
NATIONAL UNIVERSITY OF SINGAPORE

Department of Physics

# *Certification of Quantum Entanglement*

Author:
Goh Qi Xuan Benjamin

Supervisor: Prof Valerio Scarani
Mentor: Goh Koon Tong

*Thesis submitted in partial fulfillment of the requirements for the degree of Bachelor of Science (Honours)*

April 2017

# Abstract

Some entanglement witnesses in the literature rely on the assumptions of state tomography – the complete knowledge of the measurement device characteristics as well as the Hilbert space dimension of the systems being measured. On the other extreme end are the Device-Independent (DI) entanglement witnesses, which rely only on the observed correlations to certify the presence of entanglement without any further assumptions. However, such entanglement witnesses give pessimistic bounds on the amount of certifiable entanglement and are restricted to correlations that violate a Bell inequality.

There are also schemes which are considered "semi-device independent" which keep some of the assumptions of state tomography while relaxing others. One such relaxation involves only assuming the dimension of the Hilbert space of the systems, and has been tested with ideal statistics [7], certifying entanglement for correlations that do not violate a Bell inequality and thus cannot be certified in a DI fashion. This project addresses the question of quantifying the amount of entanglement given experimental correlations and only the knowledge of the Hilbert space dimension, accounting for the experimental fluctuations arising from non-ideal detectors as well as finite sample size. The performance of the semi-DI scheme is then compared against other numerical methods that certify entanglement in a DI way [13], as well as in the case when full characterization of the setup is assumed.

# Acknowledgments

This project has been a one-of-a-kind journey, and I would like to take this opportunity to express my gratitude to those who have been part of it. I would first like to thank Professor Scarani, who always gave insight, encouragement and guidance to add value to this project and keep the time I have spent in learning and contributing to discussions enjoyable. I am thankful to him also for the reason that he has given me the opportunity to work in a friendly and productive environment in CQT.

I would also like to take this chance to thank my mentor, Goh Koon Tong, for his patience and commitment to ensuring I am on track with my progress and in particular, pointing me to the resources which helped my understanding of the concepts, taking me through the most uphill part of this experience.

Next, kudos to Alessandro and Hoh Shun, who provided the experimental data on which I performed my numerical tests, and for kindly taking out time to explain the data processing to me, as their data was instrumental to the execution of this project.

In addition, I would like to express my thanks to Cai Yu, who in our numerous 'office' conversations has given me boosts of inspiration and quantum (no pun intended) leaps in understanding. His willingness in answering my queries and engaging my curiosity should not go unnoticed.

Finally, I would like to thank my family and friends, in church and outside, as well as my soulmate Grace, who have given me moral support and have always been looking out for me. Last but not least, I want to thank my pet dog Snowy, who always joins me enthusiastically during my morning exercises in preparation to start the day and face new challenges.

# Contents

# Chapter 1

## Introduction

Entanglement is a valuable resource in many applications of quantum mechanics and quantum information science. The certification of entanglement thus becomes useful in applications where the serviceability of the quantum devices employed needs to be ensured. These applications include but are not limited to quantum key distribution and quantum random number generation.

When a precise characterization of the measuring equipment is available, one can use the experimental set-up to certify the presence of entanglement, usually with the use of some entanglement witness, that can be found in a system. Such entanglement witnesses usually rely on knowledge of the measurements being performed and the Hilbert Space dimension of the system under study. For example, $\langle XX \rangle + \langle YY \rangle > 1$ is an entanglement witness, provided that the system under study is indeed bipartite qubit and the measurement settings are exactly in the X and Y bases. If either of these assumptions are violated, the observed correlations may give a false positive.

In applications related to quantum communication, there exist adversarial scenarios where a description of the measurement devices of some communicating party is not available or trusted. In these cases, it is possible to quantify entanglement solely from the observed correlations, provided the correlations violate some Bell Inequality. These entanglement witnesses are known as Device-Independent (DI) entanglement witnesses.

It should therefore be noted that the certification of entanglement depends on the level of characterization of the devices. DI certification of entanglement requires virtually no assumptions. However, DI certification schemes always require the violation of some Bell Inequality, and computational methods using semi-definite programming (such as the NPA Hierarchy) usually give pessimistic bounds on the amount of certifiable entanglement. In addition, DI schemes are also experimentally demanding, since the experimenter has to ensure that the observed Bell violation is loophole-free. To date, only three loophole-free experimental realizations have been performed. On the other hand, with the knowledge of the Hilbert Space dimension and the measurement settings, one has the same

assumptions as state tomography, which is versatile and has been routinely implemented in experiments. State tomography, in principle, allows for a full reconstruction of the state, which comes at the price of “trusting” the dimension and measurement settings.

In view of the trade-off above, it is thus useful to consider a Semi-DI approach, where the dimension is known, but the measurement settings are not. This approach was first taken by Moroder and Gittsovich (2012)[11], who discussed how to certify the presence of entanglement and provided analytical bounds on the correlations for some cases. Later work by Goh, Bancal and Scarani (2016) [7] explores the minimal amount of certifiable entanglement (in terms of the concurrence) for a given set of correlations and the knowledge of the dimension for the case of bipartite qubits. In the project, I explore the scheme proposed in Goh et al (2016) with a modification that accounts for experimental fluctuations, and implement the scheme on data obtained from experiments on entangled photon pairs generated by Type II crystals. I determine the minimum amount of entanglement certifiable using the negativity of the state, given a set of experimental correlations with fluctuations arising from non-ideal detectors and finite sample size.

# Chapter 2

# Preliminaries

This chapter provides a summary of the important notions for the project. There are well-established notions, and I have synthesized them here as part of the summary of the project repertoire and for an uninitiated reader (which was where I began) to grasp the basic operational notions.

## 2.1 Qubits

### 2.1.1 A Qubit

In QIS, a qubit is a unit of quantum information. It can be thought of as a quantum analog of a (classical) bit, which encodes either a 1 or a 0. Any two-level quantum system, such as a single photon polarization or electron spin, can be used as a physical realization of a qubit.

Unlike its classical analog, a qubit does not necessarily have to either be in the '0' state or the '1' state. Qubits are allowed to be in a superposition of the two. A qubit state can then be represented by a state vector, with basis states that a state can be measured in. One such basis would be the “standard” basis:

A general pure qubit state can be written as:

$$|\psi\rangle = e^{i\Omega}(a|0\rangle + e^{-i\phi}\sqrt{1-a^2}|1\rangle) \tag{1}$$

for some real numbers a,b, Ω and φ, bearing in mind that quantum states are unique up to a global phase and normalization. For any state |ψ>, we can also consider the projector onto that state

$$\begin{aligned} |\psi\rangle\langle\psi| = a^2|0\rangle\langle 0| + (1-a^2)|1\rangle\langle 1| \\ + a\sqrt{1-a^2}\left(e^{i\phi}|0\rangle\langle 1| + e^{-i\phi}|1\rangle\langle 0|\right) \end{aligned} \tag{2}$$

The reason it is called a projector is due to its projection property:

$$(|\psi\rangle\langle\psi|)|\phi\rangle = (\langle\psi|\phi\rangle) \cdot |\psi\rangle \quad (3)$$

The vector space that the state resides in is known as the Hilbert Space, and qubits have Hilbert space dimension 2, or alternatively can be said to reside in $\mathbb{C}^2$ . Measurements that can be made on the state can be represented by their measurement operators as matrices in $\mathbb{C}^2$. The measurement operators, denoted PI(x,a), have the following properties:

- For each measurement setting x and output a, $\Pi^x_a$ is the associated measurement operator.
- For each setting, the measurement operators are complete:

$$\sum_a \Pi^x_a = \mathbb{I} \quad (4)$$

- All the measurement operators $\Pi^x_a$ are positive semi-definite.

When the measurement operators are mutually orthogonal, the measurement is projective (and they take the form of a projector onto some subspace of the Hilbert Space of the system). Some measurements are not projective, and the most general measurement can be defined using Positive Operator Valued Measurements (POVMs). They similarly satisfy the properties for measurement operators as above, with the relaxation that they may have full rank, and may not be trace 1 unlike projective measurements.

For projective measurements on $\mathbb{C}^2$, their corresponding observables M can be constructed in the following form:

$$M = \lambda_1|\psi\rangle\langle\psi| + \lambda_2|\psi_\perp\rangle\langle\psi_\perp| \quad (5)$$

The above form is also known as the spectral decomposition of M, since $|\psi\rangle$ and $|\psi_\perp\rangle$ are orthogonal. The operator M corresponds to the physical observable, which may be spin or polarization in the direction specified by the state $|\psi\rangle$ . The operators M are typically constructed with the following properties:

1. They are Hermitian, which means that their eigenvectors are mutually orthonormal and span the Hilbert space in which they reside. Consequently,

the operator can be decomposed in terms of its eigenvectors, which corresponds to the basis in which the measurement of the associated observable takes place. Its eigenvalues are real, and are assigned to be +1 and –1.

2. They are unitary. As a consequence of their construction, measurement operators satisfy $M^2 = \mathbb{1}$. Along with being Hermitian, this means they satisfy $M^\dagger M = \mathbb{1}$.

3. They are traceless. Since their eigenvalues are +1 and –1, they must also satisfy tr(M) = 1 – 1 = 0.

Of course, the matrix representation of the observable depends on the choice of measurement basis. If the representation is constructed using the three following bases:

$$|0\rangle \widehat{=} \begin{pmatrix}1\\0\end{pmatrix}, |1\rangle \widehat{=} \begin{pmatrix}0\\1\end{pmatrix}, |\pm\rangle \widehat{=} \frac{1}{\sqrt{2}}\begin{pmatrix}1\\\pm 1\end{pmatrix}, |L\rangle \widehat{=} \frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}, |R\rangle \widehat{=} \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}$$

$$\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1| \widehat{=} \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} - \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0 & 1\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix} \quad (6)$$

$$\sigma_x = |+\rangle\langle +| - |-\rangle\langle -| \widehat{=} \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1 & 1\end{pmatrix} - \frac{1}{2}\begin{pmatrix}1\\-1\end{pmatrix}\begin{pmatrix}1 & -1\end{pmatrix} = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix} \quad (7)$$

$$\sigma_y = |L\rangle\langle L| - |R\rangle\langle R| \widehat{=} \frac{1}{2}\begin{pmatrix}1\\i\end{pmatrix}\begin{pmatrix}1 & -i\end{pmatrix} - \frac{1}{2}\begin{pmatrix}1\\-i\end{pmatrix}\begin{pmatrix}1 & i\end{pmatrix} = \begin{pmatrix}0 & -i\\i & 0\end{pmatrix} \quad (8)$$

We obtain the well known Pauli matrices. In this text, the three bases will sometimes be referred to as the computational or “standard” basis, the “plus-minus” or “diagonal” basis, and the “circular” basis respectively. This alludes to their assignment to the respective polarization states for photons. In the cases of electron spin, they are the angular momentum eigenstates along the Z, X and Y axes respectively. In some presentations, the Pauli matrices are denoted as Z, X and Y respectively.

These matrices satisfy the above three properties and also are all unitarily

equivalent to each other. They satisfy the trace orthonormality relation:

$$\mathrm{tr}(\sigma_i \sigma_j) = 2\delta_{ij} \tag{9}$$

### 2.1.2 Bloch Sphere

Pure States

For the case of qubits, the projectors for the pure states have one-to-one correspondence to points on a unit sphere. This can be seen by considering a qubit pure state, and decomposing it in terms of the Pauli matrices.

$$|\psi\rangle\langle\psi| \widehat{=} \frac{1}{2}\begin{pmatrix} 1+\cos\theta & e^{-i\phi}\sin\theta \\ e^{i\phi}\sin\theta & 1-\cos\theta \end{pmatrix} \tag{10}$$

$$= \frac{1}{2}(\mathbb{I} + \sigma_x \sin\theta\cos\phi + \sigma_y \sin\theta \sin\phi + \sigma_z \cos\theta) \tag{11}$$

$$|\psi\rangle\langle\psi| = \frac{1}{2}(\mathbb{I} + \vec{n}\cdot\vec{\sigma}) \tag{12}$$

The parameters φ and θ can then be interpreted as the azimuthal and polar coordinate respectively, and **n** denotes the Bloch vector. Each pair of anti-nodal points on the unit sphere $S_2$ corresponds to pair of mutually orthonormal basis states: so the pair of points on the z-axis (θ = 0) correspond to the standard basis, while the pair on the x-axis (θ = π/2, φ = 0) corresponds to the diagonal basis, while the pair on the y-axis (θ = π/2, φ = π/2) corresponds to the circular basis. The matrix representation of the observable can similarly be expressed using 2 degrees of freedom.

$$\vec{n}\cdot\vec{\sigma} \widehat{=} \begin{pmatrix} \cos\theta & e^{-i\phi}\sin\theta \\ e^{i\phi}\sin\theta & -\cos\theta \end{pmatrix} \tag{13}$$

Mixed states

In general, a state can also be a statistical ensemble of multiple pure states. In such a situation, a ket vector is no longer sufficient for the description of the state, but rather necessitates the density operator:

$$\rho = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} \tag{14}$$

$$\rho = \sum_i p_i |i\rangle\langle i| \tag{15}$$

It should be noted that the states |i> need not be orthogonal. A density operator ρ must be positive semi-definite, trace 1 and must be an element of a linear operator on $\mathbb{C}^2$. For n-dimensional systems, the density operator can be represented by an n-by-n matrix.

For qubits, the density operator can also be written in terms of the Pauli matrices, with the same form as for projectors of pure states. In the case of pure states, the Bloch vector has magnitude 1. For mixed states, the Bloch vector has magnitude less than 1. It is instructive to consider the following examples of density operators.

$$\rho_1 = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = |+\rangle\langle +| = \frac{1}{2}(\mathbb{I} + \sigma_x) \tag{16}$$

$$\rho_2 = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} = \frac{1}{2}|1\rangle\langle 1| + \frac{1}{2}|-\rangle\langle -| = \frac{1}{2}(\mathbb{I} + \frac{1}{2}\sigma_x + \frac{1}{2}\sigma_z) \tag{17}$$

$$\rho_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}|\psi\rangle\langle\psi| + \frac{1}{2}|\psi_\perp\rangle\langle\psi_\perp| = \frac{1}{2}\mathbb{I} \tag{18}$$

In addition, every density operator admits a spectral decomposition. i.e. For any density operator, a decomposition of the form:

$$\rho = \sum_i \lambda_i |i\rangle\langle i| \tag{19}$$

where the state vectors |i> are mutually orthonormal, is always possible.

### 2.1.3 Multiple Qubits

In cases where the system comprises more than one qubit, one would consider the state to reside in the composite Hilbert space $\mathbb{C}^{2\otimes n}$. In the case of two qubits, each one in a pure state, one would write the bipartite state as a tensor product of the two states. For example, if one qubit is in the |0> state, while the other is in the |+>

state, we have

$$|\Psi\rangle = |0\rangle \otimes |+\rangle = |0\rangle|+\rangle \tag{20}$$

Composite states can also be in superposition of any tensor product combination over the two subsystems. For example, we can consider a superposition of two qubits both being in the |0> state and both being in the |1> state.

$$|\Phi^+\rangle = \tfrac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle) = \tfrac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{21}$$

In terms of the Bloch vectors, bipartite qubit states have the following form:

$$\rho_{AB} = \tfrac{1}{4}(\mathbb{I} \otimes \mathbb{I} + \vec{n}_A \cdot \vec{\sigma} \otimes \mathbb{I} + \mathbb{I} \otimes \vec{n}_B \cdot \vec{\sigma} + \vec{n}_A \cdot T \cdot \vec{n}_B) \tag{22}$$

The notion of composite states can be extended to three or more qubits, as well as higher dimensional systems.

Mixed states arise when considering a bipartite state in a composite Hilbert space $H_1 \otimes H_2$ (say $\mathbb{C}^2 \otimes \mathbb{C}^2$), and one half of the composite system is not accessible to the observer. The state in one of the Hilbert spaces $H_1$ is then the partial trace of the composite state over $H_2$. For example, we consider again $|\Phi^+\rangle$ as denoted in (21):

$$|\Phi^+\rangle\langle\Phi^+| = \tfrac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \tag{23}$$

$$\mathrm{tr}_B(|\Phi^+\rangle\langle\Phi^+|) = \tfrac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{24}$$

It is also known that any mixed state in Hilbert space H can be seen as a reduced state of some higher dimensional pure state in $H \otimes H$ . This is known as state purification, and can be summarized as follows:

$$\begin{aligned} \forall \rho_A &= \sum_i \rho_{ii} |i\rangle\langle i| \\ \exists\ |\Psi_{AB}\rangle &= \sum_{i,i'} \sqrt{\rho_{ii}} |i\rangle_A |i'\rangle_B \\ s.t.\ \mathrm{tr}_B(|\Psi_{AB}\rangle\langle\Psi_{AB}|) &= \rho_A \end{aligned} \tag{25}$$

A simple criterion for determining if a state ρ is mixed is to take the trace of $\rho^2$. If $\mathrm{tr}(\rho^2) = 1$, the state ρ is pure, and otherwise it is mixed.

### 2.1.4 Statistics obtained by measurements

It would be useful to consider the statistics obtained by measurements on quantum states. For a state p, the expectation value ⟨A⟩ of an observable A is given by tr(ρ A). For a pure state |ψ⟩, this expectation value reduces to ⟨ψ|A|ψ⟩. In state tomography, many copies of the state are measured with the appropriate measurement operators to obtain their expectation values (known as moments), and the moments are used to reconstruct the state. For a qubit, these measurement operators will be the Pauli matrices or linear combinations thereof.

The fact that the Pauli matrices are traceless, unitary and Hermitian allows for a very neat property for measurement statistics on qubits: The expectation values for any observable $\mathbf{n}\cdot\boldsymbol{\sigma}$ for a state $\frac{1}{2}(\mathbb{1} + \mathbf{m}\cdot\boldsymbol{\sigma})$, will simply depend on the angle between the associated Bloch vectors. In particular,

$$\langle \mathbf{n}\cdot\boldsymbol{\sigma} \rangle = \mathrm{tr}(\tfrac{1}{2}(\mathbb{1} + \mathbf{m}\cdot\boldsymbol{\sigma})\ \mathbf{n}\cdot\boldsymbol{\sigma}) = \mathbf{m}\cdot\mathbf{n} \tag{26}$$

$$P(\pm|\text{measurement setting } \pm\mathbf{n}\cdot\boldsymbol{\sigma}, \text{ state } \tfrac{1}{2}(\mathbb{1} + \mathbf{m}\cdot\boldsymbol{\sigma})) = \tfrac{1}{2}(1 \pm \mathbf{m}\cdot\mathbf{n}) \tag{27}$$

In experiments, because of the finite sample size and non-ideality of the detectors, fluctuations in the data are expected. In the case of quantum state estimation (QSE), fidelity may be used as a figure of merit that quantifies the quality of the estimate[10]. Fidelity can be understood as the closeness of two states, and is defined as the following:

$$\begin{aligned} F(\rho, \sigma) &= \mathrm{tr}(\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}) \\ &= \sqrt{\langle\psi|\rho|\psi\rangle} \quad \text{if one state is pure} \end{aligned} \tag{28}$$

$$= |\langle\psi|\phi\rangle| \qquad \text{if both states are pure}$$

where the square root of a matrix is defined using the spectral decomposition theorem. For pure states, the fidelity is equivalent to the absolute value of the inner product of the two state vectors, sometimes also called the overlap of the two states.

