavakoli. *Lecture notes for quantum information theory*. Physics Department and NanoLund, Lund University. Dated 20 May 2025. Lund, Sweden, 2025.

[12] *How to draw a Bloch sphere*. TeX Stack Exchange: https://tex.stackexchange.com/questions/345420/how-to-draw-a-bloch-sphere. TeX Stack Exchange.

[13] V. Scarani. *Bell nonlocality*. Oxford: Oxford University Press, 2019. ISBN: 9780198788416. DOI: 10.1093/oso/9780198788416.001.0001.

[14] C. Brukner, M. Żukowski, and A. Zeilinger. "The essence of entanglement". In: *arXiv preprint quant-ph/0106119* (2001). arXiv: quant-ph/0106119.

[15] J. F. Clauser et al. "Proposed experiment to test local hidden-variable theories". In: *Phys. Rev. Lett.* 23 (1969), pp. 880–884. DOI: 10.1103/PhysRevLett.23.880.

[16] N. Gisin. "Bell's inequality holds for all non-product states". In: *Phys. Lett. A* 154 (1991), pp. 201–202.

[17] M. Born, ed. *The Born–Einstein letters*. New York: Walker, 1971.

[18] N. D. Mermin. "Bringing home the atomic world: quantum mysteries for anybody". In: *Am. J. Phys.* 49.10 (1981), pp. 940–943. DOI: 10.1119/1.12594.

[19] D. Bohm. *Quantum theory*. Englewood Cliffs, NJ: Prentice-Hall, 1951.

[20] R. Horodecki, P. Horodecki, and M. Horodecki. “Violating Bell inequality by mixed spin-1/2 states: necessary and sufficient condition”. In: *Phys. Lett. A* 200 (1995), pp. 340–344. DOI: 10.1016/0375-9601(95)00214-N.

[21] K. F. Pál and T. Vértesi. “Maximal violation of a bipartite three-setting, two-outcome Bell inequality using infinite-dimensional quantum systems”. In: *Phys. Rev. A* 82 (2010), p. 022116. DOI: 10.1103/PhysRevA.82.022116.

[22] P. M. Pearle. “Hidden-variable example based upon data rejection”. In: *Phys. Rev. D* 2 (1970), pp. 1418–1425. DOI: 10.1103/PhysRevD.2.1418.

[23] J. F. Clauser and M. A. Horne. “Experimental consequences of objective local theories”. In: *Phys. Rev. D* 10 (1974), pp. 526–535. DOI: 10.1103/PhysRevD.10.526.

[24] N. Brunner et al. “Detection loophole in asymmetric Bell experiments”. In: *Phys. Rev. Lett.* 98 (2007), p. 220403. DOI: 10.1103/PhysRevLett.98.220403.

[25] C. Branciard. “Detection loophole in Bell experiments: how post-selected local correlations can look non-local”. In: *Phys. Rev. A* 83 (2011), p. 032123. DOI: 10.1103/PhysRevA.83.032123.

[26] D. Collins and N. Gisin. “A relevant two-qubit Bell inequality inequivalent to the CHSH inequality”. In: *J. Phys. A: Math. Gen.* 37 (2004), pp. 1775–1787. DOI: 10.1088/0305-4470/37/5/021.

[27] M. J. W. Hall. “Local deterministic model of singlet state correlations based on relaxing measurement independence”. In: *Phys. Rev. Lett.* 105 (2010), p. 250404. DOI: 10.1103/PhysRevLett.105.250404.

[28] M. G. A. Paris, G. M. D’Ariano, and M. F. Sacchi. “Maximum-likelihood method in quantum estimation”. In: *AIP Conference Proceedings*. Vol. 568. 2001, pp. 456–467. DOI: 10.1063/1.1381908.

[29] J. Nocedal and S. J. Wright. *Numerical optimization*. New York: Springer, 1999. ISBN: 9780387987934.

[30] P. Virtanen et al. “SciPy 1.0: fundamental algorithms for scientific computing in Python”. In: *Nat. Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

A All Pure Bipartite Entangled States Violate CHSH

Here, we prove that all pure bipartite entangled states violate the CHSH inequality [16]. We begin by rearranging the Schmidt decomposition of a bipartite d-dimensional pure state (Equation 2.24)

$$|\psi\rangle = \sum_{i=1}^{d} c_i |i\rangle|i\rangle = \sum_{j=1}^{d/2} c_{2j-1}|2j-1\rangle|2j-1\rangle + c_{2j}|2j\rangle|2j\rangle \tag{A.1}$$

$$= \sum_{j=1}^{d/2} \sqrt{c_{2j-1}^2 + c_{2j}^2}\Big(\frac{c_{2j-1}}{\sqrt{c_{2j-1}^2 + c_{2j}^2}}|2j-1\rangle|2j-1\rangle + \frac{c_{2j}}{\sqrt{c_{2j-1}^2 + c_{2j}^2}}|2j\rangle|2j\rangle\Big). \tag{A.2}$$

