Calculation Of Tight Bell Inequalities Through Robustness

Author: Neil Smith

Supervisors: Dr Mark Howard and Dr Pieter Kok

May 2018

Abstract

Bell Inequalities can provide evidence of nonlocality through violations and have important applications in device-independence and self-testing. Systematically deriving Bell Inequalities has been a computationally hard problem. In this work we introduce a computationally efficient, heuristic approach for creating Bell Inequalities by calculating the robustness of a quantum state. We applied our algorithm to the simplest scenario (2,2,2) and it yielded a family of tight CHSH inequalities for which the classical bound is $s(\alpha) = 2 - 3\alpha$. It also calculates trivial higher dimensional inequalities, which are uninteresting.

Contents

1 Introduction

In this work we introduce a new method to efficiently calculate tight Bell Inequalities, by the calculation of robustness of a quantum probability distribution produced from local measurements on a spatially separated quantum state. A Bell Inequality distinguishes the probability distributions that are achievable according to a classical theory and those that cannot, e.g. those that can only be described by Quantum Mechanics. Violations of Bell Inequalities therefore provide evidence of the non-classical nature of the universe. By observing the probability distributions of the outcomes of local measurements on spatially separated systems we can show that the universe can exhibit non-local behaviours that cannot be explained through any Local Hidden Variable (LHV) theory. In order to test the locality of nature, we need to derive experimentally testable Bell Inequalities, this is the goal of this work. Bell Inequalities not only provide evidence of non-locality but also have more practical applications in quantum information processing tasks, communication complexity tasks, random number generation, randomness certification, quantum state verification or self-testing and device-independent quantum-key distribution.

Non-locality refers to the non-local correlations that can be observed between parties making local measurements on entangled quantum systems. Non-locality has the potential to be used in new communication and information-processing devices and enable us to accomplish tasks that are currently either impractical or impossible to do with current technology. The study of non-locality through Bell Inequalities will help us to realise these technologies. Non-locality is observed through the correlations between spatially separated systems. These correlations can be greater than any achievable in a Local Hidden Variable theory. These correlations are not due to some hidden variable or cause that predetermines the correlations or due to a signal traveling at or less than the speed of light. Performing measurements on one particle part of a pair of entangled particles can instantaneously influence the state of the other particle, no matter how far apart they are. Surprisingly this is not in contradiction with causality as these "instantaneous signals" between the particles do not transmit any useful information. These signals are random, performing a measurement results in a random but probabilistic outcome. Yet they can still be exploited to perform tasks that are not possible with classical physics and technologies.

In section 2 we remind the reader of the relevant concepts and terminology used in the framework to study non-locality. We introduce the joint probability distributions of local measurements on separated systems, local hidden variable theories, local deterministic behaviours and the local polytope. We will introduce the idea of a Bell Inequality and how violations of inequalities are evidence of non-locality. We dicuss how we calculate these joint probabilities in quantum mechanics and how we can parametrise local measurements and quantum states. We remind the user of the facet dimension algorithm which calculates a measure of the tightness of a given Bell Inequality by its intersection with the extremal points of the local polytope. We introduce the optimization algorithm used to calculate the inequalities and how it works, giving an example. In section 3 we discuss how we applied this algorithm to produce inequalities in the CHSH (2,2,2) scenario. In section 4 we show the tests we checked the algorithm against and the results of these tests. In section 5 we show our findings. In section 6 we summarise our findings and postulate ideas for further investigation.

2 Preliminaries

2.1 Nonlocality

Here we explain the framework in which we study non-locality. There are n parties who each have access to their own system upon which they can make measurements and observe outcomes each labelled by a set of numbers. Each party k can make a measurement $m_k \in \mathcal{M}_k$ from a set of possible measurements $\mathcal{M}_k = \{1, 2, ..., M_k\}$ out of a total possible M_k measurements. When party k makes measurement m_k they can observe one outcome $d_{m_k,k} \in \mathcal{D}_{m_k,k}$ from a set of possible outcomes $\mathcal{D}_{m_k,k} = \{1, 2, ..., D_{m_k,k}\}$ out of a total possible $D_{m_k,k}$ outcomes. Any party k can also choose not to perform a measurement. We can think of this as them choosing to perform the zeroth measurement, which we denote as $m_k = 0$.