In cases where the state is composite, and a local measurement is performed on each partition, the expectation values obtained will be:

$$\widehat{A} = \bigotimes_i \widehat{A}_i$$

$$\langle\widehat{A}\rangle = \mathrm{tr}(\rho \bigotimes_i \widehat{A}_i) \tag{29}$$

$$= \langle\psi|\bigotimes_i \widehat{A}_i|\psi\rangle \qquad \text{if state is pure} \tag{30}$$

For bipartite qubits, this expression reduces to $\langle A\rangle = \mathrm{Tr}(\rho\, A_1 \otimes A_2)$. At this point it is also useful to note that the expectation value $\langle A\rangle$ is invariant under a local unitary transformation on ρ and the measurements.[1]

### 2.1.5 Qutrits and higher dimensions

For higher dimensional systems, there exist generalizations of the Pauli matrices and bloch sphere. For example, a 3-dimensional quantum system, also known as a qutrit, can be expressed as a state vector

$$|\psi\rangle = a|0\rangle + b|1\rangle + c|2\rangle \tag{31}$$

where $|a|^2 + |b|^2 + |c|^2 = 1$. Analogously to qubits, projectors and density operators for qutrits are represented by 3-by-3 matrices. Pure states and mixed states can be distinguished in the same way that was described earlier by computing the trace of the square of the density operator. For qutrits, the analog to the Paul matrices are the Gell-Mann Matrices:

1 Of course, the local unitary transformation needs to be the same for both the state and measurements. To see why, apply local unitary transformations on the state and measurements and use the cyclic property of the trace to eliminate the matrices using $U^\dagger U = 1$.

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$
$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

There is a simple way to construct analogous matrices for higher dimensions as well. These matrices are Hermitian and traceless, and also satisfy the trace orthonormality relation $\mathrm{tr}(\lambda_i\lambda_j) = 2\delta_{ij}$. Along with the identity matrix, form a complete basis for any 3x3 Hermitian matrix. Indeed, it is known that any qutrit state can be represented as

$$\rho = \frac{1}{3}(\mathbb{I} + \vec{n} \cdot \vec{\lambda}) \tag{32}$$

for an eight-dimensional vector **n**, where $|\mathbf{n}|^2 \leq 3$. Unlike qubits, qutrits do not enjoy the same symmetries in the evaluations of expectation values stated earlier for the reason that Hermitian matrices in $\mathbb{C}^3$ cannot be unitary and traceless simultaneously. An easy way to see this is to consider the spectral decomposition of a 3x3 matrix analogously to (5)

$$M = \lambda_1|\psi_1\rangle\langle\psi_1| + \lambda_2|\psi_2\rangle\langle\psi_2| + \lambda_3|\psi_3\rangle\langle\psi_3| \tag{33}$$

if the eigenvalues are real and M is unitary, then $\lambda^2 = 1$. However, no permutation of ±1 between the three of them will satisfy the traceless requirement $\lambda_1 + \lambda_2 + \lambda_3 = 0$.

Other generalizations of Pauli Matrices

There also exist other generalizations of Pauli matrices. These have certain desirable properties such as unitarity, correspondence to the Pauli matrices, generalization to all dimensions, while trading off other properties such as Hermiticity. For example, for the generalization of $\sigma_z$, there is the clock operator:

$$Z = \sum_{i=0}^{d-1} \chi^i |i\rangle\langle i| \tag{34}$$

where χ is the d-th root of unity.

In the case of $\sigma_x$, the generalization is known as the shift operator:

$$X = \sum_{i=0}^{d-1} |i\rangle\langle i+1| \tag{35}$$

where i is evaluated modulo d.

Generalized Bloch sphere

The form (32) suggests the existence of an 8-dimensional vector space into which a qutrit density operator can be decomposed. Indeed, for any qudit (d-dimensional quantum system), there exist methods to construct a Generalized Bloch Sphere. The boundary of the Bloch sphere for any dimension can be determined by computing the positivity condition from the characteristic polynomial of the density operator, expressed in terms of the Bloch vector. The boundaries are known to be closed surfaces, and their geometries depend on the choice of basis used to decompose the density operators.

For qubits, the positivity condition is the following when decomposed in terms of the Pauli matrices:

$$\det(\rho) = \frac{1}{4}\begin{vmatrix} 1+n_z & n_x - in_y \\ n_x + in_y & 1-n_z \end{vmatrix} = \frac{1-(n_x^2+n_y^2+n_z^2)}{4} \geq 0$$

which constrains a valid Bloch vector for qubits to the 3-dimensional sphere of radius 1. In this case, all points in the sphere correspond to a valid qubit state. For qutrits, the Bloch vector for pure states (using the G matrices as a basis) has magnitude $\sqrt{3}$, but the positivity condition is more nuanced. The characteristic polynomial is given by:

$$\lambda^3 - \lambda^2 + \frac{3-\sum n_i^2}{9}\lambda - \det(\rho)$$

From the characteristic polynomial, it can be seen that $|\mathbf{n}|^2 \leq 3$ is still a necessary condition for the matrix to represent a valid state (consider the coefficient of $\lambda^1$). However, it is not a sufficient condition, and some density operators represented by points on the $|\mathbf{n}|^2 = 3$ hypersphere do not correspond to physical states. An explicit example can be found by considering the anti-nodal point of a pure state. If $|\psi\rangle\langle\psi| = 1/3(\mathbb{1} + \mathbf{n}\cdot\boldsymbol{\lambda})$, then the eigenvalues of $\mathbf{n}\cdot\boldsymbol{\lambda}$ must be 2, –1 and –1. The anti-nodal point, which corresponds to $1/3(\mathbb{1} - \mathbf{n}\cdot\boldsymbol{\lambda})$, will then represent an indefinite matrix since its eigenvalues will be 2/3, 2/3 and –1/3.

For this reason, the 8-dimensional Generalized Bloch sphere for qutrits is not a true hypersphere. In fact, it is known that no 3-D slice of the qutrit GBS is a sphere of radius $\sqrt{3}$.[8] A study characterizing the 2-D and 3-D slices of the qutrit GBS also determined that wide classes of 3-D slices have similar geometries. These geometries include spheres (not of radius $\sqrt{3}$, but radius 1), prolate ellipsoids, cones, and obese tetrahedrons[8].

An observation: Rank of the density operators on the Bloch Sphere

An elementary property of the density operator is its rank. In the case of qubits, every point on the boundary of the Bloch Sphere corresponds to a rank 1 density operator, while every point in the interior corresponds to a rank 2 operator. In higher dimensions, it can be seen by considering the eigenvalues, that singular matrices must lie on the boundary of the GBS, so all points in the interior of the qutrit GBS correspond to rank 3 density operators. It can also be shown that for any dimension, points on the boundary must necessarily represent singular density matrices, thus making the singularity of the density matrix a necessary and sufficient condition to characterize the GBS for any dimension. As it turns out, points on the boundary of the qutrit GBS not at a distance of $\sqrt{3}$ from the origin correspond to rank 2 density operators, while points on the boundary at a distance of $\sqrt{3}$ are rank 1 density operators (i.e. pure states).

In Appendix A, I describe an easy method that exploits this observation to plot 3-D sections of the Generalized Bloch sphere for qutrits (or any dimension, for that matter) numerically using MATLAB.

## 2.2 Entanglement

### 2.2.1 An Entangled State

In the quantum model, a quantum state is said to be *entangled* when the state cannot be written as a (mixture of) tensor product of states in the respective Hilbert Spaces of each sub-system. For pure states, a simple method to determine if the state is entangled is to consider the partial trace over one subsystem to obtain the reduced density operator. If the trace of the square of the reduced density operator is less than 1, the original state has to be entangled.

Recall the earlier example in (21). The state $|\Phi^+\rangle$ is a *maximally entangled state* of two qubits. The state can be written as:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle) = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

If one were to take the partial trace over one subsystem (say B), one will obtain the *maximally mixed state* of a single qubit as in (18).

$$\mathrm{tr}_B(|\Phi^+\rangle\langle\Phi^+|) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The trace of the square of the reduced density operator can then be shown to be less than one:

$$\mathrm{tr}(\tfrac{1}{4}\mathbb{I}^2) = \frac{1}{2} < 1$$

One notable feature about entangled states are the correlations exhibited when the appropriate measurements are made. Consider the following scenario:

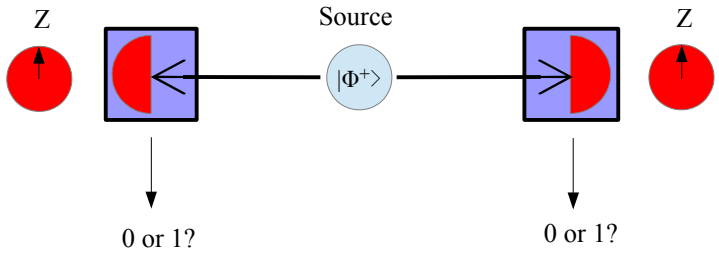



Figure 2.2.1.1: Two devices each measure one half of a maximally entangled pair of qubits, and output a measurement outcome in the form of a bit.

Let two photons be separated, and an observer at each point (call them Alice and Bob) makes a measurement on the standard (Z) basis. If the two photons are in the state $|\Phi^+\rangle$, whenever Alice obtains the outcome 0, Bob is also guaranteed to obtain the outcome 0. Similarly, if Alice obtains the outcome 1, Bob is guaranteed to obtain the outcome 1. Each pair of outcomes (00) and (11) occur with probability ½, while the outcomes (01) and (10) do not occur. The matrix of probabilities would look like this:

| | Bob gets 0 | Bob gets 1 |
|---|---|---|
| Alice gets 0 | **½** | **0** |
| Alice gets 1 | **0** | **½** |

Figure 2.2.1.2: The matrix of probabilities for Alice's and Bob's joint outcomes. In this scenario, Alice's and Bob's measurement outcomes are fully correlated.

### 2.2.2 Entanglement Measures

When characterizing an entangled state, one can speak about the amount of entanglement present. In the earlier example, we referred to $|\Phi^+\rangle$ as the maximally entangled state, and there is well-established reason for doing so. $|\Phi^+\rangle$ is maximally entangled as it gives the maximal value for entanglement measures relevant to two-qubit states[2]. For bipartite qubits one such entanglement measure is the concurrence:

$$C(\rho) = \max\{0, e_1 - e_2 - e_3 - e_4\} \tag{36}$$

where $\{e_1,e_2,e_3,e_4\}$ are the eigenvalues of the matrix $M = (\sqrt{\rho}\ \rho'\sqrt{\rho})^{½}$ in decreasing order. The density operator $\rho' = (\sigma_y \otimes \sigma_y)\rho(\sigma_y \otimes \sigma_y)$. The concurrence quantifies entanglement as follows: $C = 0$ for a separable state, $0 < C < 1$ for an non-maximally entangled state and $C = 1$ for a maximally entangled state.

For density operators of higher dimension, one can use the negativity:

$$N[\rho] = \sum_i \frac{|\lambda_i| - \lambda_i}{2} \tag{37}$$

2 The fact that the partial trace of $|\Phi^+\rangle$ is the maximally mixed state is also part of its definition.

where $\lambda_i$ are the eigenvalues of the partial transpose of ρ, commonly denoted $\rho^{T_A}$. The negativity is simply the sum of the negative eigenvalues of the partial transpose. If N[ρ] = 0, ρ is said to have Positive Partial Transpose, and is called a PPT state. If N[ρ] > 0, ρ is said to have Negative Partial Transpose, and is called an NPT state. The negativity is a convex function of the state, meaning that:

$$N[\sum_i p_i \rho_i] \leq \sum_i p_i N[\rho_i] \tag{38}$$

The relation of the negativity to entanglement is known as the Peres-Horodecki Criterion[18], which states that:

- For a 2x2 or 2x3 dimensional composite Hilbert Space, ρ is separable iff ρ is PPT.
- For any other dimension, ρ is entangled if ρ is NPT.

The necessity of PPT for separability can be easily seen in the definition of a separable state. A general separable state ρ is defined as a state which can be written in the form:

$$\rho_{AB} = \sum_i p_i \sigma_{A,i} \otimes \phi_{B,i} \tag{39}$$

for some $p_i > 0$ and density operators $\varphi_i$ and $\sigma_i$. Then the partial transpose of ρ will be positive semi-definite, since $\rho^{T_A} = \Sigma_i(p_i \sigma_i^T \otimes \varphi_i)$ is still a mixture of positive semi-definite density operators.

NPT entangled states further have the property of *distillability*, which involves transforming multiple copies of non-maximally entangled states to produce an asymptotically pure entangled state [2]. For a bipartite d-dimensional quantum system, $N \leq \frac{1}{2}(d - 1)$, with equality for maximally entangled states. Therefore, in addition to being an entanglement witness, negativity is also frequently used as a dimension witness.

The Concurrence and Negativity are entanglement monotones. Entanglement monotones are non-increasing with local operations and classical communication (LOCC) such as local unitary transformations and local measurement. In particular, local unitary transformations do not change the values of entanglement monotones, and classes of states equivalent to each other under local unitary transformations

tend to have the same value for the same entanglement measure. One example of a class of states in this sense would be the maximally entangled two-qubit state. Below are four maximally entangled two-qubit states, known as the Bell basis states:

$$|\Phi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle) \tag{40}$$

$$|\Psi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \tag{41}$$

And it can easily be shown that they are equivalent up to local unitary transformations. In addition, all states of the form:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|i,j\rangle \pm |i_{\perp},j_{\perp}\rangle) \tag{42}$$

can be shown to be maximally entangled, and they all share the same maximal values of concurrence (C=1) and negativity (N = ½).

LOCC operations can be used to perform tasks in QIS experiments. For example, suppose Alice and Bob share one pair of maximally entangled qubits either in the state $|\Psi^{-}\rangle$ or $|\Phi^{+}\rangle$, which they wish to distinguish by performing LOCC operations. Alice and Bob will simply need a classical communication channel (such as a phone or the internet), and each of them performs a projective measurement on their respective qubit in the computational basis. They then share their measurement output with each other via the classical communication channel. If they obtained the same measurement result, they know their state is $|\Phi^{+}\rangle$, otherwise it is $|\Psi^{-}\rangle$. Other LOCC operations include local unitary transformations and numerical processing.

### 2.2.3 Entanglement for Certifiable Randomness

It is known that the outcomes of quantum measurements are truly random. However, not all randomness produced by quantum measurements are certifiable. Entanglement is a resource for generating certifiable true randomness. It would be pertinent to first introduce the notion of randomness. Consider the day-to-day scenario of rolling a six-sided dice labeled '1' to '6'. Assuming an ideal fair dice, the probability of obtaining any one of the six outcomes is 1/6. Notationally, we say

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$$

The probability distribution is uniform, so if one were to guess the outcome of a particular roll without any additional information, there would be no guess that is better than another, since all outcomes are equally likely. In a sense, one can talk about the uniform probability distribution being "maximally random". A quantity that can be used to quantify randomness is the min entropy, defined as

$$H = -\log p_{guess}$$

where $p_{guess}$ is the guessing probability, or the probability of the most probable outcome. One can easily see that for any experiment with n discrete outcomes, the deterministic probability distribution has the minimum entropy $H = 0$ and the uniform probability distribution has the maximum entropy $H = \log(n)$. For this reason, H can function as a measure of randomness – a deterministic probability distribution is minimally random and a uniform probability distribution is maximally random.

One probably would not consider a deterministic probability distribution a "random" process, but any non-deterministic probability distribution would be considered "random" at least to some degree (which one could quantify using entropy). This intuitive comparison can in fact be called *uniformity*, and one should see it as distinct from randomness.

What is then worth noting is when one considers the 6-sided dice roll a random process, one actually speaks with the assumption that the experimenter does not have any information that will allow them to predict the outcome. In principle, the information is actually there – if the dynamical variables were calculated, it would greatly increase the ability of the experimenter to determine the outcome of the dice roll. The main barrier to reducing the randomness of the dice roll this way is the *computational complexity* of calculating the dynamical variables of the dice, and not any fundamental principle of physics.[3]

What this illustrates therefore is that randomness has to be spoken of relative to the information possessed by the observer. While computational complexity and uniformity makes it less likely that the outcome can be predicted by naive guessing,

3 To put it more rigorously, the dice roll is *chaotic* because its evolution is sensitive to initial conditions, nevertheless, chaotic behavior is deterministic in principle.

it is not fool-proof against an observer that has knowledge of the variables that produce the outcomes of the process. A process where no information is available that can predict its outcome is thus said to be intrinsically random. By definition, any classical process that is used to generate its outputs must necessarily have information that can predict the outcome.

On the other hand, outcomes of measurements on quantum systems are truly random, and there is no information that allows for the prediction of the outcome of a particular measurement other than the statistical behavior of ensembles of the same system under measure. However, it is also known that measurements made on separable states can always be reproduced by classical strategies. (read section 2.3.5 for an illustration of this)

Entangled states are therefore a necessary ingredient for producing behavior that allow us to certify the intrinsic randomness of a process. These will be elaborated on in the subsequent section on Bell Non-locality with examples.

## 2.3 Bell Non-locality

### 2.3.1 First notion: A Bell Experiment

In this text, the discussion shall be limited to a specific class of experiments of the following form:

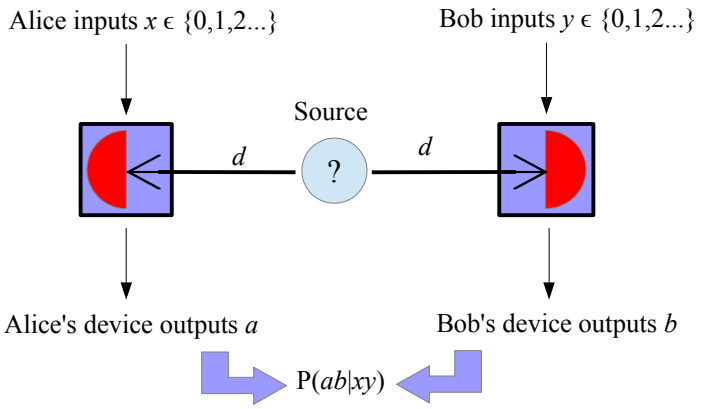



Figure 2.3.1.1: Schematic diagram of a Bell test for two parties. Communication channels between the devices are not specified in the above diagram, but it is assumed in general that the Alice and Bob may have precommunicated and the devices may be pre-loaded with some shared information.

Otherwise, Alice and Bob do not influence each other at the point of measurement.

Two parties (our beloved Alice and Bob), each own a device that reads in their inputs x and y respectively, and outputs a value a and b respectively. No *a priori* assumptions are made on the devices, and they may well be black boxes to Alice and Bob. These are a specific subset of experiments sometimes called Bell experiments. The scenario described at the end of Section 2.2.1 regarding local measurements on two maximally entangled qubits is simply an ideal quantum realization of a Bell Experiment for a particular pair of settings x and y.

### 2.3.2 The Correlation Matrix: The two setting, two output case

In most cases, the inputs x and y are picked from the integer values {0,1,2..., n – 1} for a n-input scenario. The values of a and b are chosen depending on the context, and in most studies involving two outputs either use a,b ∈ {0,1} or a,b ∈ {+1,–1}. Either way, once Alice and Bob finish their local measurements, they will communicate classically and collate their measurement outcomes to form their joint probability matrix or correlation matrix p(ab|xy):

| | P(b\|Y = 0) | | P(b\|Y = 1) | |
|---|---|---|---|---|
| | P(b = +1) | P(b = –1) | P(b = +1) | P(b = –1) |
| P(a = +1\|X=0) | **P(++\|00)** | **P(+–\|00)** | **P(++\|01)** | **P(+–\|01)** |
| P(a = –1\|X=0) | **P(–+\|00)** | **P(——\|00)** | **P(–+\|01)** | **P(——\|01)** |
| P(a = +1\|X=1) | **P(++\|10)** | **P(+–\|10)** | **P(++\|11)** | **P(+–\|11)** |
| P(a = –1\|X=1) | **P(–+\|10)** | **P(——\|10)** | **P(–+\|11)** | **P(——\|11)** |

Figure 2.3.2.1: The correlation matrix. The eight entries outside the main 4x4 block are the marginal probabilities, while the sixteen entries in the 4x4 block are the joint probabilities.