With the measurements acting independently on each Schmidt block

$$A_x = \bigoplus_j \mathbf{a}_{x,j} \cdot \boldsymbol{\sigma}_j, \qquad B_y = \bigoplus_j \mathbf{b}_{y,j} \cdot \boldsymbol{\sigma}_j. \tag{A.3}$$

We aim to recreate a partially entangled state inside this decomposition. Therefore, by defining

$$\cos\theta = \frac{c_{2j-1}}{\sqrt{c_{2j-1}^2 + c_{2j}^2}}, \quad \text{and} \quad \sin\theta = \frac{c_{2j}}{\sqrt{c_{2j-1}^2 + c_{2j}^2}}, \tag{A.4}$$

The state becomes a combination of partially entangled states

$$|\psi\rangle = \sum_{j=1}^{d/2} \sqrt{c_{2j-1}^2 + c_{2j}^2} \left(\cos(\theta)|2j-1\rangle|2j-1\rangle + \sin\theta|2j\rangle|2j\rangle \right) = \sum_{j=1}^{d/2} \sqrt{c_{2j-1}^2 + c_{2j}^2} \, |\psi(\theta_j)\rangle. \tag{A.5}$$

Here, $|\psi(\theta_j)\rangle$ is a partially entangled state with entanglement parameter θ_j. For a non-zero θ, the partially entangled state violates CHSH. Thus, the existence of at least one non-trivial Schmidt pair (i.e., $c_{2j-1}, c_{2j} \neq 0$) corresponding to non-zero θ_j is sufficient to produce CHSH violation. This is satisfied by the assumption of a bipartite entangled state, which requires non-zero c_1 and c_2.

B Separable States Only Generate Local Behavior

A separable bipartite state can be seen as a classical mixture of product states. Consequently, the correlations created by these states are just shared randomness. We show this by calculating the joint probability for a separable bipartite density operator given in Equation 2.21 [11]

$$\begin{aligned} P(a,b|x,y) &= \mathrm{Tr}\left[\rho\,(A_{a|x} \otimes B_{b|y})\right] && \text{(B.1)} \\ &= \mathrm{Tr}\Big[\sum_\lambda q_\lambda (\rho_A^\lambda \otimes \rho_B^\lambda)\,(A_{a|x} \otimes B_{b|y})\Big] && \text{(B.2)} \\ &= \sum_\lambda q_\lambda \,\mathrm{Tr}\left[\rho_A^\lambda A_{a|x}\right] \mathrm{Tr}\left[\rho_B^\lambda B_{b|y}\right] && \text{(B.3)} \\ &= \sum_\lambda q_\lambda P_{a|x}^\lambda P_{b|y}^\lambda. && \text{(B.4)} \end{aligned}$$

We have recovered the local form of the probability distribution given in **Equation 2.30**, suggesting that the separable states generate only the correlation reproducible by an LHV model.

C Strategies for the Conceptual Model

Here we prove that any assignment of predetermined outcomes to the coin in subsection 2.5.1, fails to satisfy both conditions mentioned in that section. The requirement that Alice and Bob always obtain the same outcome whenever they choose the same color implies that any predetermined assignment of outcomes must be identical for both parties. Otherwise, choosing the same color could produce different outcomes. We proceed by investigating the second condition, where Alice and Bob choose different colors. There are three possible pairs:

$$(R,G) \quad (R,B) \quad (B,G)$$

Since the choices are random, each pair occurs with the probability $\frac{1}{3}$. Next consider an assignment of outcomes to these three coins. Since each coin can only take two values, H or T, while there are three coins, at least two coins must share the same value [1]. Consequently, among these three pairs, at least one pair gives the same outcome to both parties. Let's label each assignment of H and T to the coins with a variable λ (for example, $\lambda_1 = (H,T,T)$). Therefore, for each predetermined set of coins corresponding to