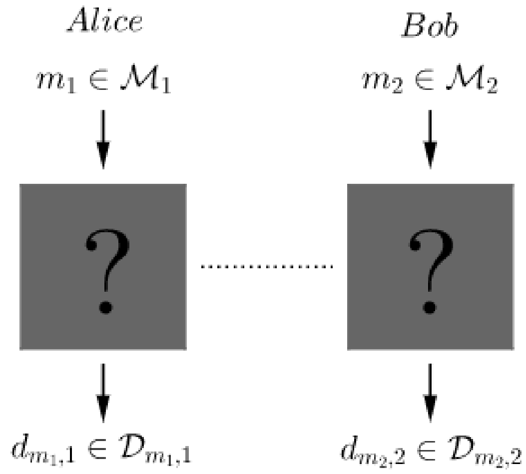

Figure 1: The Alice and Bob scenario, there are two parties who are very far apart from each other and each have access to their own system upon which they can make measurements and observe outcomes. Alice can make a measurement $m_1 \in \mathcal{M}_1$ from a set $\mathcal{M}_1$ of possible measurements and observe an outcome $d_{m_1,1} \in \mathcal{D}_{m_1,1}$ from a set $\mathcal{D}_{m_1,1}$ of possible outcomes. Similarly, Bob can make a measurement $m_2 \in \mathcal{M}_2$ from a set $\mathcal{M}_2$ of possible measurements and observe an outcome $d_{m_2,2} \in \mathcal{D}_{m_2,2}$ from a set $\mathcal{D}_{m_2,2}$ of possible outcomes.

The joint probability that party k makes a measurement m_k on their system and observes an outcome $d_{m_k,k}$ is denoted $P(d_{m_1,1} d_{m_2,2} ... d_{m_n,n} | m_1 m_2 ... m_n)$. We can associate with this probability distribution a vector (1), with elements that are the corresponding probabilities. We call this vector a behaviour as it represents a behaviour of the system. The number of elements of this vector l will just be the total number of possible ways of getting different outcomes. This can be found by calculating, for each party, the sum of the total number of possible outcomes for each of their measurements and then taking the product of all these numbers.

$$\vec{P}(d_{m_1,1} d_{m_2,2} ... d_{m_n,n} | m_1 m_2 ... m_n) \in \mathbb{R}^l \qquad l = \prod_{k=1}^{n} \left(\sum_{m_k=1}^{M_k} D_{m_k,k} \right) \tag{1}$$

2.2 Local Behaviours

A Local Hidden Variable (LHV) theory makes the assumption that any correlations observed between distant systems is due to some quantity that predetermines or causes the outcomes we observe when we make measurements upon the systems. This quantity $\lambda \in \Lambda$ is hidden from us and is known as a hidden variable. In particular, a LHV theory makes three assumptions about the way the universe behaves, these are:

1. **Locality**: The systems are influenced only by their local surroundings.
2. **Realism**: The properties of systems have well-defined values independent of external measurements.
3. **Free Will**: The parties can freely choose which measurements to perform.

The first, the assumption of locality, is the important one here. We assume that the parties are far enough apart that they cannot influence or communicate with each other, by sending a signal at or less than the speed of light. In order for the local hidden variable theory to be consistent with the special theory of relativity, the speed of this signal cannot be greater than the speed of light, the systems are unaffected by anything outside their light cones. Another way to phrase this is that each party's measurement choice does not depend on the measurement choices of the other parties and the choice of which measurement to make is made when they are far apart. The local probability distribution of one party's outcome is independent of the experiments performed by the other parties. Mathematically, this can be expressed for party 1 as (2).

$$P(d_{m_1,1}|d_{m_2,2}...d_{m_n,n}m_1m_2...m_n\lambda) = P(d_{m_1,1}|m_1\lambda) \quad \forall\, \lambda \in \Lambda \tag{2}$$

In quantum mechanics the properties of systems do not always have well-defined values until they are measured, they do not exhibit realism. In fact, the result depends on the specific experimental setup being used to measure it, this is known as Quantum Contextuality. In the assumption of free will we assume that the common cause λ is not correlated with their choice of measurements. In other words their measurement choices do not depend on the cause of the correlations, either directly or indirectly. This is also known as the assumption of measurement independence. Nothing influences the choices the parties make when performing the measurements.