### 2.3.3 Local Variables, No-Signaling and Pre-established Agreement

In everyday life, correlations between distant parties is commonplace. People communicate with each other in order to synchronize their schedules, and establish

agreements on how to take subsequent action in their own individual posts in their workplace. These two examples illustrate two classical mechanisms that can explain distal correlations: signaling and pre-established agreement. From the idea of pre-established agreement one can formally characterize local variables – that the behavior of each device must be planned in advance for all possible pairs of inputs in such a way that the device can produce its output based on the input it receives but not the other device's. This would be equivalent to distributing a list of instructions, labeled using the variable λ. One can imagine such a list to look like this:

λ = 1: x = 0 → output a with probabilities P(a|x=0,λ=1)
x = 1 → output a with probabilities P(a|x=1,λ=1)
y = 0 → output b with probabilities P(b|y=0,λ=1)
y = 1 → output b with probabilities P(b|y=1,λ=1)

λ = 2: x = 0 → output a with probabilities P(a|x=0,λ=2)
x = 1 → output a with probabilities P(a|x=1,λ=2)
y = 0 → output b with probabilities P(b|y=0,λ=2)
y = 1 → output b with probabilities P(b|y=1,λ=2)
etc.

The distribution of λ should be independent of the inputs, i.e. p(λ|x,y) = p(λ), known as the *measurement independence* criterion.[4] The process that generates a and b can be stochastic as implied by the distributions P(a|i,j) above, but the process generating a should not depend on y and b on x. Then the joint probability distribution is obtained by averaging out over λ (which may be a continuous variable):

$$P(ab|xy) = \int d\lambda \; p(\lambda)p(a|x,\lambda)p(b|y,\lambda) \tag{43}$$

A joint probability distribution created this way, called a Local Variable (LV) model, can be said to be created through classical shared randomness, or pre-established agreement.

A final note on LV models: The statistics can be explained using deterministic local variables if and only if it can be explained with pre-established agreement. This

4 If y influences λ and thus p(a|x,λ), the marginal probabilities (see point 5) of a will depend on y.

notion is important for understanding that if a LV model can be constructed to produce the probability distribution for the outputs of a process, then there exists a local variable that can deterministically predict the output of the process.

Notice how we have assumed that p(a)[5] is independent of y and p(b) of x; this is the *no-signaling* condition. One can first consider no-signaling in the context of quantum theory. With reference to the example of Alice and Bob measuring two halves of a maximally entangled state (Figure 2.2.1.1), it should be noted that no information can be transmitted to Bob faster than light via Alice's measurement despite the apparent correlation between Alice's and Bob's measurement outcomes. Bob cannot detect from his measurement outcome nor his measurement statistics the result of Alice's measurement unless they have made prior agreement on the choice of measurement basis. Indeed, if Alice and Bob were not allowed to communicate prior to the measurement, even with the prior knowledge of the state being distributed to them, Bob's (as well as Alice's) measurement statistics will appear completely random (he will simply get 0 or 1 with probability ½). This makes sense, as it is the statistics obtained from measurements made on the maximally mixed state obtained in (18) after taking the partial trace of $|\Phi^+\rangle$ over Alice's subsystem. Furthermore, the full correlation of the measurement outcome is contingent on the choice of basis – both Alice and Bob must measure in the same basis in order to obtain those correlations.

On the other hand, allowing Alice and Bob to have prior agreement on the basis negates the use of the measurement for signaling, since their measurement outcomes would already be known to each other before measurement.

In a more general sense, no signaling is an assumption that assures that both Alice and Bob cannot communicate via their observed probabilities during the process of a measurement. If one complies with the assumption, then the only remaining mechanism that is consistent with LV would be pre-established agreement.

In terms of the correlation matrix, no-signaling imposes a constraint on the probabilities that appear in the table. Mathematically, the constraint is:

$$P(a|x) = \sum_b P(ab|xy) \ \forall y \text{ and } P(b|y) = \sum_a P(ab|xy) \ \forall x \quad (44)$$

which can be verbally interpreted as saying that the marginal probability for Alice

5 p(a) and p(b) are the marginal probabilities of a and b.

(read off one of the four entries outside the square) must equal the sum of the joint probabilities across all of Bob's possible outputs for any input by Bob. One can easily show that all LV and quantum correlations must satisfy no-signaling.

### 2.3.4 The EPR paper: What is the big deal about Local Variables?

Here I shall briefly outline the arguments Einstein-Podovsky-Rosen paper[5] to provide the historical motivation for discussing local variables in physics.

Evolving out of the field theories of classical physics, locality is the idea that for any object at a point to exert influence on another object at another point, something must exist in the space between them (such as a wave or a particle) that can mediate the interaction. The Theory of Relativity further limits the speed at which the wave or particle (collectively called fields) can propagate between two points to the speed of light. Developments in physical theories preceding quantum mechanics were developed to be consistent with the principle of locality, which was essential for the preservation of causality.

Realism is the idea the an objective description of reality exists independent of the observation of observer. The fact that even if the measurement result of a measurement that has not yet been performed does not exist, it is still a real entity and not a creation of the observer's mind. In the EPR paper, it defines an “element of reality” in the following fashion: if one can predict with certainty the outcome of a measurement of a quantity that has not yet been performed, that quantity corresponds to some element of reality.

Einstein's principle of local realism holds to two key ideas: (1) that cause-and-effect is limited by the speed of light, and (2) that a particle must possess a pre-existing value for any possible measurement that can be performed on it, even measurements that have not been performed. Local realism is a feature of Einstein's field equations, but Quantum Mechanics was deemed by him to be inconsistent with the principle, a representative example being measurements made on a distant pair of entangled states, which Einstein termed “spooky action at a distance”.

In that paper, Einstein held the view that Quantum Mechanics was an incomplete theory for several reasons. Contending that a complete theory must predict the outcome of any possible measurement on a system with certainty even if the

corresponding measurement has not yet been performed lead him to conclude that Born's rule was an unsatisfactory explanation. He was not in favor of the idea that an individual measurement outcome cannot be predicted but only statistics produced by ensembles are well-defined.

The existence of non-commutating observables in QM, whose values cannot be simultaneously measured, led to what Einstein believed to be a contradiction if the wavefunction is assumed to provide a complete description of the state. In the EPR paper, a counterexample of two physical observables of an entangled pair of particles was presented involving the momenta and the position of the two particles. However, the example was contrived such that the observables under study were actually commuting, which was used to (wrongly) conclude that there has to be a way to simultaneously measure the position and momenta of any particle. The argument is briefly outlined here:

Consider two particles, with their position $q_1$, $q_2$ and momenta $p_1$, $p_2$. Using the commutation relation $[q,p] = ih\mathbb{1}$, it can be checked that $q_1 - q_2$ are $p_1 + p_2$ are compatible observables. It is then possible to construct a simultaneous eigenfunction of both observables:

$$(q_1 - q_2)|\psi\rangle = x_0|\psi\rangle \ , \ (p_1 + p_2)|\psi\rangle = 0 \tag{45}$$

What this means is that if an experimenter were to measure the position of particle 1, the position of particle 2 will automatically be known. Similarly, the experimenter can then measure the momentum of particle 2, and determine the momentum of particle 1. Apparently, the position and momentum of both particles are now simultaneously known!

The EPR paper said that therefore, that there must exist some underlying mechanism (later known as LV or LHV/LR models) influencing the variables to give the observed effect of correlations produced by entangled states, non-commutating observables and Born's Rule. However, the EPR paper did not provide such a theory that can provide the description of these mechanisms.

The subtle aspect that EPR missed in presenting the above argument was that while the measurement outcomes of the above experiment can be reproduced by local variables, not all measurement outcomes on entangled states can be reproduced by

LV.[6] Therefore LV cannot be an explanation for QM as a whole, even if some measurements on certain entangled states admit LV explanations.

### 2.3.5 Measurements statistics of single qubits can be reproduced by LV

Here an example of how an LV model can reproduce the measurement statistics of a single qubit is provided. For any qubit state ½(𝟙 + **n·σ**), the probabilities of obtaining the measurement outcome a = +1 or –1 for the observable A = **A·σ** is given by P(a|A) = ½(1 + a **m·A**), and so 〈A〉 = **m·A** .

To reproduce these statistics, consider a vector **L** uniformly distributed[7] over the unit sphere $S^2$ in $\mathbf{R}^3$, and the system is represented by the vector **m**, and the measurement setting by **A**. Then the measurement outcome is computed via a = sign((**m** – **L**)·**A**) = +1 or –1. One can then show that

$$\langle \widehat{A} \rangle = \int_{S^2} d\theta d\phi \frac{1}{4\pi} \sin\theta \ \mathrm{sign}[(\vec{m} - \vec{L}) \cdot \vec{A}] p(\vec{L}) \tag{46}$$

Caveat: If measurement statistics on a single qubit can be reproduced by LV, then local measurements on any separable qubit state can always be reproduced by LV, since the LV strategy can always be implemented independently on each subsystem and mixed classically.

### 2.3.6 Bell Inequalities

A Bell inequality can be understood operationally as a condition on the observed statistics that rules out LV explanations. From here it will become useful to clarify the following notions. Looking back at the Bell experiment selected earlier in section 2.1.1, let $M_a$ and $M_b$ be the number of settings available for Alice and Bob respectively, and $m_a$ and $m_b$ be the number of outputs available for each of their devices.

In the case of $m_a = m_b = M_a = M_b = 2$ (sometimes known as the 2,2,2-scenario), the correlation matrix denoted by p(ab|xy), despite having 24 entries (see Figure 2.3.2.1), can be reduced to eight independent terms. For each setting x and y, under

6 It can be checked that the "counterexample" given by EPR can be reproduced by LV.

7 **L** does not need to be generated randomly – it can also be pre-registered with uniform relative frequency (i.e. deterministically)

no-signaling, one should easily see that the entries have the following relationship:

| | p(b=0\|y) | p(b=1\|y) = 1 – p(b=0\|y) |
|---|---|---|
| p(a=0\|x) | p(00\|xy) | p(01\|xy) = p(a=0\|x) – p(00\|xy) |
| p(a=1\|x) = 1 – P(a=0\|x) | p(10\|xy) = p(b=0\|y) – p(00\|xy) | p(11\|xy) = 1 – (the other three) |

Figure 2.3.6.1: Joint probability matrix for one pair of settings x,y, showing the interdependence of the entries.

The eight independent terms are therefore:

- p(a=0|x=0), p(a=0|x=1), p(b=0|y=0), p(b=0|y=1), also known as the marginals
- p(00|00), p(00|01), p(00|10) and p(00|11), the joint probabilities

In other contexts it may be more useful to consider the correlation coefficients $E_{xy}$= p(a=b|xy) – p(a≠b|xy), and marginals $\langle A_x \rangle$ = p(a=0|x) – p(a=1|x), $\langle B_y \rangle$ = p(b=0|y) – p(a=1|y). The correlation coefficients are sometimes referred to as the expectation values of the observable or moments $A_xB_y$, which they would be if the outcomes a,b are labeled ±1. The following notation is used interchangeably:

$$E_{xy} = \langle A_x B_y \rangle \tag{47}$$

Whichever choice of the eight numbers one uses, they define a vector in $\mathbf{R}^8$, known as the correlation vector **P**.

$$\mathbf{P} = (A_0, B_0, A_1, B_1, A_0B_0, A_0B_1, A_1B_0, A_1B_1)$$

A valid correlation vector will have all eight of its components have magnitude less than or equals one.

### 2.3.7 The Local Polytope and CHSH inequality

One can thus view probability distributions for the 2,2,2-scenario as living in an 8-dimensional space. The probability distributions that can be produced by LV thus

form a set on $\mathbf{R}^8$ space, and this set L is a convex set. It is known for any scenario (more settings or outputs) that this is also the case. This is simply due to the fact that any two LV probability distributions can be taken in convex combination to produce another LV probability distribution. A convex set can be characterized by knowledge of its extremal points, and any LV probability distribution can be decomposed as a convex sum of deterministic LV probability distributions. One can thus see that the deterministic points form the extremal points of the convex set $L$[8]. For 222, L has 16 extremal points, representing the 16 possible deterministic LV probability distributions. A convex set with a finite number of extremal points is referred to as a polytope.

An eight-dimensional polytope is bounded by seven-dimensional hypersurfaces succinctly known as facets, and each facet must have at least eight extremal points on it, all of which must be on the same side of the polytope. One can define a vector $\mathbf{n}$ in $\mathbf{R}^8$ such that for all points on a facet, they satisfy $\mathbf{n}\cdot\mathbf{P} = f$. Then all points in the polytope L must satisfy $\mathbf{n}\cdot\mathbf{P} \leq f$.

For higher values of ma, mb, Ma, Mb, L can be completely embedded in a fashion similar to above in $\mathbf{R}^D$ space, where $D = M_a M_b (m_a - 1)(m_b - 1) + M_a (m_a - 1) + M_b(m_b - 1)$.

It is known that, for any scenario, all facets of L are either Bell inequalities or positivity conditions (sometimes known as trivial facets). A classic example of a Bell Inequality would be the Clauser-Horne-Shimony-Holt (CHSH) Inequality, which can be represented by:

$$S(\mathbf{P}) = E_{00} + E_{01} + E_{10} - E_{11} \leq 2 \tag{48}$$

It is known that there are eight non-trivial facets of L, which correspond to the eight variations of the CHSH inequality and they are all equivalent in the sense that the eight facets are symmetric up to a relabeling of inputs x and y. These facets are non-trivial in the sense that a valid correlation vector that does not belong to L, such as P = (+1,+1,+1,–1), can violate this inequality giving S(P) = 4. It can be seen from here, that for the 2,2,2-scenario, the CHSH inequality is the only Bell inequality. It was derived first by Clauser, Horne, Shimony and Holt in 1969 [4]

8 This is implied by the fact that any LV model is simply a mixture of deterministic strategies – it's why the variable in LV models, such as the example in section 2.3.5 can simply be pre-registered.

and is the most studied Bell inequality of all to date.

### 2.3.8 Loophole-free Bell Violations

Bell inequalities are the constraints of LV, and as such the violation of Bell inequalities can be used as the certification that simulation of the statistics using LV is impossible. However, a Bell experiment can falsely violate a Bell inequality with a LV strategy by having one of the detectors refuse to provide an output. For example, Pearle (1970)[15] gives an explicit case of an LV model that reproduces statistics produced by a maximally entangled state of two spin-½ particles if one of the outputs of the LV model (corresponding to the 'no-detection' outcome) is discarded in the data processing. This leads to recognizing an important pitfall in tests aimed at certifying non-locality: post-selection is not an allowed data processing method. To avoid this pitfall, Alice and Bob can either: (1) include the no-detection outcome as an additional outcome, or (2) lump the no-detection outcome into one of the outcomes by default (either +1 or –1). Either way, Alice and Bob should consider the whole sample when computing the statistics in order to ensure the violation of the Bell inequality is conclusive.

In most experimental implementations of Bell experiments, the devices are characterized – meaning that their nature and efficiencies are known, and the cause of the devices' lowered efficiencies are ensured to not depend on the input choice. In the absence of characterization, it leaves room for possible Bell violations solely with the use of LV. The earlier example is a detection loophole, and there are various other loopholes that can be opened due to experimental limitations. Another example of a loophole for Bell violation is the locality loophole. The locality loophole is opened if there exists a way for the two devices to communicate classically, such as failing to place them sufficiently far apart during the measurement process, allowing for signaling to occur.

To date, only three experimental implementations of loophole-free Bell tests have been performed.[6][9][17] These implementations use various methods to close the loopholes, such as using detectors of high efficiency for the detection loophole, fast selection of measurement basis (combined with sufficient spatial separation) for the locality loophole, and high quality entanglement sources to reduce noise. It is worthwhile to point out here the authors themselves concede that closing loopholes for Bell violation is an experimentally demanding feat, and only these three

relatively recent (at the time of writing) experiments have been able to close the most significant loopholes simultaneously. These three experiments obtain a violation of the CHSH inequality to some degree of significance used to test the null hypothesis which is LV. The p-values come in the range of $10^{-7}$ to $10^{-31}$.

### 2.3.9 Other remarks

Violation of the CHSH Inequality by Quantum States

One then can consider an example of a quantum scheme that violates the CHSH inequality. Suppose we set up a Bell experiment where Alice and Bob share two halves of a maximally entangled qubit state. They choose the following measurement settings:

| | |
|---|---|
| Alice | |
| X = 0 | ½(1 + aZ) |
| X = 1 | ½(1 + aX) |
| Bob | |
| Y = 0 | ½(1 + b/√2(Z+X)) |
| Y = 1 | ½(1 + b/√2(–Z+X)) |

It can be shown that the probabilities obtained are

$$p(ab|xy) = \frac{1}{4}\left(1 + \frac{(-1)^{xy}ab}{\sqrt{2}}\right) \tag{49}$$

Putting this into the CHSH inequality gives $S = 2\sqrt{2}$, which is the maximal violation of the CHSH inequality for quantum states.

Do all statistics produced by entangled states violate Bell Inequalities?

Once one understands that although correlations that do violate Bell inequalities must necessarily be attained by entangled states, but some correlations produced by entangled states can still be reproduced by local realistic models, the EPR contradiction disappears. Consider the measurement scheme of the BBM92 protocol:

| | |
|---|---|
| Alice | |
| X = 0 | ½(1 + aZ) |
| X = 1 | ½(1 + aX) |
| Bob | |
| Y = 0 | ½(1 + aZ) |
| Y = 1 | ½(1 + aX) |

The statistics obtained are given by $p(ab|xy) = ¼(1 + \delta_{xy}ab)$, and the correlations are producible by both an entangled state as well as a local variable model (see the later section on Device Independence for a proof). It should then come as no surprise that the statistics produced by the BBM92 protocol do not violate the CHSH inequality:

$$S_{BBM92} = E_{00} - E_{01} + E_{10} + E_{11} = 1 - 0 + 0 + 1 = 2 \tag{50}$$

This example, along with the example presented in (LV and pre-established agreement), also illustrate a small but subtle point about the power of LV. While one may be in a hurry to prove "quantum superiority" over LV, local variables are resources in their own right, and they can be used for relevant purposes. Implementations (particularly experimental) of Bell experiments usually also require that Alice and Bob pick their inputs randomly with equal frequency, but it would be costly to require that they also use a quantum mechanism to generate their inputs – it makes little sense to have Alice measure a single qubit in order to decide which measurement setting to select for the measurement of another qubit. In this case, the use of LV would be sufficient for practical purposes.

One will find that, the BBM92 measurements on a different state, or the CHSH measurements on the maximally entangled state, indeed do rule out pre-established agreement – so the logical conclusion is that a local realistic model underlying all of QM cannot be found.

# Chapter 3

# Device-Independence and Semi-Device-Independence

In this chapter, I outline the notions in the study of Device-independence and Semi-device-Independence relevant to my project. I provide an example using an elementary QKD protocol that demonstrates the trade-off between full characterization and device-independence in terms of the security of the protocol.

## 3.1 Device-Independence

### 3.1.1 Introducing Device Independence

Notice that in our discussion on Bell Inequalities, there was no need for any consideration of how the statistics are obtained – the fact that measurements on quantum states incidentally produce correlations that violate the CHSH inequality has nothing to do with the fact that violation of Bell inequalities rule out LV explanations. One can say that any violation of a Bell inequality implies the presence of non-locality, regardless of the choice of explanation for the non-locality (of which QM is the most famous one). Thus Bell inequalities provide a way to certify non-locality in a device-independent way.