[1] pigeonhole principle

a λ, both parties will, with a probability of at least $1/3$, choose a color pair providing the same outcome, it follows that $P(\text{same}|\text{different colors}, \lambda) \geq \frac{1}{3}$. However, the source may generate different assignments λ with different frequencies. For example, the assignment (H,H,H) may occur more often than (T,H,H). We denote the probability distribution for these predetermined assignments by $\rho(\lambda)$. The overall probability of obtaining the same outcome when different colors are chosen is given by

$$P(\text{same}|\text{different colors}) = \sum_{\lambda} \rho(\lambda) P(\text{same}|\text{different colors},\lambda), \quad \text{where} \quad 0 \leq \rho(\lambda) \leq 1. \tag{C.1}$$

But as we discussed, for every individual set $P(\text{same}|\text{different colors},\lambda) \geq \frac{1}{3}$. Since $P(\text{same}|\text{different colors})$ is a convex combination of these individual probabilities, we conclude that

$$P(\text{same}|\text{different colors}) \geq \frac{1}{3} > \frac{1}{4}. \tag{C.2}$$

This result suggests that no predetermined local strategy can satisfy both conditions.

D Alternative Derivation of Tsirelson's Bound

We wish to provide an alternative derivation of the Tsirelson's bound for the CHSH inequality that relies less heavily on operator methods than the proof in subsection 2.5.3. The derivation follows the treatment represented in [13]. By inspecting Equation 2.29, one can rewrite the CHSH operator in the following form

$$\mathbf{S} = \frac{1}{\sqrt{2}} \left(A_0^2 \otimes \mathbb{I}_B + A_1^2 \otimes \mathbb{I}_B + \mathbb{I}_A \otimes B_0^2 + \mathbb{I}_A \otimes B_1^2 \right) - \frac{\sqrt{2}-1}{8} \mathbf{M},$$

where $\mathbf{M}$ is a positive operator expressed by

$$\begin{aligned}
\mathbf{M} = & \left[(\sqrt{2}+1)(A_0 \otimes \mathbb{I}_B - \mathbb{I}_A \otimes B_0) + A_1 \otimes \mathbb{I}_B - \mathbb{I}_A \otimes B_1 \right]^2 && \text{(D.1)} \\
& + \left[(\sqrt{2}+1)(A_0 \otimes \mathbb{I}_B - \mathbb{I}_A \otimes B_1) + A_1 \otimes \mathbb{I}_B - \mathbb{I}_A \otimes B_0 \right]^2 && \text{(D.2)} \\
& + \left[(\sqrt{2}+1)(A_1 \otimes \mathbb{I}_B - \mathbb{I}_A \otimes B_0) + A_0 \otimes \mathbb{I}_B + \mathbb{I}_A \otimes B_1 \right]^2 && \text{(D.3)} \\
& + \left[(\sqrt{2}+1)(A_1 \otimes \mathbb{I}_B + \mathbb{I}_A \otimes B_1) - A_0 \otimes \mathbb{I}_B - \mathbb{I}_A \otimes B_0 \right]^2. && \text{(D.4)}
\end{aligned}$$

Since $\mathbf{M} \geq 0$, the second term is non-positive due to the minus sign. The maximum is attained when this term vanishes, i.e. when $\mathbf{M} = 0$. Consequently,

$$\begin{aligned}
\|\mathbf{S}\|_\infty & \leq \left\| \frac{1}{\sqrt{2}} \left(A_0^2 \otimes \mathbb{I}_B + A_1^2 \otimes \mathbb{I}_B + \mathbb{I}_A \otimes B_0^2 + \mathbb{I}_A \otimes B_1^2 \right) \right\|_\infty && \text{(D.5)} \\
& \leq \frac{1}{\sqrt{2}} \left(\left\|A_0^2 \otimes \mathbb{I}_B\right\|_\infty + \left\|A_1^2 \otimes \mathbb{I}_B\right\|_\infty + \left\|\mathbb{I}_A \otimes B_0^2\right\|_\infty + \left\|\mathbb{I}_A \otimes B_1^2\right\|_\infty \right). && \text{(D.6)}
\end{aligned}$$

Additionally, for projective observables with eigenvalues ± 1, we have $A_x^2 = B_y^2 = \mathbb{I}$, and hence $\left\|A_x^2\right\|_\infty = \left\|B_y^2\right\|_\infty = 1$ [1], leading to

$$\|\mathbf{S}\|_\infty \leq \frac{1}{\sqrt{2}} \left(\left\|A_0^2\right\|_\infty + \left\|A_1^2\right\|_\infty + \left\|B_0^2\right\|_\infty + \left\|B_1^2\right\|_\infty \right) = 2\sqrt{2}. \tag{D.7}$$

[1] For a more general proof without the assumption of projective observables see [13]