If we make these three assumptions about how the universe works then there are only a certain set of behaviours that we would be able to observe, these are known as the local behaviours. All the local behaviours can be expressed in the form (3), where $q(\lambda) \geq 0$ is a probability distribution and $\int q(\lambda)\, d\lambda = 1$.

$$P(d_{m_1,1}d_{m_2,2}...d_{m_n,n}|m_1m_2...m_n) = \int_\Lambda q(\lambda)p(d_{m_1,1}|m_1\lambda)p(d_{m_2,2}|m_2\lambda)...p(d_{m_n,n}|m_n\lambda)\, d\lambda \tag{3}$$

Note that not all of these probability distributions represent physically realisable behaviours of systems, what we normally refer to as the local behaviours are infact a subset of these probability distributions (3) that also obey the normalisation (4) and no-signalling conditions (5), these are also known as the local correlations $\mathcal{L}$. The normalisation condition requires that the sum of all the probabilities for the difference outcomes for a given set of measurements is unity.

$$\sum_{d_{m_1,1},d_{m_2,2},...,d_{m_n,n}} P(d_{m_1,1}d_{m_2,2}...d_{m_n,n}|m_1m_2...m_n) = 1 \quad \forall \quad m_i \in M_i \quad \forall \quad i \in \{1,...,n\} \tag{4}$$

When we use the term "local behaviours" we are really refering to this subset of local behaviours. Any correlation or probability distribution that cannot be explained through a local hidden variable theory is known as a non-local correlation.

2.3 No-Signalling Correlations

Local hidden variable theories impose more assumptions upon the scenario than just the assumption that the parties cannot communicate faster than the speed of light. We can also consider what behaviours could occur given just this assumption about the parties, this is known as a no-signalling scenario. We assume that the parties are far enough apart that they cannot get any information about the other parties' experiments, this means that the probability of one party getting a given outcome is independent of any of the other parties' experiments. Mathematically this can be expressed as (5), which is equivalent to requiring that the marginals of the joint probability distribution are well-defined.

$$\sum_{d_k} P(d_{m_1,1}...d_{m_k,k}...d_{m_n,n}|m_1...m_k...m_n) = \sum_{d_k} P(d_{m_1,1}...d_{m_k,k}...d_{m_n,n}|m_1...m_k'...m_n) \tag{5}$$
$$\forall\, m_k, m_k' \in M_K, \quad \forall\, k \in \{1,...,n\}, \quad \forall\, d_i, m_i \in D_i, M_i\ (i \neq k)$$

The behaviours that satisfy these no-signalling conditions form the no-signalling set.

2.4 The Local Polytope

The set of local behaviours $\mathcal{L}$ form a convex set that is also a polytope. A polytope is a convex set with a finite number of extremal points[1], it is a generalisation of a three-dimensional polyhedron, it is an object with "flat" sides. The local behaviours therefore form a shape with "flat" sides in probability distribution space. The extremal behaviours (6) are those probability distributions that can be expressed as a deterministic[2]product of the local probability distributions of each party. These are the "corners" of this polytope.

$$P(d_{m_1,1}d_{m_2,2}...d_{m_n,n}|m_1m_2...m_n) = D_1(d_{m_1,1}|m_1)D_2(d_{m_2,2}|m_2)...D_n(d_{m_n,n}|m_n) \tag{6}$$

It is possible to determine the exact form of each of the extremal behaviours, each element of the probability distribution, by combining the deterministic property of the behaviours (6) with the no-signalling (5) and normalisation conditions (4) into a single set of constraints (7).