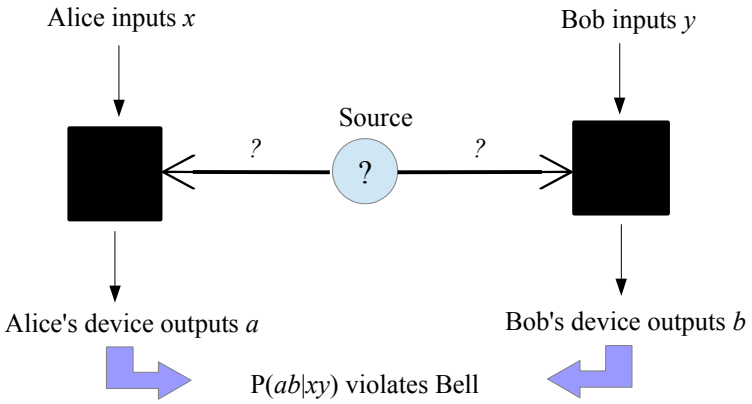



Figure 3.1.1.1 A Bell test where the devices are black boxes and the nature of the source is unknown is known is known as a device-independent (DI) Bell test. If the correlations between the outputs violate Bell inequalities, non-locality can be certified in a DI way.

In the specific case where a Bell violation is observed in the context of quantum theory, it certifies the presence of entanglement solely by the knowledge of the

observed correlations.

In the field of quantum communication, entangled states are a necessary ingredient for generating certifiable randomness and secrecy. However, most implementations, both theoretical and experimental, require at least some assumptions regarding the nature of their devices. One of the basic assumptions include the physical degree of freedom being measured and the fact that the device does not write in nor read from any other degrees of freedom. While in most laboratory implementations the experimenters are reasonably well-protected against adversarial mechanisms, these assumptions made on the experimental setups can open loopholes which may compromise the tasks the setups seek to perform in "real-world" applications of these setups. i.e. Applications of quantum communication typically require characterization of the devices in order to work – be it randomness generation or secrecy extraction – failure to comply with the assumptions made on the devices will compromise the security of the protocols.

In the case where one has minimal characterization, one has Device Independence (DI), which describes a suite of tests that can certify entanglement with the sole assumption of no-signaling. The disadvantage of DI is that implementations of DI only work with correlations that violate Bell inequalities, which must be loophole-free – making them experimentally demanding. Furthermore, DI entanglement witnesses usually give pessimistic bounds. On the other hand, knowledge of the measurement settings and the Hilbert Space dimension will allow for a reconstruction of the state (up to some precision), which requires more assumptions and thus opens more potential loopholes. This trade-off can be illustrated by an example.

### 3.1.2 An Example Using Secrecy Extraction

The correlations obtained from measurements on entangled states, along with prior agreement on the choice of basis, can be used as a resource for secrecy extraction. However, the security of the protocols depend on the level of characterization of the setup. For illustration purposes, consider a simple example of a Quantum Key Distribution (QKD) protocol, with ideal states and detectors. In the Bennett-Brassard-Mermin (1992) protocol, two parties Alice and Bob share two halves of a maximally entangled bipartite qubit state[3]. They first communicate classically, and establish two choices of measurement basis, say Z and X.

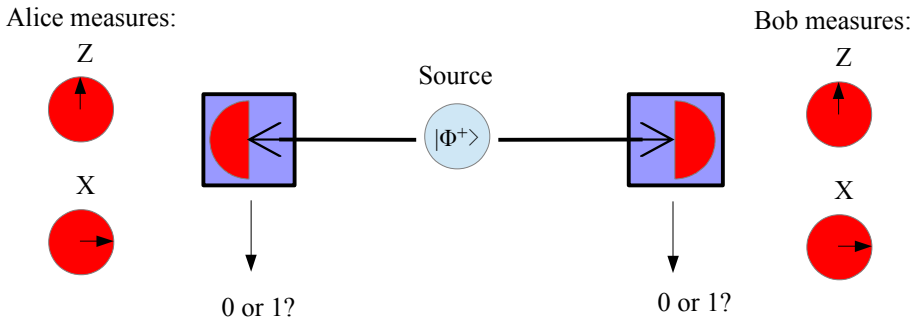



Figure 3.1.2.1: The scheme of the BBM92 protocol. Alice and Bob share two halves of a maximally entangled qubit state. Their detectors are set to measure either in the Z or X basis, and they collate their measurement outcomes (either 0 or 1) after finishing their local measurements on a sequence of qubit pairs.

Alice and Bob then share two halves of a sequence of maximally entangled qubits, and make a sequence of measurements on each qubit they receive. Alice and Bob will flip a coin for each qubit that they measure and select a measurement basis based on the coin flip result (thus randomizing the basis choice). Alice and Bob will each then have a list of measurement outcomes (a bit string) that would resemble:

| Qubit sequence no. | Measurement result | Basis Choice |
|---|---|---|
| 1 | 0 | Z |
| 2 | 1 | Z |
| 3 | 1 | Z |
| 4 | 1 | X |
| 5 | 1 | X |
| 6 | 0 | Z |
| 7 | 0 | X |
| 8 | 1 | X |
| 9 | 0 | Z |
| 10 | 1 | X |

Figure 3.1.2.2 A table simulating what the list of outcomes Alice or Bob would look like after performing their measurements.

At first glance, both Alice and Bob will on average have ½ of their measurement results being 0 and half being 1 for each setting, as well as in total. If Alice's measurement outcome is a and Bob's is b, then P(a = 0|Z or X) = ½ = P(a = 1|Z or X) and P(b = 0|Z or X) = ½ = P(b = 1|Z or X). This can be calculated from the fact that the partial trace of the maximally entangled bipartite qubit state over either Alice or Bob will obtain the maximally mixed state for the other.

The aforementioned probabilities are the marginal probabilities, and indeed neither Alice nor Bob will be able to learn anything about what the other party did upon performing their local measurement (such as which basis they chose), until they reveal their choice of basis via some other subluminal means. If they were to mutually reveal their basis choice and collate their joint probabilities, they will be able to construct their joint probability matrix, p(ab|xy), where x and y are Alice's and Bob's basis choice respectively.

The joint probability matrix will look like:

| | | Bob measures Z | | Bob measures X | |
|---|---|---|---|---|---|
| | | Obtains 0 | Obtains 1 | Obtains 0 | Obtains 1 |
| Alice measures Z | Obtains 0 | **½** | **0** | **¼** | **¼** |
| | Obtains 1 | **0** | **½** | **¼** | **¼** |
| Alice measures X | Obtains 0 | **¼** | **¼** | **½** | **0** |
| | Obtains 1 | **¼** | **¼** | **0** | **½** |

Figure 3.1.2.3 The correlation matrix displaying the probability of a joint outcome given the measurement settings of Alice and Bob. If Alice and Bob select the same measurement setting, their results are fully correlated, otherwise they are maximally random (uniform probability for all possible joint outcomes).

In the BBM92 protocol, Alice will reveal her basis choice (with the sequence numbers) to Bob via the classical channel. Bob will then answer on the classical channel which measurements they had performed with the same basis choice, and both Alice and Bob will discard those bits in their bit string which are obtained via a different measurement basis from the other. On average, this will retain half of

their original bit string. In the ideal case, where the state and detectors are perfect, their bit strings will be fully correlated and secret.[9]

Introducing the adversary Eve

At this point, it is useful to consider if the protocol is secure. The field of QKD deals with wide ranges of attacks by adversaries, which will not be discussed in detail. This example shall consider an adversary that attempts to reproduce the behavior of the protocol using LV resources. If Alice and Bob ensure that they indeed have a source of qubit pairs[10] in the state $|\Phi^+\rangle$, and that their detectors are indeed measuring in the Z and X bases, the protocol is secure against such a reproduction.

Now we consider relaxing these assumptions. What if an adversary Eve wanted to devise a counterfeit set-up that can reproduce the same behavior as the maximally entangled state under Alice and Bob's measurement scheme? It can be shown that there is indeed such a strategy possible, and this allows Eve to have full knowledge of the bit string that Alice and Bob end up with.

So now suppose Alice and Bob's measurement devices were pre-programmed by Eve in the following way, and then later distributed to them as QKD devices:

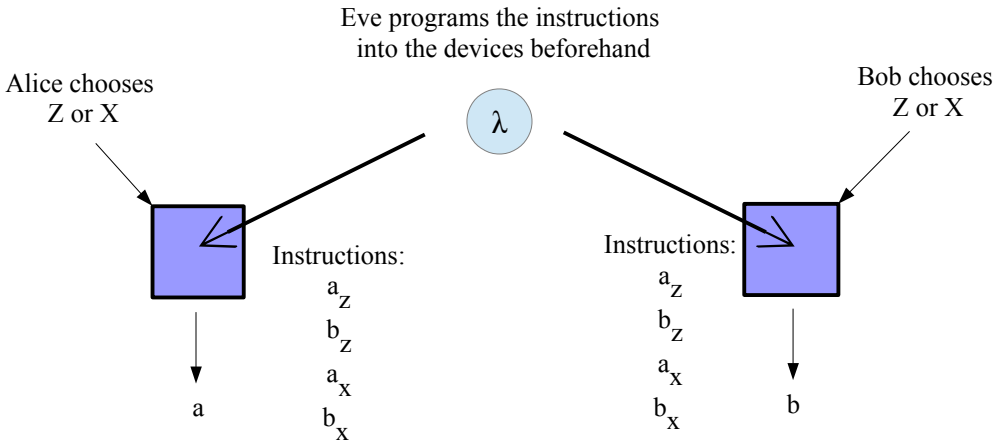



Figure 3.1.2.4: If Eve were the manufacturer of the QKD devices used by Alice and Bob, the above diagram illustrates what she could possibly do: she could pre-program the devices with instructions on how to respond locally to the inputs selected by their users.

9 There are other problems not addressed here: before the point of measurement, the state is quantum, but once the outputs are generated, everything is classical – so the secrecy is only guaranteed right after the point of measurement.

10 One example would be a non-linear crystal providing down-converted pairs of photons.

| p(λ) | λ | $a_z$ | $b_z$ | $a_x$ | $b_x$ |
|---|---|---|---|---|---|
| ¼ | 1 | + | + | + | + |
| ¼ | 2 | + | + | – | – |
| ¼ | 3 | – | – | + | + |
| ¼ | 4 | – | – | – | – |

Figure 3.1.2.5: The list of instructions that Eve could preload into Alice and Bob's devices to generate the statistics of the BBM92 protocol.

We can see that this strategy will be able to reproduce the exact correlations produced by an "honest" measurement of a maximally entangled state under the BBM92 protocol. The expectation values can be explicitly worked out to be:

$$\langle ZZ \rangle = \langle XX \rangle = \tfrac{1}{2} + \tfrac{1}{2} - 0 - 0 = 1 \tag{52}$$

$$\langle ZX \rangle = \langle XZ \rangle = \tfrac{1}{4} + \tfrac{1}{4} - \tfrac{1}{4} - \tfrac{1}{4} = 0 \tag{53}$$

$$\langle Z \rangle = \langle X \rangle = \tfrac{1}{2} - \tfrac{1}{2} = 0 \tag{54}$$

So under this strategy, as far as Alice and Bob are concerned, they will still obtain a random string of bits that they mutually share and are apparently secret. However, since the devices are pre-programmed, the bit string will be deterministic with respect to Eve. This is a consequence of the fact that randomness is relative and depends on the information possessed by the observer. In modern applications, generators of classical randomness such as random number generators rely on uniformity and computational complexity to prevent prediction of the generators' outputs. However, these generators are not truly random by design. For example, classical random number generators use a *seed state* that determines its outputs. Although the output is random with respect to the end user, it is deterministic with respect to the designer (in the above example, Eve is the designer who decides how λ is chosen and output).

Back to LHV, one last time

In this example, the "hidden variable" is λ, which determines which of the four pre-established sets of outputs are selected in each run. The model is local as the output

of each device is only influenced by variables in its subluminal vicinity (the pre-programming and distribution by Eve are all bound by relativistic limits), and hidden because the model generates the quantum behavior of the system without being part of the quantum theory. However, the fact that there exist some quantum behavior that can be reproduced by a classical (LV/LHV) model, does not mean that an LV explanation for all of QM exists. Finally, it should be noted that the use of the term LHV is more a matter convention – in the absence of context (in the above example, the context is QKD) there is no necessity for the pre-established agreement to be hidden.[11]

Considerations in experimental implementations

The aforementioned possibility of reproducing the correlations in the BBM92 protocol therefore requires that the characterization of the devices be done in order to ensure its security. In practice, very often one does not have full control over the characterization of the device, and the imperfections in real detectors open loopholes for adversarial attacks, breaking the security of these QKD protocols.

Certifying that the source is indeed entangled, is therefore one way to assure the protocol is secure. In the case of full characterization, one has already made the necessary assumptions which would lead one to reconstruct the state fully and conclude that the source is indeed quantum. In the DI case, the statistics must violate Bell inequalities in order to certify entanglement, which the BBM92 statistics do not. Experimental implementations that attempt to certify non-locality and entanglement in a DI fashion need to show that the experimental correlations violate some Bell inequality to a sufficient degree of significance (i.e. after accounting for fluctuations in the data), and also ensure the Bell violation is loophole-free.

### 3.1.3 The Quantum Set and No-Signaling Polytope

It was mentioned earlier in this text that no-signaling imposes a constraint on the correlation matrices one can write down. In fact, it is known that for any scenario $\{m_a,m_b,M_a,M_b\}$, there exists a superset of the local polytope L known as the no-

11 In this case, the variable being hidden does provide an advantage, since Eve would certainly not tell anyone that the devices are actually pre-programmed.

signaling polytope, commonly labeled NS. This polytope has extremal points that are either the local deterministic points or the non-local Popescu-Rohrlich Box (PR-box)[16], which are the correlations that attain the algebraic limit of the Bell correlations. In the case of the 2,2,2-scenario, the algebraic limit of the CHSH correlation is S = 4, which can be attained by the following zero-marginal correlation matrix:

| | ½ | ½ | ½ | ½ |
|---|---|---|---|---|
| ½ | **0** | **½** | **½** | **0** |
| ½ | **½** | **0** | **0** | **½** |
| ½ | **0** | **½** | **0** | **½** |
| ½ | **½** | **0** | **½** | **0** |

Figure 3.1.3.1: The correlation matrix for a PR-box.

This correlation matrix can easily be verified to satisfy no-signaling. One can see that the CHSH correlation evaluates to $S = -E_{00} - E_{10} - E_{11} + E_{01} = 4$ (the overall negative sign justified by swapping the outputs a and b for one device). The correlation vector is given by PR = (0,0,0,0,–1,1,–1,–1), so NS can also be embedded in eight-dimensional real space $\mathbf{R}^8$.

Quantum correlations are a subset of no-signaling correlations, characterized by $p(ab|xy) = Tr(\rho\ \Pi^x_a \otimes \Pi^y_b)$, where the measurement operators Π satisfy the conditions stipulated in section 2.1.1, and ρ is a density operator. The set of correlations attainable by quantum states and measurements is known as the Quantum Set Q, and it is known that Q is a convex set. Unlike L and NS, Q is an elliptope, with a continuously infinite number of extremal points. The difficulty in characterizing Q lies in the fact that Q bounds correlations attainable by states of any dimension, including infinite-dimensional systems.

The following is a 2-D slice of the no-signaling polytope studied during this project, for the 2,2,2-scenario.

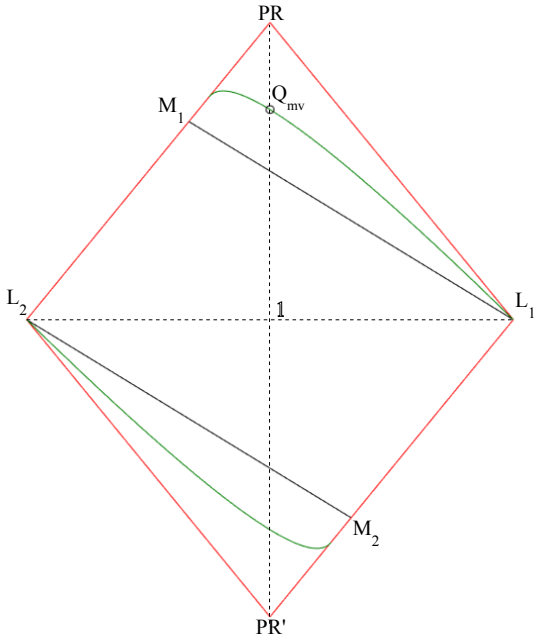



Figure 3.1.3.2: A 2-D slice of the eight-dimensional no-signaling polytope. The extremal point of NS (red) are the local deterministic points and PR-boxes, while the quantum set (green) is an elliptope. The statistics on the green curve are attained by bipartite measurements on the singlet given by in Figure 4.2.1.2.

In the 2,2,2-scenario, the boundary of Q is known to be attainable by measurements on pure qubit states, and the maximal violation (also known as the Tsirelson Bound) of the CHSH inequality on Q is $2\sqrt{2}$. The following subsections briefly outline two methods for computing Tsirelson bounds, the second of which is relevant for certifying entanglement in a DI fashion.

The Bell Operator

A Bell inequality can be visualized in the no-signaling polytope as a facet, and any facet can be described using the equation **n**·**P** = f, where **n** is a normal vector and **P** is the correlation vector. The representative CHSH inequality simply has **n** = (0,0,0,0,1,1,1,–1). One can then consider an operator known as the Bell Operator B, where

$$\begin{aligned}\widehat{B} = n_1\widehat{A}_0 + n_2\widehat{B}_0 + n_3\widehat{A}_1 + n_4\widehat{B}_1 + n_5\widehat{A}_0\widehat{B}_0 \\ + n_6\widehat{A}_0\widehat{B}_1 + n_7\widehat{A}_1\widehat{B}_0 + n_8\widehat{A}_1\widehat{B}_1\end{aligned} \tag{55}$$

and the expectation value of B is

$$\langle \widehat{B} \rangle = \mathrm{tr}(\rho \widehat{B}) \tag{56}$$

The expectation value of B will simply be the Bell correlation obtained by the state ρ under the measurements operators, which can be computed if the dimension of the state is assumed. In the case of qubits, $A_x$ and $B_y$ will simply be linear combinations of the Pauli matrices, and can be parametrized using the Bloch Sphere. The maximal expectation value of B can be obtained via an analytical or numerical optimization. The Tsirelson bounds obtained using the Bell Operator always provide a lower bound on the maximal violation.

### 3.1.4 The NPA Hierarchy and Local-Level Moment Matrix: Certifying Entanglement

The Navascues, Pironio and Acin (NPA) hierarchy[13], is a set of necessary conditions that bound the correlations attainable by quantum states. The hierarchy is complete, in the sense that any correlations that are not found in Q will fail to satisfy the conditions at some level of the hierarchy, typically labeled $Q_1$ to $Q_n$.

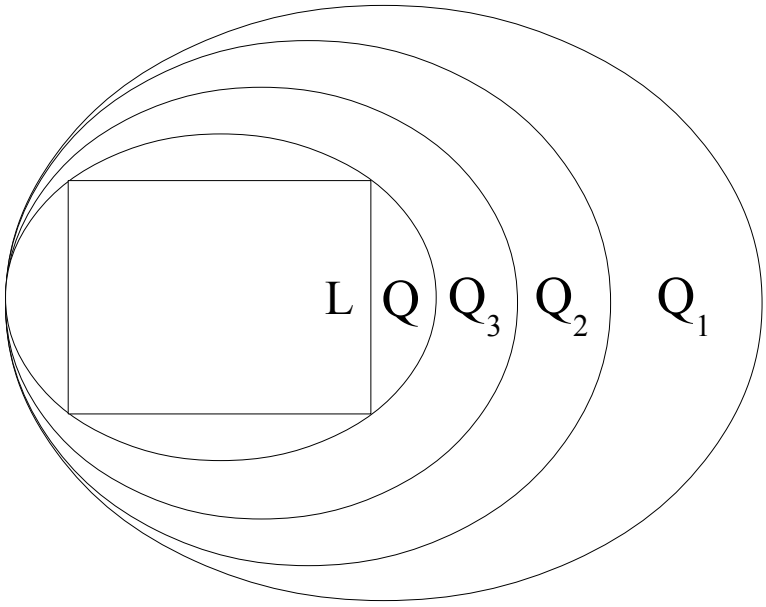



Figure 3.1.4.1: A geometric visualization of the hierarchy, $Q_1 \supseteq Q_2 \supseteq Q_3 \ldots \supseteq Q$. Any correlations not in Q will fail the condition at some level of the hierarchy.