E Horodecki Criterion for Partially Entangled States

Partially entangled states can produce a CHSH violation as high as $S_{\max} = 2\sqrt{1+\sin^2(2\theta)}$, where θ is the entanglement parameter. However, to prove this, we begin by proving the Horodecki Criterion. Consider the density matrix for a two-qubit state (**Equation 2.17**) where the correlation matrix elements are given by **Equation 2.18**. If we denote the two largest eigenvalues of $T^{\mathrm{T}}T$ by λ_1 and λ_2, then the Horodecki Criterion states that the state violates the CHSH inequality if and only if:

$$\lambda_1 + \lambda_2 > 1, \tag{E.1}$$

and the maximal CHSH violation achievable by this state is given by

$$S_{\max} = 2\sqrt{\lambda_1 + \lambda_2}. \tag{E.2}$$

To construct the Bell parameter, we find an expression for $\langle A_x B_y \rangle$ using the Equation 2.20

$$\begin{aligned}
\langle A_x B_y \rangle &= \sum_{a,b} (-1)^{a+b} P(a,b|x,y) && \text{(E.3)}\\
&= \sum_{a,b} (-1)^{a+b} \frac{1}{4}\Big(1 + (-1)^a\, \mathbf{a}_x \cdot \mathbf{r} + (-1)^b\, \mathbf{b}_y \cdot \mathbf{s} + (-1)^{a+b}\, \mathbf{a}_x^{\mathrm{T}}\, T\, \mathbf{b}_y\Big) && \text{(E.4)}\\
&= \sum_{a,b} \frac{1}{4}\Big((-1)^{a+b} + (-1)^b\, \mathbf{a}_x \cdot \mathbf{r} + (-1)^a\, \mathbf{b}_y \cdot \mathbf{s} + \mathbf{a}_x^{\mathrm{T}}\, T\, \mathbf{b}_y\Big). && \text{(E.5)}\\
& && \text{(E.6)}
\end{aligned}$$

It is evident that the first three terms average to zero, leading to

$$\langle A_x B_y \rangle = \mathbf{a}_x^{\mathrm{T}}\, T\, \mathbf{b}_y. \tag{E.7}$$

Therefore, the CHSH parameter (Equation 2.29) can be rewritten as

$$S = \mathbf{a}_0^{\mathrm{T}}\, T\, \mathbf{b}_0 + \mathbf{a}_0^{\mathrm{T}}\, T\, \mathbf{b}_1 + \mathbf{a}_1^{\mathrm{T}}\, T\, \mathbf{b}_0 - \mathbf{a}_1^{\mathrm{T}}\, T\, \mathbf{b}_1 = \mathbf{a}_0^{\mathrm{T}} T(\mathbf{b}_0 + \mathbf{b}_1) + \mathbf{a}_1^{\mathrm{T}} T(\mathbf{b}_0 - \mathbf{b}_1). \tag{E.8}$$

Thus, the expression consists of two inner products, which gain their maximum value when the pair of vectors are aligned. In other words, the best strategy for Alice is to align her measurment vectors $\mathbf{a}_0$ and $\mathbf{a}_1$ with $T\,(\mathbf{b}_0 + \mathbf{b}_1)$ and $T\,(\mathbf{b}_0 - \mathbf{b}_1)$ respectively. Therefore

$$\max_{\mathbf{a}_0}[\mathbf{a}_0 \cdot T\,(\mathbf{b}_0 + \mathbf{b}_1)] = |T\,(\mathbf{b}_0 + \mathbf{b}_1)|, \quad \text{and} \quad \max_{\mathbf{a}_1}[\mathbf{a}_1 \cdot T\,(\mathbf{b}_0 - \mathbf{b}_1)] = |T\,(\mathbf{b}_0 - \mathbf{b}_1)|. \tag{E.9}$$

$$\max_{\mathbf{a}_0,\mathbf{a}_1} S = |T\,(\mathbf{b}_0 + \mathbf{b}_1)| + |T\,(\mathbf{b}_0 - \mathbf{b}_1)| \tag{E.10}$$

To further simplify, we define two pairs of mutually orthogonal unit vectors $\hat{c}_0$ and $\hat{c}_1$ such that

$$\mathbf{b}_0 + \mathbf{b}_1 = 2\cos\theta\, \hat{c}_0, \qquad \mathbf{b}_0 - \mathbf{b}_1 = 2\sin\theta\, \hat{c}_1, \quad \text{where} \qquad \theta \in \left[0, \frac{\pi}{2}\right]. \tag{E.11}$$