$$\sum_{d_{m_k,k}} D_k(d_{m_k,k}|m_k) = 1 \quad \forall\, m_k \in \mathcal{M}_k \quad \forall\, k \in \{1,...,n\} \tag{7}$$

We can determine from these constraints (7) the total number of extremal behaviours n_{ex}. This can be calculated by taking the product of the total number of possible outcomes $D_{m_k,k}$ for each measurement m_k of each party k (8).

$$n_{ex} = \prod_{k,m_k} D_{m_k,k} \tag{8}$$

The dimension of the local polytope is calculated using (9) and can be found by considering the normalization and no-signalling conditions as shown in Theorem 3.1 of [6].

$$\dim \mathcal{L} = \prod_{k=1}^{n} \left(\sum_{m_k=1}^{M_k} (D_{m_k,k} - 1) + 1 \right) - 1 \tag{9}$$

The local behaviours can be written as a linear combination, a convex combination, of the extremal behaviours (10). Here $c_i > 0$ are a series of coefficients that obey the constraint $\sum_i c_i = 1$. The local polytope is the convex hull of the local behaviours.

$$\mathcal{L} = \left\{ \vec{P}(d_{m_1,1}...d_{m_n,n}|m_1...m_n) \,|\, P(d_{m_1,1}...d_{m_n,n}|m_1...m_n) = \sum_i c_i D_1(d_{m_1,1}|m_1)D_2(d_{m_2,2}|m_2)...D_n(d_{m_n,n}|m_n) \right\} \tag{10}$$

2.5 Bell Inequalities

Bell Inequalities are expressions that can be used to help distinguish between what is and what isn't a local behaviour[3]. All classical behaviours satisfy the inequality, but non-classical behaviours may not. Some non-classical behaviours may satisfy the inequality and because of this not all Bell Inequalities may be able to perfectly distinguish between classical and non-classical behaviours. Bell Inequalities that can completely distinguish between classical and non-classical behaviours are called tight Bell Inequalities. Violations of Bell Inequalities therefore provide evidence of the non-classical nature of the universe. By observing the probability distributions

[1]For a more detailed explanation of an extremal point see [5] and for polytopes see [7].

[2]A deterministic probability is one in which the probability of an outcome is either one or zero $D_k(d_{m_k,k}|m_k) \in \{0,1\}$.

[3]There are many different types of inequalities that can be used to distinguish different sets of behaviours or correlations, which are all given different names. Bell Inequalities are named after John Bell the first to use them.

of the outcomes of local measurements on spatially separated systems, and checking that they satisfy a tight Bell Inequality, we can show that the universe can exhibit non-local behaviours that cannot be explained through any Local Hidden Variable (LHV) theory.

Here we consider only those Bell Inequalities that are linear in the probabilities and so represent half-spaces in probability distribution space. We can associate with every linear Bell Inequality a plane in probability space that defines the boundary of the half-space, this is sometimes called the Bell plane. Tight Bell Inequalities represent the facets of the local polytope, the intersection of all the tight Bell Inequalities is the local polytope. Other Inequalities may just touch an edge of the polytope or be some distance from it.

Bell Inequalities are not only used to provide evidence of the non-classical nature of the universe but are also related to device-independence and have practical applications. Some of which are quantum information processing tasks, communication complexity tasks, random number generation, randomness certification, quantum state verification or self-testing, and device-independent quantum-key distribution.

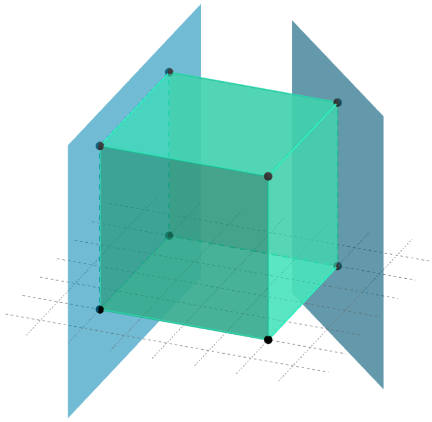

Figure 2: The geometric relationship between Bell Inequalities and the local polytope. Each linear Bell Inequality represents a half-space or series of planes of probability space. Bell planes are the planes corresponding to the equality. Tight Bell Inequalities represent the facets of the local polytope. There are an infinite number of Bell Inequalities for each scenario, some represent hyperplanes parallel to the facet, some are not. Some Bell planes are closer to the facet and others are further away.