The conditions are formulated as a semi-definite programming (SDP) problem, which maps a state to a moment matrix, usually denoted χ. χ is constructed in the

following way:

$$\chi = \sum_{i,j} |i\rangle\langle j| \mathrm{tr}(\rho\ O_i^\dagger O_j) \tag{57}$$

$$O_i = A_i \otimes B_i \tag{58}$$

$$A_i = \prod_k^{|A_i|} A_{k,i}\ ,\ A_{k,i} \in \{\mathbb{I}, A_0, A_1, ...\} \tag{59}$$

$$n = \text{order of } \chi = \max\{|O_i|, |O_j|\} = \max\{|A_i| + |B_i|\} \tag{60}$$

where $|A_i|$ and $|B_i|$ are the number of non-identity operators concatenated together. Some moments in $\chi$ can be found in the observed statistics p(ab|xy), while the others are unknown and left to run free, subject to the positivity of $\chi$.[12] This defines a convex constraint, which the optimization of a desired objective function would be subjected to. If a moment matrix $\chi$ includes moments that are up to 2n-fold products of the measurement operators, one may say that $\chi$ describes $Q_n$. The more moments are included in $\chi$, the stricter the constraint, and the closer one gets to describing Q.

When evaluating Tsirelson bounds, the objective function is chosen to be the Bell expression **N·P**, and all moments are left free.

$$\min\ N \cdot P\ ,\ \chi \geq 0 \tag{61}$$

The bounds computed using the NPA hierarchy are upper bounds to the maximal violation of the Bell inequality on $Q_n$, and in the case of 2,2,2, they are tight when the marginals are zero.

When it comes to DI certification of entanglement, a useful and closely related hierarchy is the local-level moment matrices developed by Moroder et al. (2013) [12]. These matrices include moments that are up to order 2n for each partition in the measurement, and may be referred to as $Q_{Ln}$. In general, $Q_n \supseteq Q_{Ln} \supseteq Q_{2n}$ They are defined in a similar fashion as (57):

(62)

12 The mapping of ρ to χ is a positive map.

$$\chi = \sum_{ijkl} |ij\rangle\langle kl| \ \mathrm{tr}(\rho A_i^\dagger A_k \otimes B_j^\dagger B_l)$$

$$A_i = \prod_k^{|A_i|} A_{k,i} \ , A_{k,i} \in \{\mathbb{I}, A_0, A_1, ...\} \tag{63}$$

$$n = \text{order of } \chi = \max|A_i| = \max|B_i| \tag{64}$$

χ can then be referred to as a moment matrix of local level n (or order n), containing all the 2n-fold products of the local measurement operators. These matrices have a bipartite structure, which can be used to compute their partial transpose and negativity. Since negativity is an entanglement monotone, the negativity of χ serves as a valid device-independent lower bound on the negativity of the state.

### 3.1.5 A final note on Device-independence: Self-testing

A particular family of quantum correlations have an additional property known as self-testing [19]. If a set of correlations self-test, it is possible to identify the underlying state solely from the observed correlations, up to a local isometry. Self-testing statistics also have another neat feature: if a set of statistics are self-testing, they must necessarily lie on a boundary point or extremal point of the quantum set. This feature allows for another way of characterizing Q – if the statistics fulfill the conditions for self-testing [19], one can immediately say that the statistics lie on the boundary of Q without any additional knowledge. One example of statistics that self-test are the quantum correlations that produce the maximal violation of CHSH.

## 3.2 Semi-Device-Independence

### 3.2.1 Semi-Device-Independence

Semi-Device-Independence describes a range of tests that bring back some of the assumptions on the devices in Bell experiments. These assumptions include the Hilbert space dimension and the measurements being performed. In this text, the former assumption is dealt with: the Hilbert Space dimension of the system. First,

one can consider some recent approaches that use the assumption that the measured state is a two-qubit state. Recent work by Moroder et al. (2012) provided analytical bounds for the correlations that certify the presence of entanglement[11], while Goh et al. (2016) [7] proposed a scheme that not only certifies the presence of entanglement, but puts a lower bound on its amount. Goh's scheme will be briefly described in the subsequent section.

Bringing back the assumption of the dimension is useful for when the degree of freedom being measured is known, and has been implemented in experiments that attempt to certify the presence of entanglement[1]. Indeed, in most practical situations, it would seem quite ridiculous to perform any experiment where the experimenter does not even know whether the laser they are using is even a laser. In addition, Goh et al (2016) showed that the assumption of the dimension allows for quantification of entanglement for statistics that do not violate any Bell inequality. In the 2,2,2-scenario, the CHSH inequality is the only inequality that needs to be considered – if the statistics do not violate CHSH, they will not violate any Bell inequality.

In addition, in comparison to full characterization, assumption on the dimension is useful for cases where the non-ideality in the detectors are not modeled, for reasons including the lack of knowledge or the difficulty thereof.

### 3.2.2 Semi-DI Scheme by Goh (2016)

Goh (2016) computes the minimum amount of certifiable entanglement for a set of observed statistics, supplemented with the assumption of the dimension of the measured states, which was restricted to be qubits. The entanglement is quantified using the concurrence, an entanglement monotone. The scheme is set up as an optimization problem, with the following framework:

$$\min C(\rho)\ s.t.\ \mathrm{tr}(\rho\ \Pi_a^x \otimes \Pi_b^y) = p(ab|xy) \tag{65}$$

$$\rho \geq 0\ ,\ \sum_a \Pi_a^x = \sum_a \Pi_b^y = \mathbb{I}\ ,\ \Pi_a^x, \Pi_b^x \geq 0 \qquad \rho \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2)$$

In Goh's implementation of this optimization scheme, the statistics serve as a constraint, while the state and measurements are left free. The result provides a lower bound on the amount of entanglement present in the state that is compatible

with the observed statistics, produced by the associated measurements.

Goh (2016) dealt with ideal statistics, as well as one case of simulated noisy statistics where the detectors were assumed to have an uncorrelated 'no-detection' outcome. The study found that in certain cases, the violation of a Bell inequality is no longer necessary if the assumption of qubits was supplemented, and statistics that do not violate the CHSH inequality can certify the same amount of entanglement as statistics that do. Of particular note are the BBM92 statistics, which uniquely identify the state to be the maximally entangled state $|\Phi^+\rangle$ with just the knowledge of the dimension (the measurements need not even be characterized!). This is an example of Semi-DI self-testing, in comparison to the CHSH statistics which allow for DI self-testing.

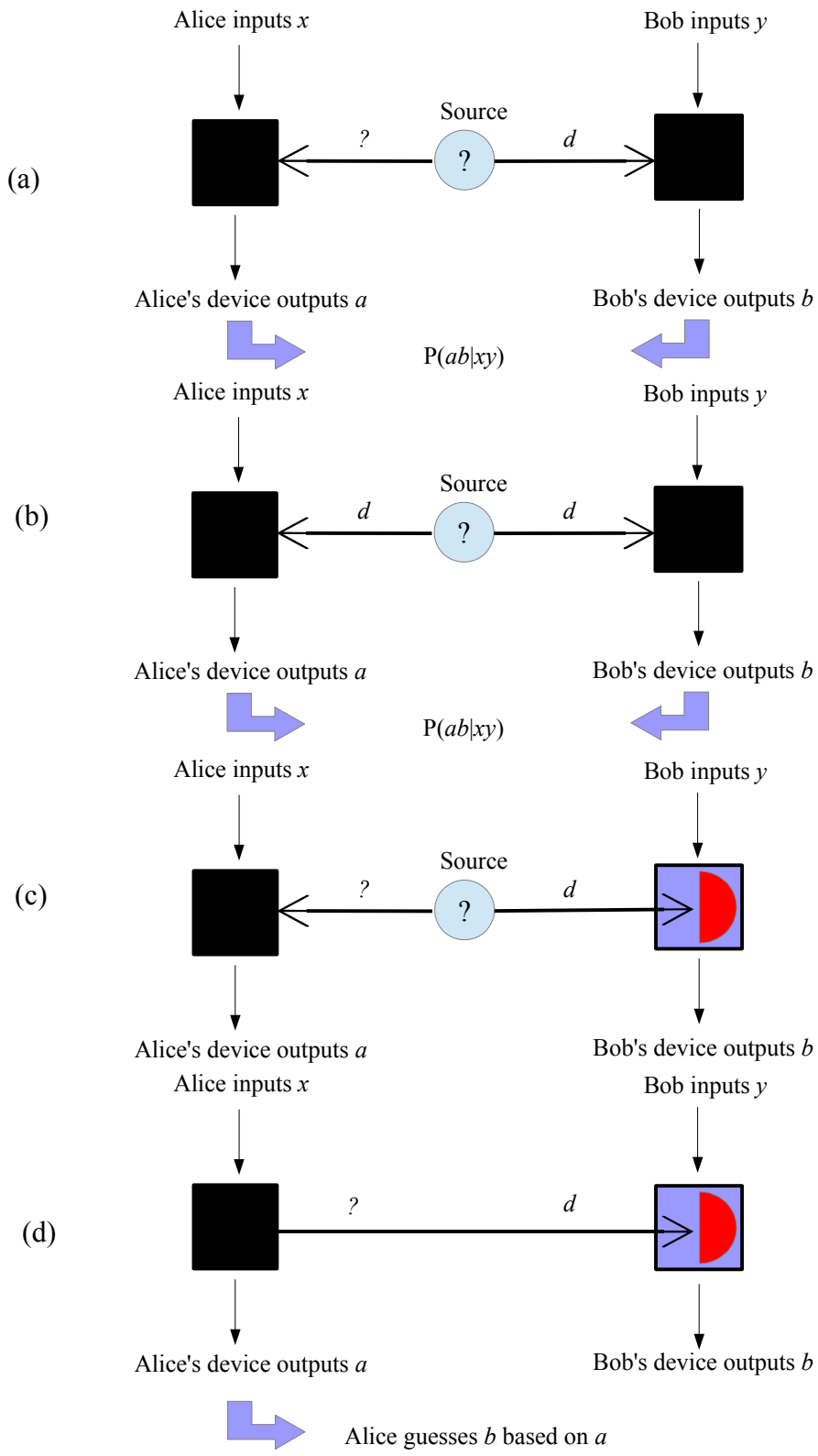



Figure 3.2.2.1: A diagram featuring various levels of characterization of a Bell test. (a) when the dimension on only one party is assumed, as in some semi-DI BB84 protocols (b) when the dimension on both parties' is known (c) when one side has full characterization and the other has none (d) Steering, where one untrusted party guesses the outcome of the measurement for the other

# Chapter 4

# Main Result

## 4.1 General Framework

In this project, I work with a modification on the scheme by Goh et al (2016). The modified scheme seeks to certify lower bounds on the amount of entanglement needed in an experimental source of qubit pairs to produce the observed correlations within specified error bounds, supplemented with the assumption that the state is bipartite qubits.

The minimum amount of entanglement certifiable by observing the relative frequencies $\hat{p}(ab|xy)$ for a known Hilbert space dimension is obtained via an optimization:

$$\begin{aligned} &\min\ E[\rho_{AB}] \\ &s.t.\ \left|p(ab|xy) - \mathrm{tr}(\rho_{AB}\ \Pi_x^a \otimes \Pi_y^b)\right| \leq \epsilon_{a,b,x,y}\ \forall\ x,y,a,b \\ &\rho_{AB} \geq 0, \rho_{AB} \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B) \\ &\Pi_x^a, \Pi_y^b \geq 0\ ,\ \sum_a \Pi_x^a = \sum_b \Pi_y^b = \mathbb{I} \end{aligned}$$

where E is an entanglement monotone of the state $\rho_{AB}$, which is the negativity [18] for any arbitrary dimension, and the concurrence [20] for $\dim(H_A) = \dim(H_B) = 2$. The measurements are allowed to be POVMs, and the optimization runs over all states and measurements compatible with the observed statistics within the error margin specified by the error matrix $\varepsilon_{abxy}$.

Analytical solutions for such an optimization are difficult, due to the large number of parameters. Furthermore, it is known that the subset of Q for finite dimensional systems is not convex[14], so the optimization is not a semi-definite program. Fortunately, in the case where the state is assumed to be qubits, the number of parameters is sufficiently small that heuristic optimization algorithms are reliable.

I also perform a semi-definite program to compare the amount of entanglement

certifiable in the DI case, and use two noise models, Bell-diagonal and isotropic, to perform a comparison in the case of full characterization. The figure of merit of choice is the negativity. In the following sections, I first outline the experimental scheme and data, followed by the three certification tests, and end off with section 4.6 that compares the three characterization levels and concludes the study.

## 4.2 Experimental Scheme and Data

### 4.2.1 Experimental Setup

In this section, I detail the experimental set-up as reported by the experimenters.[13]

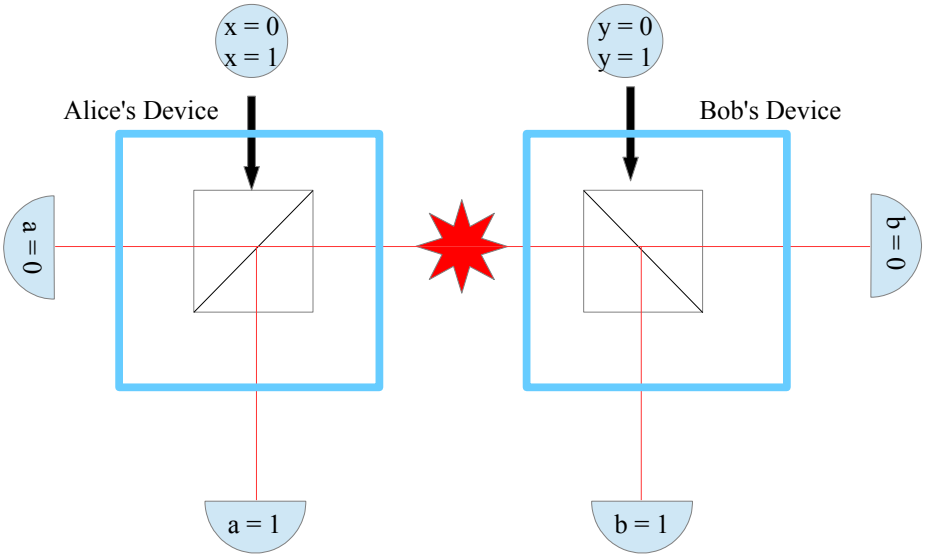



Figure 4.2.1.1: Schematic of the experimental setup. The source sends one half of a pair of photons generated through Type II SPDC, which go through a PBS into either the 0 or 1 channel.

The degree of freedom measured in the experimental set-up is the polarization of photons. Photon pairs were generated used a Type II down-conversion process, and each pair of signal and idler were sent through a polarizing beam splitter. Data was taken when both the a and b channels clicked exactly once. The axis of polarization of the PBS was chosen depending on the inputs x and y in the following way: Let $A_x$, $B_y$ be the observables corresponding to the polarization direction chosen for the PBS on settings x,y. Then $A_x$, $B_y$ represented on the Bloch sphere, will look like below.

13 Kudos to Alessandro and Hoh Shun for the data!

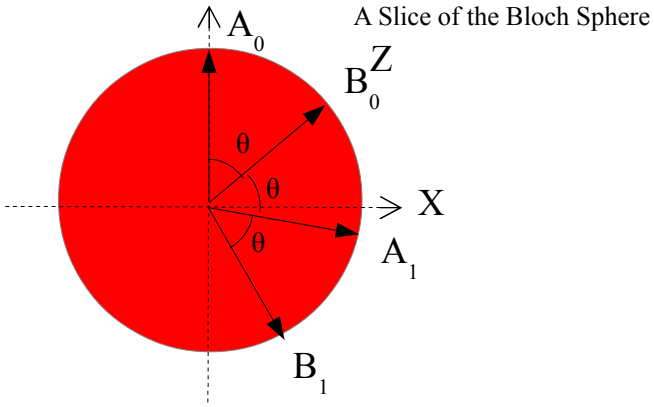



Figure 4.2.1.2: A Slice of the Bloch Sphere, representing the measurement settings.

The range of θ was taken over θ = 0 to θ = 0.44 rad, with a sample size of $n \sim 330000$ for each measurement setting x,y and a given angle θ. No error bars were provided for the Bloch angles.

### 4.2.2 The Correlation and Error Matrices

The data was given in the form of a list of quats (0, 1 2 and 3), corresponding to the four possible measurement outcomes (00, 01, 10 and 11). I wrote a MATLAB script named "Pabxy22" (script in Appendix D) that reads the list of quats and record their frequencies, and computes their relative frequencies $\hat{p}$ from the total number of quats for each setting (x,y).

$$\widehat{p}(ab|xy) = \frac{N_{ab,xy}}{N_{xy}}$$

The correlation matrix estimate is constructed by repeating the above computation for each pair of measurement settings. Under the assumption of Independent-Identical-Distribution (IID), the frequencies of photon detection for each channel pair (a,b) and the total frequency of detection on all channel pairs are treated as independent random variables N, so $\sigma_N = \sqrt{N}$ is an estimator for the uncertainty for each of them. The errors in the relative frequencies were computed using the variance formula:

$$\epsilon_{abxy} = \sqrt{\left(\frac{\partial \widehat{p}}{\partial N_{abxy}}\sigma_{N_{abxy}}\right)^2 + \left(\frac{\partial \widehat{p}}{\partial N_{xy}}\sigma_{N_{xy}}\right)^2}$$

$$= \sqrt{\left(\frac{\sqrt{N_{abxy}}}{N_{xy}}\right)^2 + \left(\frac{N_{abxy}}{N_{xy}^2}\sqrt{N_{xy}}\right)^2}$$

$$= \sqrt{\frac{N_{abxy}}{N_{xy}^2} + \frac{N_{abxy}^2}{N_{xy}^3}}$$

where $N_{abxy}$ is the frequency of detection at each pair of output channels and $N_{xy}$ is the total frequency of detection for a setting x,y. This defines the error matrix $\varepsilon_{abxy}$. with one error bar for each of the 16 probability estimates.