Subsequently, we need to carry out the following maximization

$$\max_{\theta,\hat{c}_0,\hat{c}_1} 2(\cos\theta|T\cdot\hat{c}_0| + \sin\theta|T\cdot\hat{c}_1|) \tag{E.12}$$

We rearrange the expression

$$\max_{\theta,\hat{c}_0,\hat{c}_1} 2\sqrt{|T\cdot\hat{c}_0|^2+|T\cdot\hat{c}_1|^2}(\cos\theta\sin\alpha+\sin\theta\cos\alpha) = \Big(\cos\theta\frac{|T\cdot\hat{c}_0|}{\sqrt{|T\cdot\hat{c}_0|^2+|T\cdot\hat{c}_1|^2}}+\sin\theta\frac{|T\cdot\hat{c}_1|}{\sqrt{|T\cdot\hat{c}_0|^2+|T\cdot\hat{c}_1|^2}}\Big). \tag{E.13}$$

This allows us to define a helper angle $\alpha_{\hat{c}_0,\hat{c}_1}$ such that

$$\cos\alpha_{\hat{c}_0,\hat{c}_1} = \frac{|T\cdot\hat{c}_1|}{\sqrt{|T\cdot\hat{c}_0|^2+|T\cdot\hat{c}_1|^2}}, \quad \text{and} \quad \sin\alpha_{\hat{c}_0,\hat{c}_1} = \frac{|T\cdot\hat{c}_0|}{\sqrt{|T\cdot\hat{c}_0|^2+|T\cdot\hat{c}_1|^2}} \tag{E.14}$$

The simplified expression becomes

$$\max_{\theta,\hat{c}_0,\hat{c}_1} 2\sqrt{|T\cdot\hat{c}_0|^2+|T\cdot\hat{c}_1|^2}(\cos\theta\sin\alpha_{\hat{c}_0,\hat{c}_1}+\sin\theta\cos\alpha_{\hat{c}_0,\hat{c}_1}) = \max_{\theta,\hat{c}_0,\hat{c}_1} 2\sqrt{|T\cdot\hat{c}_0|^2+|T\cdot\hat{c}_1|^2}\sin\big(\alpha_{\hat{c}_0,\hat{c}_1}+\theta\big) \tag{E.15}$$

This expression acquires its maximum value at $\theta = \frac{\pi}{2} - \alpha_{\hat{c}_0,\hat{c}_1}$, consequently

$$\max_{\hat{c}_0,\hat{c}_1} S = \max_{\hat{c}_0,\hat{c}_1} 2\sqrt{|T\cdot\hat{c}_0|^2+|T\cdot\hat{c}_1|^2} = \max_{\hat{c}_0,\hat{c}_1} 2\sqrt{\hat{c}_0\cdot T^{\mathrm{T}}T\cdot\hat{c}_0+\hat{c}_1\cdot T^{\mathrm{T}}T\cdot\hat{c}_1} = 2\sqrt{\lambda_1+\lambda_2} \tag{E.16}$$

In the last step we chose $\hat{c}_0$ and $\hat{c}_1$ as two eigenvectors of $T^\dagger T$ corresponding to its two largest eigenvalues λ_1 and λ_2.

For a partially entangled state, according to **Equation 2.27**, we have $T^T T = \mathrm{diag}\big(\sin^2(2\theta),\ \sin^2(2\theta),\ 1\big)$, thus $\lambda_1 = 1$ and $\lambda_2 = \sin^2 2\theta$. Therefore, the maximal CHSH violation for a partially entangled state is given by

$$S_{\max} = 2\sqrt{1+\sin^2 2\theta}. \tag{E.17}$$

F Tables

λ	(a_0, a_1, b_0, b_1)	q_λ
0	(0,0,0,0)	0.000000000000
1	(0,0,0,1)	0.366267385926
2	(0,0,1,0)	0.000000000000
3	(0,0,1,1)	0.000000000007
4	(0,1,0,0)	0.366270135013
5	(0,1,0,1)	0.000000000010
6	(0,1,1,0)	0.133732223754
7	(0,1,1,1)	0.000000000008
8	(1,0,0,0)	0.000000000000
9	(1,0,0,1)	0.133730255726
10	(1,0,1,0)	0.000000000000
11	(1,0,1,1)	0.000000000009
12	(1,1,0,0)	0.000000000007
13	(1,1,0,1)	0.000000000008
14	(1,1,1,0)	0.000000000009
15	(1,1,1,1)	0.000000000010

Table F.1: Deterministic strategies $\lambda = (a_0, a_1, b_0, b_1)$ and corresponding probabilities q_λ for $R = 0.8$.