A Bell Inequality is an inequality involving a sum of probabilities $P_{i_j,j}$ with coefficients $c_{i_j,j}$ as shown in (11). The equality relation defines the Bell hyperplane.

$$S = \sum_{i,j} c_{i_j,j} P_{i_j,j} = \vec{P} \cdot \vec{c} \leq S_{max} \tag{11}$$

We can also express these Bell Inequalities as an inner product of a vector of probabilities $\vec{P}$ that defines the space with a vector of coefficients $\vec{c}$ that determines the orientation of the corresponding set of planes of the Bell Inequality in this probability space. The vector $\vec{c}$ is the normal vector of these planes. In general, there may not be a unique way in which to express a particular Bell Inequality. Two Bell Inequalities are equivalent if they can be obtained from each other by any relabeling of the measurement settings, outcomes and parties.

The tight Bell Inequalities have a dimension (12) that is one less than that of the polytope (9).

$$\dim \mathcal{I} = \dim \mathcal{L} - 1 = \prod_{k=1}^{n} \left(\sum_{m_k=1}^{M_k} (D_{m_k,k} - 1) + 1 \right) - 2 \tag{12}$$

The first Bell Inequality was introduced by Bell in his 1964 paper [3] as a way of providing a test of whether or not quantum mechanics could be completed through the use of hidden variables. A violation of a Bell Inequality provides evidence that at least one of the assumptions made in a local hidden variable theory is incorrect, indeed violations were observed and they showed that no completion in this way was possible. To date, experiments have ruled out realism, showing that nature behaves nonlocally.[4]

An example of a tight Bell Inequality is the CHSH Inequality [4], given in (13), it is one of the most famous and well-studied Bell Inequalities.

$$\begin{aligned} P(11|11) - P(12|11) - P(21|11) + P(22|11) \\ P(11|12) - P(12|12) - P(21|12) + P(22|12) \\ P(11|21) - P(12|21) - P(21|21) + P(22|21) \\ -P(11|22) + P(12|22) + P(21|22) - P(22|22) \leq 2 \end{aligned} \tag{13}$$

It is the simplest tight Bell Inequality for which non-local correlations can be observed. It represents the correlations that can be observed between two spatially separated systems where the two parties each have a choice of two measurements to perform and can only observe two different possible outcomes for each of these measurements. This is usually denoted the $(2,2,2)$ scenario.

2.6 Measurements And Probabilities In Quantum Mechanics

In this section we revise how to calculate joint probabilities in quantum mechanics. Suppose there are n parties that each have access to their own individual system and can perform measurements upon it and observe outcomes. Suppose party k can choose from a set of measurements $\mathcal{M}_k = \{Q_{1,k}, Q_{2,k}, ..., Q_{M_k,k}\}$, each described by a measurement operator $Q_{m_k,k}$. The spectral decomposition of a measurement operator $Q_{m_k,k}$ is given by (14).

$$Q_{m_k,k} = \sum_{d_{m_k,k}=1}^{D_{m_k,k}} \lambda_{m_k,k}^{d_{m_k,k}} \Pi_{m_k,k}^{d_{m_k,k}} = \sum_{d_{m_k,k}=1}^{D_{m_k,k}} \lambda_{m_k,k}^{d_{m_k,k}} \left| e_{m_k,k}^{d_{m_k,k}} \right\rangle \left\langle e_{m_k,k}^{d_{m_k,k}} \right| \tag{14}$$