For the DI certification scheme, it is more useful to compute the correlation vectors and propagate their associated errors. The statistics are experimental, which means that the estimated correlation matrix will not satisfy no-signaling. For the marginals, this means that in general, $A_x(y = 0) \neq A_x(y = 1)$. A unique value is obtained by taking the average over y:

$$A_x = \frac{1}{2}\sum_y \left(\sum_b \widehat{p}(a = 1, b|xy) - \widehat{p}(a = -1, b|xy)\right)$$

$$\epsilon_{A_x} = \frac{1}{2}\sqrt{\sum_y \sum_b \left(\epsilon_{1,bxy}^2 + \epsilon_{-1,bxy}^2\right)}$$

The same is done for $B_y$:

$$B_y = \frac{1}{2}\sum_x \left(\sum_a \widehat{p}(a, b = 1|xy) - \widehat{p}(a, b = -1|xy)\right)$$

$$\epsilon_{B_y} = \frac{1}{2}\sqrt{\sum_x \sum_a \left(\epsilon_{a1xy}^2 + \epsilon_{a,-1,xy}^2\right)}$$

The joint correlations are obtained as expected, and their error propagation is straightforward using the variance formula:

$$E_{xy} = A_x B_y = \widehat{p}(a = b|xy) - \widehat{p}(a \neq b|xy)$$

$$\epsilon_{A_x B_y} = \sqrt{\sum_{ab} \epsilon^2_{abxy}}$$

Geometric Visualization on the NS polytope

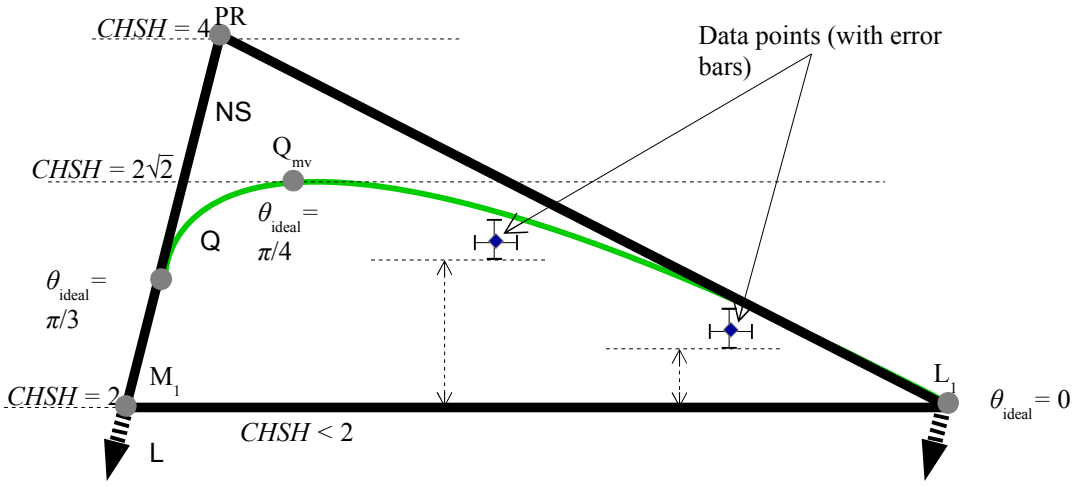



Figure 4.2.3.1: The non-local region of the NS polytope slice in figure 3.1.5.2.

Therefore each correlation matrix estimate can be visualized on the NS polytope as a point estimate with error bars. I therefore seek to find the minimum amount of entanglement certifiable amongst all points within the error margins at the various levels of characterization. The statistics obtained from the measurements considered in this project approximate the singlet state

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$$

## 4.3 Full Characterization Case

I now study the case when the knowledge of the measurement settings, as reported by the experimenters, is considered in the test for entanglement. With known measurement settings and dimension, the optimization is run over the state, which contains the only remaining free parameters. The algorithm can be summarized as follows:

$$\begin{aligned}
&\min\ N[\rho] \\
&s.t. |\widehat{p}(ab|xy) - \mathrm{tr}(\rho\ \Pi_x^a \otimes \Pi_y^b)| \leq \epsilon_{abxy} \forall a, b, x, y \\
&\Pi_0^a = \frac{1}{2}(\mathbb{I} + a\sigma_z), \Pi_1^a = \frac{1}{2}(\mathbb{I} + a(\sigma_z \cos\ 2\theta + \sigma_x \sin\ 2\theta)) \\
&\Pi_0^b = \frac{1}{2}(\mathbb{I} + b(\sigma_z \cos\ \theta + \sigma_x \sin\ \theta)), \\
&\Pi_1^b = \frac{1}{2}(\mathbb{I} + b(\sigma_z \cos\ 3\theta + \sigma_x \sin\ 3\theta))
\end{aligned}$$

I wrote two scripts, "minfidT2" and "minfidT3", which perform the above optimization for two different models for the state. The algorithms read in the Bloch angle as a parameter, and have the negativity as the objective function to be minimized under the constraint that the state is a two-qubit state and the statistics are compatible within errors. The script for minfidT3 can be found in Appendix D.

### 4.3.1 Noise Models

The state is modeled using two closely related noise models. The first one is known as the isotropic noise model, which is defined as:

$$\rho = v|\Psi^-\rangle\langle\Psi^-| + \frac{1-v}{4}\ \mathbb{I}$$

This isotropic noise model is assumes that the noise in the state comes from all states equally (hence the term isotropic). It is often seen in studies where one wants to model an uncharacterized noisy quantum channel. More general variations exist that use selected unitary operators that perform rotations on the state to model decoherence. The rotated states are modeled to occur in a given probability mixture – the isotropic noise model is the case where the rotated states are equiprobable, orthogonal, and form a basis that spans the Hilbert space of the state. It is thus clear that the noise model can thus be chosen according to the nature of the system under

study. A model characteristic of Type II down-converted photons takes the form of

$$\rho = p|\Psi^-\rangle\langle\Psi^-| + \frac{1-p}{2}\left(|01\rangle\langle 01| + |10\rangle\langle 10|\right)$$

The isotropic noise model assumes equal decoherence in all directions. However, for entangled pairs of photons generated from Type II SPDC, the conservation of angular momentum allows one to consider modeling the decoherence such that it preserves the anticorrelation of the photon polarizations.

In this study, the second noise model I choose to work with is a Bell-diagonal noise model, defined by:

$$\begin{aligned}\rho = {} & v_1|\Psi^-\rangle\langle\Psi^-| + v_2|\Psi^+\rangle\langle\Psi^+| \\ & + v_3|\Phi^+\rangle\langle\Phi^+| + (1 - v_3 - v_2 - v_1)|\Phi^-\rangle\langle\Phi^-|\end{aligned}$$

This noise model is good for the purposes of this study for the reason that it captures the behavior of the Type II down converted photons (by simply setting $v_1 + v_2 = 1$), and accommodates both models mentioned earlier while allowing for small probabilities of decoherence in the correlated directions. It also uses very few parameters, which gives high reliability when used in a heuristic optimization program.

### 4.3.2 Analytical Calculations

With a few-parameter model, I performed an analytical pre-calculation, and cast the objective function and constraints explicitly in terms of the parameters v. The correlations obtained by the measurements on the singlet state $|\Psi^-\rangle$ are given by:

$$p(ab|xy) = \frac{1}{4}\left(1 - ab\cos\left[(1 + 2\delta_{x,0}\delta_{y,1})\theta\right]\right)$$

for the other three Bell states, the correlations are:

$$|\Psi^+\rangle: \quad p(ab|xy) = \frac{1}{4}\left(1 - ab\cos\left[(1 + 2x + 2y)\theta\right]\right)$$

$$|\Phi^-\rangle : \quad p(ab|xy) = \frac{1}{4}\left(1 + ab\cos\left[(1+2x+2y)\theta\right]\right)$$

$$|\Phi^+\rangle : \quad p(ab|xy) = \frac{1}{4}\left(1 + ab\cos\left[(1+2\delta_{x,0}\delta_{y,1})\theta\right]\right)$$

These correlations are substituted into the noise models to obtain the correlations produced by the noisy state. For the isotropic noise model, the correlations read:

$$p(ab|xy) = \frac{1}{4} - \frac{v\ ab\cos\left[(1+2\delta_{x,0}\delta_{y,1}\theta\right]}{4}$$

For the Bell-diagonal noise model, the correlations are:

$$p(ab|00) = \frac{1}{4} + \frac{1-2(v_1+v_2)}{4}\cos\,\theta$$

$$p(ab|01) = \frac{1}{4} + \frac{1-2(v_1+v_2)}{4}\cos\,3\theta$$

$$p(ab|11) = \frac{1}{4} + \frac{v_3-v_1}{4}\cos\,\theta + \frac{1-v_3-v_1-2v_2}{4}\cos 5\theta$$

$$p(ab|10) = \frac{1}{4} + \frac{v_3-v_1}{4}\cos\,\theta + \frac{1-v_3-v_1-2v_2}{4}\cos 3\theta$$

The negativity of the state can also be calculated analytically in both cases:

$$N[\rho] = \max\left\{0, \frac{3v-1}{4}\right\}$$

$$N[\rho] = \max\left\{0, \frac{2v_1-1}{2}, \frac{2v_2-1}{2}, \frac{2v_3-1}{2}, \frac{1-2(v_3+v_2+v_1))}{2}\right\}$$

These are substituted into the script and a heuristic optimization is run over the parameters v using minfidT2 and minfidT3.

### 4.3.3 Results

Unfortunately, with the exception of the two smallest angular settings, the optimization problem found no compatible state when the Bloch angles reported by the experimenters were substituted into the scripts[14]. Given that the model for the state is sufficiently general for the nature of the setup, it is likely that there is noise also in the measurements. If that were the case, in the absence of error bars for the Bloch angles (since they were not provided), there is a large space of parameters to explore – maybe the angles between the settings were not the same, or the PBS has circular basis noise, or the angles simply were not set correctly, etc. The point is, there is no characterization provided for the uncertainty in the measurement settings, and I thus cannot make an informed choice on how to model it.

I chose to make one modification: rather than running the scripts for the reported Bloch angles, I varied the Bloch angle until I found the closest possible value of $\theta$ where a feasible solution exists. The angles obtained this way deviated from the reported angles by as little as 7.14% up to as high as 30% (for $\theta = 0.08$). The mean absolute deviation from the reported Bloch angles is 0.03075 rad or, in terms of the polarization angles[15], about $0.89^{o}$. The optimization then certifies near-maximal entanglement for all $\theta > 0$, and obtains fidelity $F > 0.97$ to the singlet state. For the Bell-diagonal model, the statistics for $\theta = 0$ certify no entanglement, and can simply be understood by noting that both $|\Psi^{-}\rangle$ and $|\Psi^{+}\rangle$ obtain the same statistics for the measurement settings, and so an equal mixture of both is also a compatible state.

14 It can be easily judged whether an optimization is feasible when the number of parameters is so small.

15 Polarization angles are in 2-to-1 correspondence to the Bloch angles, allowing this simple conversion.

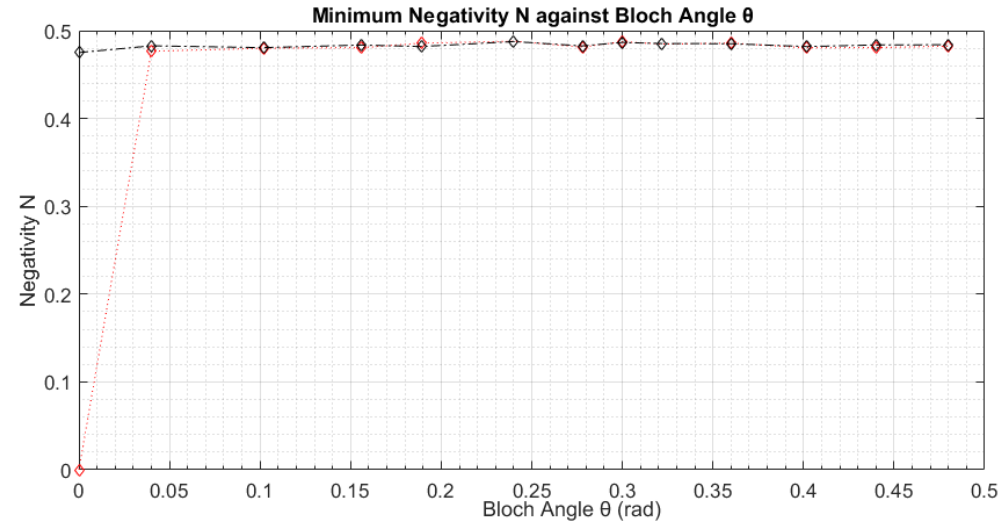



Figure 4.3.3.1 Plot of N[ρ] against θ. We find that other than θ = 0, both the isotropic model $N_{isotropic}$ (black data points) and the Bell-diagonal model $N_{bell\text{-}diagonal}$ (red data points) certify near-maximal entanglement.

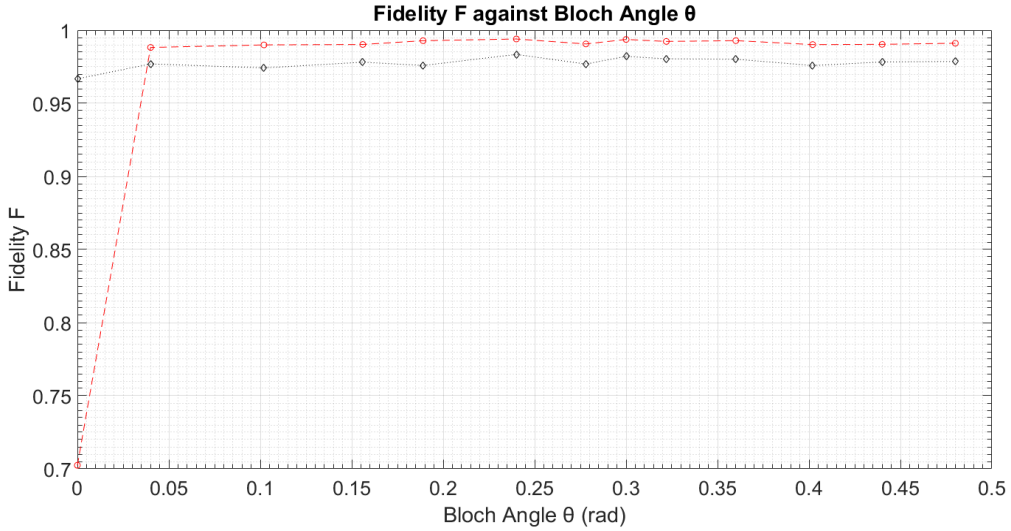



Figure 4.3.3.2 Plot of fidelity to the singlet against θ. We find that other than θ = 0, both the isotropic model $N_{isotropic}$ (black data points) and the Bell-diagonal model $N_{bell\text{-}diagonal}$ (red data points) have F > 0.97.

## 4.4 Device-Independent Case

### 4.4.1 Optimization Scheme

For the DI case, I run a semi-definite program using the method of Moroder et al (2013)[12], with minor modifications to account the experimental errors. The optimization scheme can be summarized as:

$$\min\ N = \chi_{-11}$$

$$s.t.\ \chi_{-} = \chi_{+} - \chi$$

$$\chi, \chi_{+}^{T_A}, \chi_{+}^{T_A} - \chi^{T_A} \geq 0$$
$$|\langle A_x B_y \rangle - \widehat{E}_{xy}| \leq \epsilon_{A_x B_y}$$

where the matrices χ are obtained via a positive map from an underlying quantum state of unknown dimension, where the matrix elements include the expectation values of the observables $A_x$ and $B_y$, the joint correlations and higher order moments thereof. The known moments are obtained from the correlation matrix, which are allowed to vary within the experimental error bounds. The algorithm obtains the minimum amount of entanglement certifiable based on the element of χ_ which corresponds to the negativity of the underlying state (exactly which element it is depends on construction). I wrote a script "DIOptData" which performs this SDP (script can be found in Appendix D).

The moment matrix is constructed by first considering the following decomposition of the underlying state (which may be of any dimension):

$$\rho = \rho_{+} - \rho_{-}$$

for any matrices $\rho_+$ and ρ_. If one imposes the constraint that both $\rho_+$ and ρ_ are PPT, one can work out that if ρ were NPT, then tr(ρ_) = N[ρ]. Moroder et al. (2013) then constructs the local-level moment matrix mapping ρ onto a difference between two moment matrices:

$$\rho = \rho_{+} - \rho_{-} \Rightarrow \chi[\rho] = \chi_{+} - \chi_{-}$$

where χ is defined according to (62) – (64) in section 3. In this study I consider $Q_{L1}$ (or local level 1), which gives a 9-by-9 moment matrix of the form:

$$\chi = \sum_{i,j,k,l=1}^{3} |i\rangle\langle j|\ \mathrm{tr}(\rho\ A_i A_k \otimes B_j B_l)$$

$$A_i \in \{\mathbb{I}, A_0, A_1\}, B_j \in \{\mathbb{I}, B_0, B_1\}$$

(exact form of $\chi$ in Appendix B)

The matrix $\chi_+$ is parametrized with a similar structure i.e. where similar moments appear in $\chi$, they must also be the same in $\chi_+$, since the same map is applied to both of them. The map is a positive map, so

$$\rho_+^{T_A}, \rho_-^{T_A} \geq 0 \Rightarrow \chi_+^{T_A}, \chi_-^{T_A} \geq 0$$

The moment matrices $\chi$ can give a lower bound on the negativity, since negativity is an entanglement monotone. If the topleft most element of $\chi$ is chosen to be the expectation value of the identity, it simply gives the trace of the underlying state of $\chi$ . For the DI case, therefore, I perform a SDP to certify the minimal entanglement certifiable with the topleft most element $\chi_-$ as the figure of merit:

$$\langle 11|\chi_-|11\rangle = \mathrm{tr}(\rho_-)$$

since $\mathrm{tr}(\rho_-) = N[\rho]$.

Accounting for experimental errors

The known moments (i.e. $A_0$, $B_0$, $A_1$, $B_1$, $A_0B_0$, $A_0B_1$, $A_1B_0$ , $A_1B_1$) in $\chi$ are fixed for an optimization problem where the statistics p(ab|xy) are error-free, but I insert addition parameters, allowing the known moments to vary between their lowest and highest possible values. i.e.

$$\langle A_x B_y \rangle = \sum_{ab} (\widehat{p}(a = b|xy) - \widehat{p}(a \neq b|xy)) + v_i \cdot \epsilon_{A_x B_y}$$

where the parameters $|v_i| \leq 1$.

### 4.4.2 Results

After accounting for the errors, the observed statistics violate the CHSH inequality for all $\theta \geq 0.102$ (considering the feasible Bloch angle), and thus certify the presence of entanglement for $\theta$ in that range. The amount of certifiable entanglement increases as the Bloch angle increases. The power of DI can be seen by making the following observation: in the full characterization case, the amount of certifiable entanglement depends on the Bloch angle input into the script– I simply chose the closest angle to the reported values that were feasible. If the reported knowledge on the measurements are known to be inaccurate, one can still certify the presence of entanglement using DI schemes if the statistics violate Bell

inequalities.

The bounds established by the DI optimization scheme, however, are loose: the optimal negativity only reaches up to ~0.2 for the whole range of θ, and certifies no entanglement for low angles. Fortunately, the angles at which the statistics become achievable by separable states are also the angles at which the reported measurement settings have feasible solutions for the fully characterized case.

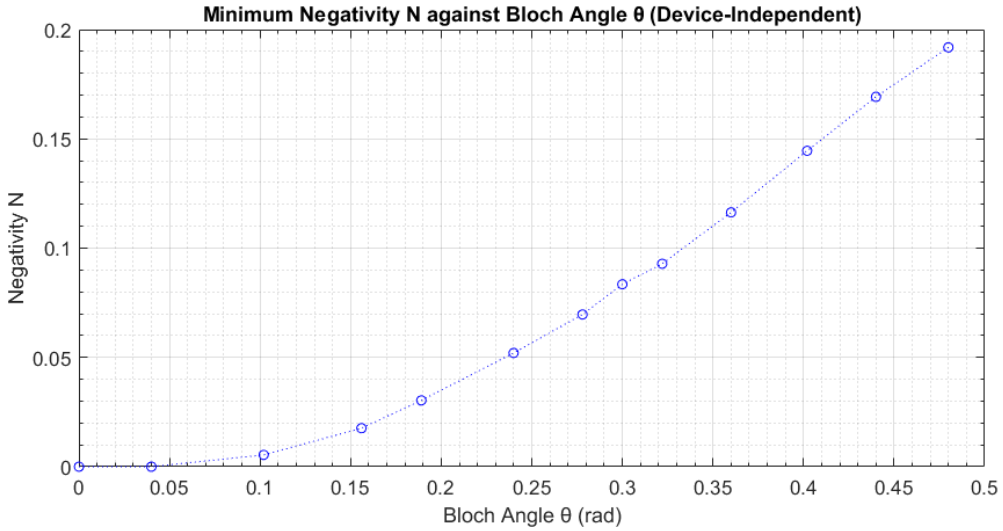



Figure 4.4.2.1 Plot of minimum negativity against the (feasible) Bloch angle.