The $\Pi_{m_k,k}^{d_{m_k,k}}$ is the projector of the measurement operator $Q_{m_k,k}$ associated with finding measurement outcome number $d_{m_k,k}$ when party k performs measurement m_k on their system. The $\lambda_{m_k,k}^{d_{m_k,k}}$ and $\left| e_{m_k,k}^{d_{m_k,k}} \right\rangle$ are the corresponding eigenvalues and eigenvectors. In the (2,2,2) CHSH scenario (15), there are two parties and each can only make two different measurements. For each of these measurements there are two different possible outcomes and so there are two different eigenvalues and eigenvectors.

$$\begin{aligned} Q_{1,1} &= \sum_{d_{1,1}=1}^{2} \lambda_{1,1}^{d_{1,1}} \Pi_{1,1}^{d_{1,1}} = \sum_{d_{1,1}=1}^{2} \lambda_{1,1}^{d_{1,1}} \left| e_{1,1}^{d_{1,1}} \right\rangle \left\langle e_{1,1}^{d_{1,1}} \right| = \lambda_{1,1}^{1} \left| e_{1,1}^{1} \right\rangle \left\langle e_{1,1}^{1} \right| + \lambda_{1,1}^{2} \left| e_{1,1}^{2} \right\rangle \left\langle e_{1,1}^{2} \right| \\ Q_{2,1} &= \sum_{d_{2,1}=1}^{2} \lambda_{2,1}^{d_{2,1}} \Pi_{2,1}^{d_{2,1}} = \sum_{d_{2,1}=1}^{2} \lambda_{2,1}^{d_{2,1}} \left| e_{2,1}^{d_{2,1}} \right\rangle \left\langle e_{2,1}^{d_{2,1}} \right| = \lambda_{2,1}^{1} \left| e_{2,1}^{1} \right\rangle \left\langle e_{2,1}^{1} \right| + \lambda_{2,1}^{2} \left| e_{2,1}^{2} \right\rangle \left\langle e_{2,1}^{2} \right| \\ Q_{1,2} &= \sum_{d_{1,2}=1}^{2} \lambda_{1,2}^{d_{1,2}} \Pi_{1,2}^{d_{1,2}} = \sum_{d_{1,2}=1}^{2} \lambda_{1,2}^{d_{1,2}} \left| e_{1,2}^{d_{1,2}} \right\rangle \left\langle e_{1,2}^{d_{1,2}} \right| = \lambda_{1,2}^{1} \left| e_{1,2}^{1} \right\rangle \left\langle e_{1,2}^{1} \right| + \lambda_{1,2}^{2} \left| e_{1,2}^{2} \right\rangle \left\langle e_{1,2}^{2} \right| \\ Q_{2,2} &= \sum_{d_{2,2}=1}^{2} \lambda_{2,2}^{d_{2,2}} \Pi_{2,2}^{d_{2,2}} = \sum_{d_{2,2}=1}^{2} \lambda_{2,2}^{d_{2,2}} \left| e_{2,2}^{d_{2,2}} \right\rangle \left\langle e_{2,2}^{d_{2,2}} \right| = \lambda_{2,2}^{1} \left| e_{2,2}^{1} \right\rangle \left\langle e_{2,2}^{1} \right| + \lambda_{2,2}^{2} \left| e_{2,2}^{2} \right\rangle \left\langle e_{2,2}^{2} \right| \end{aligned} \tag{15}$$

To calculate the joint probability of party k performing measurement m_k and observing outcome number $d_{m_k,k}$ we take tensor product of the corresponding projectors of each party, apply this to the state ρ describing the whole system and calculate the trace as shown in (16).

$$P(d_{m_1,1} d_{m_2,2} ... d_{m_n,n} | m_1 m_2 ... m_n) = \mathrm{Tr}\left(\Pi_{m_1,1}^{d_{m_1,1}} \otimes \Pi_{m_2,2}^{d_{m_2,2}} \otimes ... \otimes \Pi_{m_n,n}^{d_{m_n,n}} \; \rho \right) \tag{16}$$

[4] Free will has not been ruled out but it is in fact scientifically impossible to rule out free will.