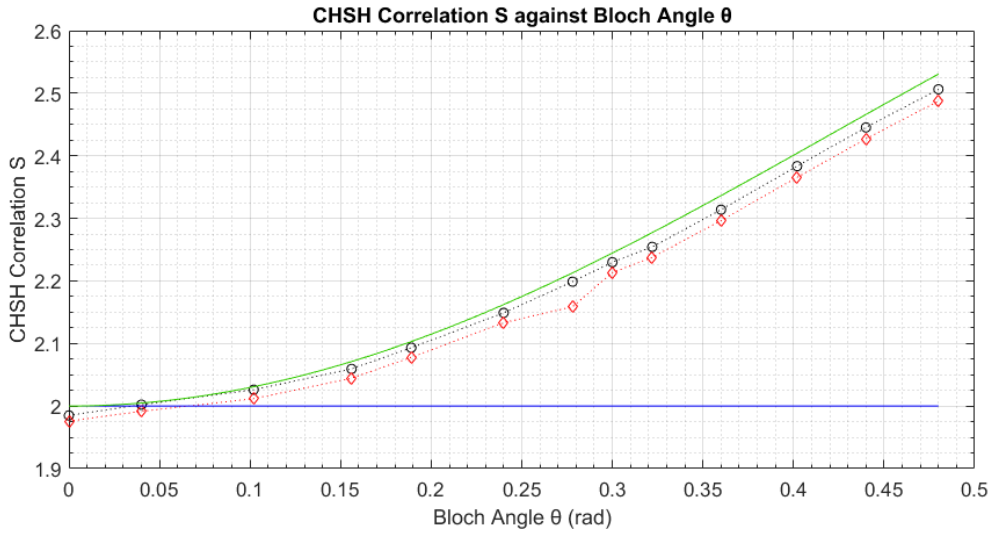



Figure 4.4.2.2 Plot of CHSH correlation against the (feasible) Bloch angle. The plot compares CHSH for ideal statistics (green curve), the experimental values before considering the error (black points) and after (red points)

### 4.4.3 Performance of the Algorithm on the whole range of θ

The optimization was also performed on ideal statistics, using correlations obtained by the measurements on a singlet state, which are known to be self-testing.
The statistics violate CHSH for a large range of θ, but are no longer on the boundary of Q for $\theta \in (\pi/3, 2\pi/3)$, and stop violating CHSH for $\cos\theta \leq \frac{1}{2}(\sqrt{3} - 1)$. While the statistics self-test outside of these ranges (see Appendix C for a short proof), the algorithm fails to reflect this fact despite having the correct qualitative behavior at crucial points. This inconsistency can be chalked up to the possibility that self-testing can somehow only be recovered at some level of the hierarchy, which would be an interesting direction for future exploration.

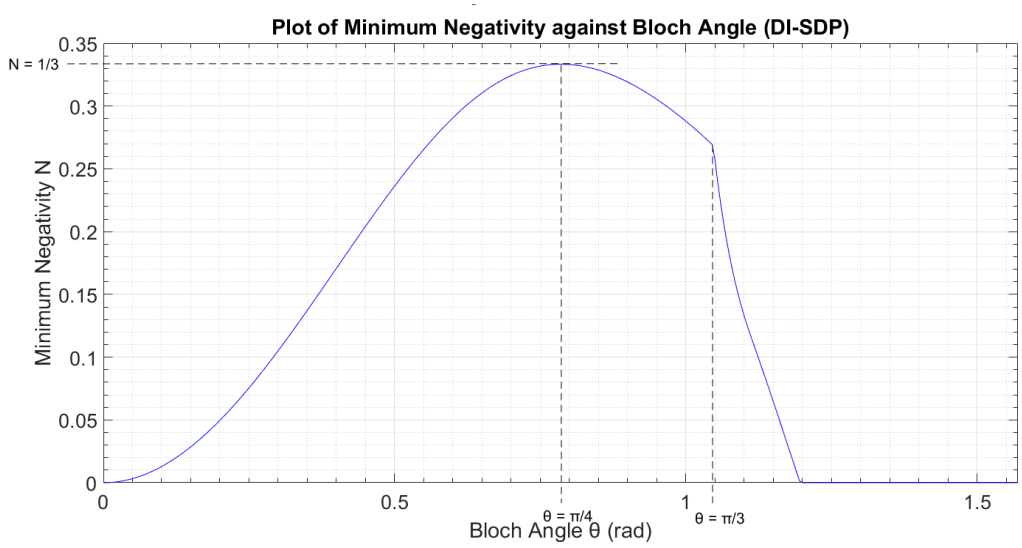



Figure 4.4.3.1 Plot of minimum negativity against the Bloch angle for ideal statistics. The qualitative features are correct: maximal entanglement is certified at the CHSH point ($\theta = \pi/4$) and the kink occurs when the statistics stop self-testing ($\theta = \pi/3$), dropping to zero when the statistics stop violating CHSH.

## 4.5 Semi-Device-Independent Case

### 4.5.1 Optimization Scheme

Now I focus on the main approach of the project outlined in section 4.1: certifying the minimum amount of entanglement compatible with a set of observed correlations within error margins, supplemented with the assumption of the Hilbert space dimension. The figure of merit chosen is the negativity.

$$\begin{aligned}
&\min\ N[\rho_{AB}] \\
&s.t. |\widehat{p}(ab|xy) - \mathrm{tr}(\rho_{AB}\ \Pi_x^a \otimes \Pi_y^b)| \leq \epsilon_{abxy}\ \forall a,b,x,y \\
&\rho_{AB} \in \mathcal{L}\left\{\mathcal{H}_A \otimes \mathcal{H}_B\right\}
\end{aligned}$$

$$\rho_{AB}, \Pi_x^a, \Pi_y^b \geq 0\ ,\ \sum_a \Pi_x^a = \sum_b \Pi_y^b = \mathbb{I}$$

The optimization is run over all states and measurements of dimension 2, subject to the constraint that the expectation values are compatible with the observed statistics within errors. In order to perform the task I wrote the script "ErrOptCorr", which reads in the correlation matrix, error matrix, and desired sample size, and outputs the state and measurements corresponding to the optimal result.

### 4.5.2 Self-Testing Consistency Check

A consistency check of the algorithm was done by running the algorithm on near-ideal statistics. I selected three points on the boundary of Q (refer to 4.2.3.1), $\theta_{ideal} = \pi/12$, $\pi/6$ and $\pi/4$, and input error bounds of the size ranging from $10^{-5}$ to $10^{-7}$. In all three cases, the result obtained for the negativity is ½, consistent with the fact that these statistics self-test.

### 4.5.3 Results

With the exception of $\theta = 0$, the semi-DI scheme certifies entanglement for all angles in the dataset. Because the scheme lets the measurements run free, entanglement is assured regardless of the accuracy of the reported measurement settings, even if we know that "something went wrong" from the full-characterization algorithm. Furthermore, the scheme certifies entanglement for statistics that do not violate the CHSH inequality once the errors are accounted for. At higher angles ($\theta > 0.24$), the certification power of the semi-DI scheme approaches that of the full characterization scheme – that is, the assumption of the measurements gives less additional advantage when the statistics become increasingly non-local.

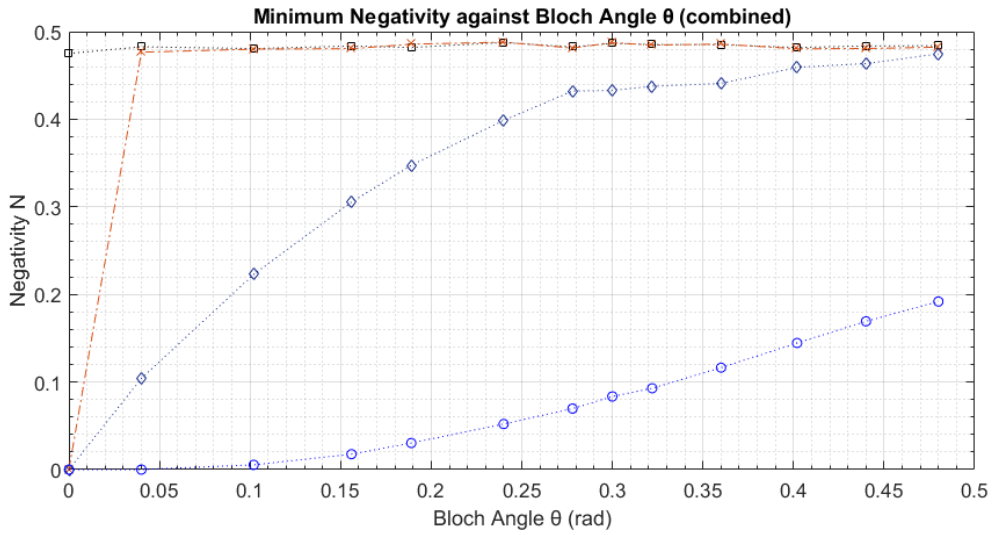



Figure 4.5.3.1 Plot of minimum negativity against the Bloch angle, comparing the three schemes: the full characterization scheme (black for isotropic model, red for Bell-diagonal model), the DI scheme (blue circles) and the semi-DI scheme (blue diamonds). At higher angles, $N_{SDI}$ approaches $N_{DI}$.

# Chapter 5

# Summary and Potential Future Work

Using experimental correlations, this project has demonstrated that a semi-DI implementation conclusively certifies entanglement in cases where the certification may be compromised by inaccurate characterization of the setup. In comparison to DI, the additional assumption on the dimension gives more certification power, useful in cases where the physical degree of freedom is known, but the measurements are not. In the case of this project, I dealt with statistics obtained from measurements on pairs of photons generated from Type II SPDC. The semi-DI scheme is also consistent with self-testing, certifying maximal entanglement for near-ideal correlations on the boundary of Q.

The DI scheme, taken from Moroder et al (2013), certifies non-maximal entanglement even on ideal statistics from measurement on the singlet, in contradiction with self-testing. The possible reasons for this include the insufficiency of the hierarchy at local-level 1, which is non-trivial given the current knowledge that the correlations in $Q_1$ (and by extension, any higher level such as $Q_{L1}$) and Q are already known to coincide.

It would thus be interesting to explore what level of the hierarchy restores self-testing when applied on the statistics obtained by measurements on the singlet, where one would expect to see the negativity = ½ throughout the range of angles. The moment matrices, while being a handy tool, are limited in their ability to detect flat regions on Q. Current literature has explored analytical expressions for the boundary Q, which are more effective at this task, making this a rich field for future exploration as well.

In the semi-DI case, the hope is to explore the relaxation of other assumptions such as IID, and implementing algorithms on experimental data which account for no-detection outcomes (i.e. data that is not post-selected), by treating the measurements as a three-outcome POVM.

# Appendix

### A. Boundary of the Generalized Bloch Sphere

The parametrization of qubit states in terms of the Pauli matrices is well known. This is the result of a study I had done when I briefly considered extending the Semi-DI optimization scheme to qutrits.

*Proposition A.1: A matrix is singular if and only if its Bloch vector is found on the boundary of the Generalized Bloch Sphere.*

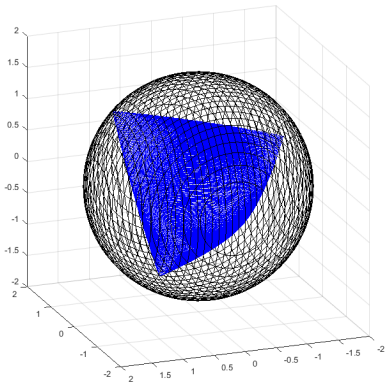
The fact that singular matrices lie on the boundary (and span the boundary) of the Generalized Bloch Sphere can be observed easily from the following argument: for any Hermitian matrix of d-dimensions, there exists a decomposition of the form

$$\rho = \frac{1}{d}(1 + \vec{n} \cdot \vec{G})$$

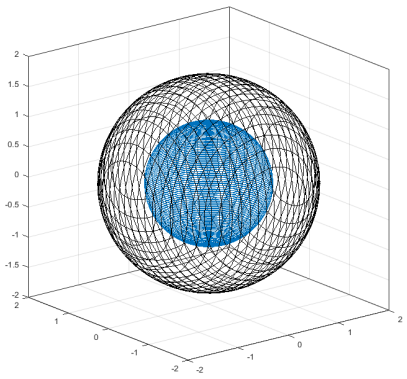
where '1' represents the identity matrix. If the density operator ρ is singular, one of the eigenvalues of **n·G** must be –1. Then going any further away from the origin in the direction of **n** will always make ρ indefinite. Thus if the Bloch sphere is characterized by $\rho \geq 0$, the boundary points will consist of singular matrices.

Therefore, exploiting this observation allows one to define an optimization problem to obtain 3D slices of the GBS. Choose spherical coordinates as parameters for the 3D space, and run the optimization over the radial coordinate r for a given φ and θ. Then according to the Proposition A.1, the objective function can be chosen to be the square of the minimum eigenvalue of the matrix. The optimization will converge with high reliability to the radial coordinate corresponding to the boundary of the GBS. This is captured by the script I wrote named "BlochArrayReal1", which has been used to plot the seven different geometries of 3D slices of the qutrit GBS.

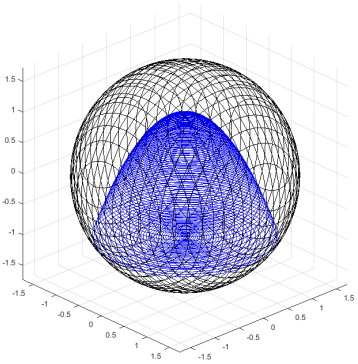
The following table showcases the 7 geometries of 3D slices of the qutrit GBS obtained from this optimization method.



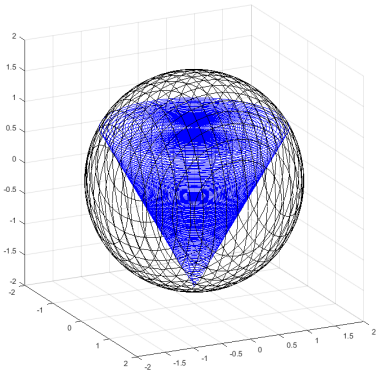
Obese Tetrahedron



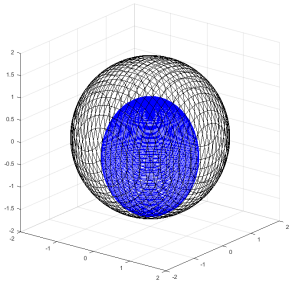
Sphere



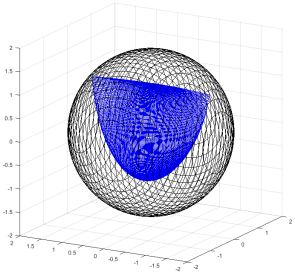
Paraboloid



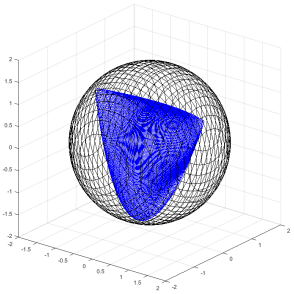
Cone



Ellipsoid



RS1



RS2

**B. Moment Matrix χ for local-level 1**

For local level 1, χ can be mnemonically treated as a tensor product of two 3x3 matrices of "local moments"

$$\chi = \begin{pmatrix} 1 & A_0 & A_1 \\ A_0 & 1 & A_0A_1 \\ A_1 & A_1A_0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & B_0 & B_1 \\ B_0 & 1 & B_0B_1 \\ B_1 & B_1B_0 & 1 \end{pmatrix}$$

where the strings containing A and B are concatenated upon multiplication, and can thereafter be treated as operators that satisfy the appropriate commutation relations. Once the full 9x9 matrix is constructed, the unknown moments can be set as free parameters for problem solving.

**C. Self-Testing of the statistics measured from the singlet**

A set of statistics self-test if they fulfill the following relation:

$$\left(\sum_{xy\neq ij} \arcsin\, E_{xy}\right) - \arcsin\, E_{ij} = \pm\pi \text{ for some } i,j \in \{0,1\}$$

and provided arccos($E_{xy}$) = 0 for at most one pair of (x,y). The statistics (the explicit expressions for them are found in section 4.3.2) for measurements settings according to Figure 4.2.1.2 give:

$$\arcsin(E_{xy}) = \tfrac{1}{2}\pi - (1 + 2\delta_{x,0}\delta_{y,1})\theta$$

which satisfies the self-testing criteria for (i,j) = (0,1) as long as $\theta \in (0, \pi/3]$ or $[2\pi/3, \pi)$.

## D. MATLAB codes for the main results

In this section I attach the codes used for the optimization problems and number-crunching. The syntax used is relatively simple, and should be understandable for those who have any experience with MATLAB. Otherwise, most of the syntax I used can be easily picked up by a simple Google search.

**D.1.** Pabxy22 – The code that produces the correlation and error matrices

```
function x = Pabxy22

x = zeros(16,2);

str = readarray;

numbers = numformat(str);
init =1;
for m = 1:4
    if m<4
        a = findindex(numbers,Inf,init);
    else
        a = length(numbers);
    end
    b = a - init+1;
    Y = zeros(b,1);
    for i =init:a
        j = i-init+1;
        Y(j) = numbers(i);
    end
    Ntot = length(Y);
    for j = 1:4
        if j<4
        k = sum(Y==j-1);
        x(4*(m-1)+j,1) = k/Ntot;
        else if j == 4
                x(4*(m-1)+j,1) = 1 - x(4*(m-1)+1,1) - x(4*(m-1)+2,1) - x(4*(m-
1)+3,1);
            end
        end
        x(4*(m-1)+j,2) = sqrt(k/Ntot^2 + k^2/Ntot^3);
    end
init = a+1;
end

end
```

and its child functions:

findindex – segregates the quats based on measurement setting

```
function x = findindex(r,a,init)
x = 0;
i = init;
```

```
while i <= length(r)
    if r(i) == a
        x = i;
        i = length(r);
    end
    i = i + 1;
end
if x == 0
    x = 'no such element exists';
end
end
```

---

readarray – reads the quats from the input file 'numbers.txt'

---

```
function x = readarray

Y = fopen('numbers.txt');
formatspec = '%c';
x = fscanf(Y,formatspec);

end
```

---

**D.2.** MinfidT3 – the script that optimizes negativity with respect to the Bell-diagonal noise model

---

```
function [x,p,fid] = minfidT3(P,t,MAXREP)

%F = @(v) sqrt(v(1));
N = @(v) 1/4*(abs(1-2*v(1))+abs(1-2*v(2))+abs(1-2*v(3))-(1-2*v(1))-(1-2*v(2))-(1-
2*v(3))+abs(2*(v(1)+v(2)+v(3))-1) - (2*(v(1)+v(2)+v(3))-1));

options = optimoptions('fmincon','Display','iter');
options.MaxFunEvals = 30000;
options.MaxIter = 30000;
trialf = zeros(1,MAXREP);
trialx = zeros(3,MAXREP);
k = 1;
i = 1;
while k <= MAXREP

[y,V,exitflag] = fmincon(N,rand(3,1),[1 1 1],1,[],[],[0 0 0],[1 1
1],@(v0)fidcon3(v0,P,t),options);

if exitflag > 0
    trialf(k) = V;
    trialx(:,k) = y;
    k = k+1;
end

if V == 0 && exitflag > 0
    k = MAXREP + 1;
end
```

```
    i = i+1;
end
jmin = 1;
for j = 2:MAXREP
    if trialf(j) < trialf(jmin)
        jmin = j;
    end
end

x = trialf(jmin);
p = trialx(:,jmin);
fid = sqrt(trialx(1,jmin));
```

and its constraint function:

fidcon3

```
function [c,ceq] = fidcon3(v,P,t)

P0000 = 1/4 + (1-2*(v(1)+v(2)))/4*cos(t);
P0100 = 1/4 - (1-2*(v(1)+v(2)))/4*cos(t);
P1000 = 1/4 - (1-2*(v(1)+v(2)))/4*cos(t);
P1100 = 1/4 + (1-2*(v(1)+v(2)))/4*cos(t);

P0001 = 1/4 + (1-2*(v(1)+v(2)))/4*cos(3*t);
P0101 = 1/4 - (1-2*(v(1)+v(2)))/4*cos(3*t);
P1001 = 1/4 - (1-2*(v(1)+v(2)))/4*cos(3*t);
P1101 = 1/4 + (1-2*(v(1)+v(2)))/4*cos(3*t);

P0010 = 1/4 + (v(3)-v(1))/4*cos(t) + (1-v(3)-v(1)-2*v(2))/4*cos(3*t);
P0110 = 1/4 - (v(3)-v(1))/4*cos(t) - (1-v(3)-v(1)-2*v(2))/4*cos(3*t);
P1010 = 1/4 - (v(3)-v(1))/4*cos(t) - (1-v(3)-v(1)-2*v(2))/4*cos(3*t);
P1110 = 1/4 + (v(3)-v(1))/4*cos(t) + (1-v(3)-v(1)-2*v(2))/4*cos(3*t);

P0011 = 1/4 + (v(3)-v(1))/4*cos(t) + (1-v(3)-v(1)-2*v(2))/4*cos(5*t);
P0111 = 1/4 - (v(3)-v(1))/4*cos(t) - (1-v(3)-v(1)-2*v(2))/4*cos(5*t);
P1011 = 1/4 - (v(3)-v(1))/4*cos(t) - (1-v(3)-v(1)-2*v(2))/4*cos(5*t);
P1111 = 1/4 + (v(3)-v(1))/4*cos(t) + (1-v(3)-v(1)-2*v(2))/4*cos(5*t);

T = [P0000; P0100; P1000; P1100; P0001;P0101;P1001;P1101;P0010;P0110;P1010;P1110;...
    P0011;P0111;P1011;P1111];

ceq = [];

c = max(abs(T-P(:,1))) - max(P(:,2));
```

**D.3.** DIOptData – the script that performs the SDP for the DI certification scheme

```
function [result ,y, z] = DIOptData(P)

J = sqrt(-1);
```

```
%A0 = 1/2;
%B0 = 1/2;
%A1 = 1/2;
%B1 = 1/2;
%A0B0 = 1/4*(1+cos(t));
%A0B1 = 1/4*(1+cos(3*t));
%A1B0 = 1/4*(1+cos(t));
%A1B1 = 1/4*(1+cos(t));
v = sdpvar(1,24);

A0 = 1/2*((P(1,1)+P(2,1)) + (P(5,1)+P(6,1))) +
v(17)/sqrt(4)*sqrt(P(1,2)^2+P(2,2)^2+P(5,2)^2+P(6,2)^2);
B0 = 1/2*((P(1,1)+P(3,1)) + (P(9,1)+P(11,1))) +
v(18)/sqrt(4)*sqrt(P(1,2)^2+P(3,2)^2+P(9,2)^2+P(11,2)^2);
A1 = 1/2*( (P(9,1)+P(10,1)) + (P(13,1)+P(14,1)) )+
v(19)/sqrt(4)*sqrt(P(9,2)^2+P(10,2)^2+P(13,2)^2+P(14,2)^2);
B1 = 1/2*( (P(13,1)+P(15,1)) + (P(5,1)+P(7,1)))+
v(20)/sqrt(4)*sqrt(P(13,2)^2+P(15,2)^2+P(5,2)^2+P(7,2)^2);

A0B0 = (P(1,1)) + v(21)*P(1,2);
A0B1 = (P(5,1)) + v(22)*P(5,2);
A1B0 = (P(9,1))+ v(23)*P(9,2);
A1B1 = (P(13,1)) + v(24)*P(13,2);

 u = sdpvar(1,25);
 Chi_Plus = [u(25) u(1) u(2) u(3) u(4) u(5) u(6) u(7) u(8);...
            0     u(1) u(9)+J*u(10) u(4) u(4) u(11)+J*u(12) u(7) u(7) u(13)+J*u(14);...
            0     0    u(2) u(5) u(11)-J*u(12) u(5) u(8) u(13)-J*u(14) u(8);...
            0     0    0 u(3) u(4)  u(5) u(15)+J*u(16) u(17)+J*u(18)
u(19)+J*u(20);...
            0  0 0 0 u(4) u(11)+J*u(12) u(17)+J*u(18) u(17)+J*u(18)
u(21)+J*u(22);...
            0     0    0    0    0  u(5) u(19)+J*u(20) u(23)+J*u(24)
u(19)+J*u(20);...
            0     0    0    0    0    0    u(6) u(7) u(8);...
            0     0    0    0    0    0    0    u(7) u(13)+J*u(14);...
            0     0    0    0    0    0    0    0    u(8)];



 Chi_Plus = (Chi_Plus + Chi_Plus') - diag(diag(Chi_Plus));

Chi = [1 A0 A1 B0 A0B0 A1B0 B1 A0B1 A1B1;...
       0 A0 v(1)+J*v(2) A0B0 A0B0 v(3)+J*v(4) A0B1 A0B1 v(5)+J*v(6);...
       0 0 A1 A1B0 v(3)-J*v(4) A1B0 A1B1 v(5)-J*v(6) A1B1;...
       0 0 0 B0 A0B0  A1B0 v(7)+J*v(8) v(9)+J*v(10) v(11)+J*v(12);...
       0 0 0 0 A0B0 v(3)+J*v(4) v(9)+J*v(10) v(9)+J*v(10) v(13)+J*v(14);...
       0 0 0 0 0 A1B0 v(11)+J*v(12) v(15)+J*v(16) v(11)+J*v(12);...
       0 0 0 0 0 0 B1 A0B1 A1B1;...
       0 0 0 0 0 0 0 A0B1 v(5)+J*v(6);...
       0 0 0 0 0 0 0 0 A1B1];

Chi = (Chi + Chi')- diag(diag(Chi));

F = [Chi>=0, Tx(Chi_Plus,2,[3 3])>=0, Tx(Chi_Plus - Chi,2,[3 3])>=0, abs(v(17))<=
1,abs(v(18))<= 1,abs(v(19))<= 1,...
    abs(v(20))<= 1,abs(v(21))<= 1,abs(v(22))<= 1,abs(v(23))<= 1,abs(v(24))<= 1];
```

```
N = Chi_Plus(1,1) - Chi(1,1);

obj = N;
optimize(F,obj);
result = value(Chi_Plus(1,1)-Chi(1,1));
y = value(Chi_Plus);
z = value(v);
```

## **D.4.** ErrOptCorr – the script the performs the semi-DI certification scheme

```
function [rho, p, px0, px1, py0, py1, NEG, TOTREP] = ErrOptCorr(P,MAXREP)
 % J = sqrt(-1);
 S0 = [1 0;0 1];
 Sx = [0 1;1 0];
 % Sy = [0 -J;J 0];
 Sz = [1 0;0 -1];

N = @(s) negativity(assign(s(1:9),4,1) + Imassign(s(10:15),4),[2 2]);
r = @(s) assign(s(1:9),4,1) + Imassign(s(10:15),4);

PIx_0 = @(s) 1/2*(S0 + s(16)*Sz);
PIy_0 = @(s) 1/2*(S0 + s(17)*Sz);

PIx_1 = @(s) 1/2*(S0 + s(18)*Sx + s(19)*Sz);
PIy_1 = @(s) 1/2*(S0 + s(20)*Sx + s(21)*Sz);


options = optimoptions('fmincon','Display','iter');
options.MaxFunEvals = 30000;
options.MaxIter = 30000;
trialf = zeros(1,MAXREP);
trialx = zeros(21,MAXREP);
k = 1;
i = 1;
while k <= MAXREP

[x,neg,exitflag] = fmincon(N,2*rand(21,1)-1,[],[],[],[],[],
[],@(p0)Errcon(p0,P),options);

if exitflag > 0
    trialf(k) = neg;
    trialx(:,k) = x;
    k = k+1;
end

if neg == 0 && exitflag > 0
    k = MAXREP + 1;
end

    i = i+1;
end
jmin = 1;
for j = 2:MAXREP
```

```
    if trialf(j) < trialf(jmin)
        jmin = j;
    end
end

p = trialx(:,jmin);
rho = r(p);
px0 = PIx_0(p);
px1 = PIx_1(p);
py0 = PIy_0(p);
py1 = PIy_1(p);
NEG = trialf(jmin);
TOTREP = i;

end
```

---

and its constraint function

Errcon

---

```
function [c, ceq] = Errcon(p,P)

% J = sqrt(-1);
S0 = [1 0;0 1];
Sx = [0 1;1 0];
% Sy = [0 -J;J 0];
Sz = [1 0;0 -1];

rho = assign(p(1:9),4,1) + Imassign(p(10:15),4);

PIx_0 = 1/2*(S0 + p(16)*Sz);
PIy_0 = 1/2*(S0 + p(17)*Sz);

PIx_1 = 1/2*(S0 + p(18)*Sx + p(19)*Sz);
PIy_1 = 1/2*(S0 + p(20)*Sx + p(21)*Sz);

T = zeros(16,1);

for i = 0:1
    for j = 0:1

        if i == 0 && j == 0

        T((2*i+j)*4+1) = sum(diag(rho*kron(PIx_0,PIy_0)));
        T((2*i+j)*4+2) = sum(diag(rho*kron(PIx_0,eye(2)-PIy_0)));
        T((2*i+j)*4+3) = sum(diag(rho*kron(eye(2)-PIx_0,PIy_0)));
        T((2*i+j)*4+4) = sum(diag(rho*kron(eye(2)-PIx_0,eye(2)-PIy_0)));

        end

        if i == 0 && j == 1

        T((2*i+j)*4+1) = sum(diag(rho*kron(PIx_0,PIy_1)));
        T((2*i+j)*4+2) = sum(diag(rho*kron(PIx_0,eye(2)-PIy_1)));
```

```
        T((2*i+j)*4+3) = sum(diag(rho*kron(eye(2)-PIx_0,PIy_1)));
        T((2*i+j)*4+4) = sum(diag(rho*kron(eye(2)-PIx_0,eye(2)-PIy_1)));

        end

        if i == 1 && j == 0

        T((2*i+j)*4+1) = sum(diag(rho*kron(PIx_1,PIy_0)));
        T((2*i+j)*4+2) = sum(diag(rho*kron(PIx_1,eye(2)-PIy_0)));
        T((2*i+j)*4+3) = sum(diag(rho*kron(eye(2)-PIx_1,PIy_0)));
        T((2*i+j)*4+4) = sum(diag(rho*kron(eye(2)-PIx_1,eye(2)-PIy_0)));

        end

        if i == 1 && j == 1

        T((2*i+j)*4+1) = sum(diag(rho*kron(PIx_1,PIy_1)));
        T((2*i+j)*4+2) = sum(diag(rho*kron(PIx_1,eye(2)-PIy_1)));
        T((2*i+j)*4+3) = sum(diag(rho*kron(eye(2)-PIx_1,PIy_1)));
        T((2*i+j)*4+4) = sum(diag(rho*kron(eye(2)-PIx_1,eye(2)-PIy_1)));
        end
    end
end
MXE = max(abs(T - P(:,1)) - max(P(:,2)));
c = [MXE -min(eig(rho)) p(16)^2-1 p(17)^2-1 p(18)^2+p(19)^2-1 p(20)^2+p(21)^2-1];
ceq = [];

end
```

---

The child functions assign and Imassign simply construct Hermitian matrices from an array of real numbers. Assign generates a symmetric real matrix, while Imassign generates an antisymmetric imaginary-valued matrix.

And finally,

BlochArrayReal1

---

```
function x = BlochArrayReal1(s,k2,k3,MAXREP,geo)
J = sqrt(-1);
x = zeros(s^2,3);
T = zeros(MAXREP,3);
if strcmp(geo,'tetra') == 1
    N = @(p)(min(eig((1/3*[1 p(1)*sin(p(2))*cos(p(3))
p(1)*sin(p(2))*sin(p(3));p(1)*sin(p(2))*cos(p(3)) 1 p(1)*cos(p(2));
p(1)*sin(p(2))*sin(p(3)) p(1)*cos(p(2)) 1])))^2;
else if strcmp(geo,'sphere') == 1
        N = @(p)(min(eig((1/3*[1 -J*p(1)*sin(p(2))*cos(p(3))
-J*p(1)*sin(p(2))*sin(p(3)); J*p(1)*sin(p(2))*cos(p(3)) 1 -J*p(1)*cos(p(2));
J*p(1)*sin(p(2))*sin(p(3)) J*p(1)*cos(p(2)) 1])))^2;
    else if strcmp(geo,'para') == 1
            N = @(p)(min(eig((1/3*[1+p(1)*cos(p(2)) 0 0; 0 1-p(1)*cos(p(2))
p(1)*sin(p(2))*exp(-J*p(3)); 0 p(1)*sin(p(2))*exp(J*p(3)) 1])))^2;
        else if strcmp(geo,'cone') == 1
```

```
                N = @(p)(min(eig((1/3*[1+1/sqrt(3)*p(1)*cos(p(2))
p(1)*sin(p(2))*exp(-J*p(3)) 0; p(1)*sin(p(2))*exp(J*p(3)) 1+1/sqrt(3)*p(1)*cos(p(2))
0; 0 0 1-2/sqrt(3)*p(1)*cos(p(2))])))))^2;
            else if strcmp(geo,'ellip') == 1
                    N = @(p)(min(eig((1/3*[1+1/sqrt(3)*p(1)*cos(p(2)) 0 0; 0
1+1/sqrt(3)*p(1)*cos(p(2)) p(1)*sin(p(2))*exp(-J*p(3)); 0 p(1)*sin(p(2))*exp(J*p(3))
1-2/sqrt(3)*p(1)*cos(p(2))])))))^2;
                else if strcmp(geo,'rs1') == 1
                        N = @(p)(min(eig((1/3*[1+p(1)*cos(p(2))
p(1)*sin(p(2))*cos(p(3)) p(1)*sin(p(2))*sin(p(3));p(1)*sin(p(2))*cos(p(3)) 1-
p(1)*cos(p(2)) 0;p(1)*sin(p(2))*sin(p(3)) 0 1])))))^2;
                    else if strcmp(geo,'rs2') == 1
                            N = @(p)(min(eig((1/3*[1+1/sqrt(3)*p(1)*cos(p(2))
p(1)*sin(p(2))*cos(p(3)) p(1)*sin(p(2))*sin(p(3));p(1)*sin(p(2))*cos(p(3))
1+1/sqrt(3)*p(1)*cos(p(2)) 0;p(1)*sin(p(2))*sin(p(3)) 0 1-
2/sqrt(3)*p(1)*cos(p(2)) ])))))^2;
                        end
                    end
                end
            end
        end
    end
end
options = optimset('Display','off');
for i = 1:s
    for j = 1:s
        x((i-1)*s+j,2) = i*(k2)/s;
        x((i-1)*s+j,3) = -k3 + j*2*k3/s;
    end
end
for j = 1:s^2
    a = x(j,2);
    b = x(j,3);
    for m=1:MAXREP
        T(m,:) = fmincon(N,[rand(1) a b],[],[],[],[],[],
[],@(p)nonlcon3(p,a,b),options);
    end
    x(j,1) = mode(T(:,1));
end
end
```

and its constraint function

nonlcon3

```
function [ c,ceq ] = nonlcon3(a,k2,k3 )
%UNTITLED2 Summary of this function goes here
%   Detailed explanation goes here
c = -a(1) ;
ceq = [a(2) - k2 a(3)-k3];


end
```

# Bibliography


[1] Bennet, A., Vértesi, T., Saunders, D. J., Brunner, N., & Pryde, G. (2014). Experimental Semi-Device-Independent Certification of Entangled Measurements. *Physical Review Letters, 113*(8). doi:10.1103/physrevlett.113.080405

[2] Bennett, C. H., Bernstein, H. J., Popescu, S., & Schumacher, B. (1996). Concentrating partial entanglement by local operations. *Physical Review A,53*(4), 2046-2052. doi:10.1103/physreva.53.2046

[3] Bennett, C. H., Brassard, G., & Mermin, N. D. (1992). Quantum cryptography without Bell's theorem. *Physical Review Letters, 68*(5), 557-559. doi:10.1103/physrevlett.68.557

[4] Clauser, J. F., Horne, M. A., Shimony, A., & Holt, R. A. (1970). Proposed Experiment to Test Local Hidden Variable Theories. *Physical Review Letters,24*(10), 549-549. doi:10.1103/physrevlett.24.549

[5] Einstein, A., Podolsky, B., & Rosen, N. (1935). Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review, 47*(10), 777-780. doi:10.1103/physrev.47.777

[6] Giustina, M., et al (2016). Significant-Loophole-Free Test of Local Realism with Entangled Photons. *Conference on Lasers and Electro-Optics*. doi:10.1364/cleo_qels.2016.fw4c.3

[7] Goh, K. T., Bancal, J., & Scarani, V. (2016). Measurement-device-independent quantification of entanglement for given Hilbert space dimension. *New Journal of Physics, 18*(4), 045022. doi:10.1088/1367-2630/18/4/045022

[8] Goyal, S. K., Simon, B. N., Singh, R., & Simon, S. (2016). Geometry of the generalized Bloch sphere for qutrits. *Journal of Physics A: Mathematical and Theoretical, 49*(16), 165203. doi:10.1088/1751-8113/49/16/165203

[9] Hensen, B., Bernien, H., Dréau, A. E., Reiserer, A., Kalb, N., Blok, M. S., . . . Hanson, R. (2015). Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature, 526*(7575), 682-686. doi:10.1038/nature15759

[10] Keyl, M. (2006). Quantum State Estimation And Large Deviations. *Reviews in Mathematical Physics, 18*(01), 19-60. doi:10.1142/s0129055x06002565

[11] Moroder, T., & Gittsovich, O. (2012). Calibration-robust entanglement detection beyond Bell inequalities. *Physical Review A, 85*(3). doi:10.1103/physreva.85.032301

[12] Moroder, T., Bancal, J., Liang, Y., Hofmann, M., & Gühne, O. (2013). Device-Independent Entanglement Quantification and Related Applications. *Physical Review Letters, 111*(3). doi:10.1103/physrevlett.111.030501

[13] Navascués, M., Pironio, S., & Acín, A. (2008). A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics, 10*(7), 073013. doi:10.1088/1367-2630/10/7/073013

[14] Pál, K. F., & Vértesi, T. (2009). Concavity of the set of quantum probabilities for any given dimension. *Physical Review A, 80*(4). doi:10.1103/physreva.80.042114

[15] Pearle, P. M. (1970). Hidden-Variable Example Based upon Data Rejection. *Physical Review D, 2*(8), 1418-1425. doi:10.1103/physrevd.2.1418

[16] Popescu, S., & Rohrlich, D. (1994). Quantum nonlocality as an axiom. *Foundations of Physics, 24*(3), 379-385. doi:10.1007/bf02058098

[17] Shalm, L. K., et al. (2016). A Strong Loophole-free Test of Local Realism. *Conference on Lasers and Electro-Optics*. doi:10.1364/cleo_qels.2016.fw4c.1

[18] Vidal, G., & Werner, R. F. (2002). Computable measure of entanglement. *Physical Review A, 65*(3). doi:10.1103/physreva.65.032314

[19] Wang, Y., Wu, X., & Scarani, V. (2016). All the self-testings of the singlet for two binary measurements. *New Journal of Physics, 18*(2), 025021. doi:10.1088/1367-2630/18/2/025021

[20] Wootters, W. K. (1998). Entanglement of Formation of an Arbitrary State of Two Qubits. *Physical Review Letters, 80*(10), 2245-2248. doi:10.1103/physrevlett.80.2245