		Outcomes (d_1, d_2)			
		11	12	21	22
Measurements (m_1, m_2)	11	$P(11\|11)$	$P(12\|11)$	$P(21\|11)$	$P(22\|11)$
	12	$P(11\|12)$	$P(12\|12)$	$P(21\|12)$	$P(22\|12)$
	21	$P(11\|21)$	$P(12\|21)$	$P(21\|21)$	$P(22\|21)$
	22	$P(11\|22)$	$P(12\|22)$	$P(21\|22)$	$P(22\|22)$

Table 1: An example of the tabular form of probabilities in the (2,2,2) scenario.

Again, as an example we consider the (2,2,2) scenario. To calculate the probability that party 1 performs measurement 2 and gets outcome 1 and that party 2 performs measurement 1 and gets outcome 2 we take the corresponding projectors of the operators associated with the measurement and tensor product them. We apply this new operator to the quantum state ρ of the whole system and calculate the trace as shown in (17).

$$P(12|21) = \mathrm{Tr}\left(\Pi^1_{2,1} \otimes \Pi^2_{1,2}\ \rho\right) \tag{17}$$

Each party performs local measurements $Q_{m_k,k}$ on their own individual system, these local measurement operators can be represented as a state $|Q\rangle$ that can be parametrised by a series of angles. In the case of measurements with two outcomes i.e. local measurements on a qubit, this state can be represented as a state on the Bloch sphere and can be parametrised by two angles θ and ϕ as in (18).

$$|Q\rangle = \begin{pmatrix} \sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} e^{i\phi} \end{pmatrix} \tag{18}$$

The local measurement operators $Q_{m_k,k}$ expressed in terms of these angles would then be written as (19)

$$Q_{m_k,k} = |Q\rangle\langle Q| = \begin{pmatrix} \sin^2\frac{\theta}{2} & \sin\frac{\theta}{2}\cos\frac{\theta}{2}e^{-i\phi} \\ \sin\frac{\theta}{2}\cos\frac{\theta}{2}e^{i\phi} & \cos^2\frac{\theta}{2} \end{pmatrix} \tag{19}$$

In the case of the (2,2,2) scenario, the two parties each have access to a qubit, the whole state $|\psi\rangle$ can be parametrised by 4 complex coefficients (20), or equivalently by 6 real coefficients $\{\theta_1, \theta_2, \theta_3, \phi_1, \phi_2, \phi_3\}$.

$$|\psi\rangle = c_1|00\rangle + c_2|01\rangle + c_3|10\rangle + c_4|11\rangle \qquad |c_1|^2 + |c_2|^2 + |c_3|^2 + |c_4|^2 = 1 \tag{20}$$

For each complex number we have a parameter for both the real and imaginary parts, eight coefficients in total. We can pull out a global phase and force one of the coefficients to be real, leaving us with 7 real parameters. But the normalisation condition on the amplitude of the wavefunction imposes a constraint on these parameters, once we have determined the first 6 the 7th is already determined. So in total we are left with 6 real parameters. This parameterisation is just like that for a single qubit pure state on the Bloch Sphere (18), if there are N complex coefficients then there are $2N-2$ real parameters, which are angles. We can then vary each of these six angles from 0 to 2π to explore all of state space. The relationship between these six real parameters and the four complex coefficients is given in (21).

$$\begin{aligned} c_1 &= \cos\theta_1\cos\theta_2\cos\theta_3 \\ c_2 &= e^{i\phi_1}\cos\theta_2\cos\theta_3\sin\theta_1 \\ c_3 &= e^{i\phi_2}\cos\theta_3\sin\theta_2 \\ c_4 &= e^{i\phi_3}\sin\theta_3 \end{aligned} \tag{21}$$

The probabilities b and corresponding coefficients of a Bell Inequality y can each be expressed in a vectorial form like in (23) but they are also often expressed in a tabular form as shown in Table 1.

2.7 The Facet Dimension Algorithm

In previous work, we developed an algorithm to calculate a measure of the tightness of a given Bell Inequality. The algorithm calculates the Bell value of the Bell operator associated with the inequality, determined by a list of the coefficients of the probabilities. It does this for every extremal point of the local polytope. It keeps track of the largest value of the Bell expression and calculates how many extremal points give rise to this maximum value. It then calculates how many of these extremal behaviours are linearly independent, this is known as the facet dimension of the Bell Inequality plane that passes through these points. If the facet dimension is equal to the spatial dimension of the Bell Inequality then we know for certain that this Inequality is tight. If the facet dimension is less than this then the Bell Inequality is not perfectly tight but still touches an edge of the local polytope and still has some power to distinguish between classical and non-classical correlations. The algorithm requires the list of coefficients y of probabilities that describe the Bell Inequality and a list of the maximum number of possible outcomes $D_{m_k,k}$ for each measurement m_k of each party k which describes the scenario of the Bell experiment. The algorithm outputs the facet dimension of the given Bell Inequality and its maximum classical bound.

2.8 The CVX Algorithm And Robustness

In this work we introduce a new efficient method to calculate tight Bell Inequalities using a convex (CVX) optimization algorithm to calculate the robustness of a quantum state. Suppose we perform measurements on some quantum state and end up with a probability distribution b, which may lie outside the local set of behaviours. Every behaviour even these outside the local set can be written as a linear combination of the local deterministic extremal points. These coefficients can be written as a vector x, and this can be written mathematically as $Ax = b$ where A is a matrix that has columns that are the behaviours of these extremal points. There are an infinite number of different solutions x to this equation, but we are only interested in the one that minimizes the L1 norm of x. The robustness of the quantum state is the minimum value of the sum of the absolute values of the elements of x. We therefore have to solve the following optimization problem.

$$\begin{aligned} \underset{x\in\mathcal{R}}{\text{minimize}} \quad & ||x||_1 = \sum_{i}^{n_{ext}} |x_i| \\ \text{subject to} \quad & Ax = b \end{aligned} \tag{22}$$

The L1 norm describes the "distance" of a hyperplane associated with the probability distribution b from the local polytope. A LHV probability distribution will have a robustness of 1, because any local probability is a affine mixture of the deterministic extremal behaviours. The quantum probabilities distributions are further away and can be a negative non-affine mixture of these extremal behaviours and so can have a robustness greater than 1. The L1 norm describes how much this quantum probability distribution violates the Bell Inequality, it describes how much greater the quantum correlation is to the maximum possible local correlation. The optimization algorithm finds the Bell Inequality for which this violation is greatest and so by solving this optimization problem it is possible to derive tight Bell Inequalities. In order to check their tightness we can apply the facet dimension algorithm we developed in previous work as mentioned in section 2.7 to calculate the facet dimension of the Bell Inequality. If the facet dimension equals the spatial dimension of the Inequality then the inequality is tight, the higher the better.

With every minimization problem there is a corresponding maximization problem, we call the variable y associated with the corresponding maximization problem the dual variable of x, it is this quantity that we are interested in. The vector y is a list of coefficients of the probabilities of a Bell Inequality. The dual variable y that corresponds to the optimal solution of x represents a Bell Inequality that touches the local polytope. In minimizing the norm of x we are also minimizing the distance of this Bell plane from the local polytope. The inequalities y that are produced by this algorithm may not always be normalised and so are scaled such that the magnitude of the largest of any element y_i is unity. This is equivalent to just multiplying the whole Bell Inequality by a constant factor, the facet dimension is left unchanged and the Bell value is multiplied by this same factor.

		Outcomes (d_1, d_2)			
		11	12	21	22
Measurements (m_1, m_2)	11	0.8159	-1.0000	-1.0000	0.8159
	12	0.9386	-0.8773	-0.8773	0.9386
	21	0.9386	-0.8773	-0.8773	0.9386
	22	-0.8773	0.9386	0.9386	-0.8773

Table 2: The normalised list of probability coefficients y corresponding to the optimal solution x in (23).

As an example, in the simplest (2,2,2) scenario the equation $Ax = b$ takes the form (23) when the two parties perform the same two measurements $(\theta, \phi) = (\pi/2, 0)$ and $(\pi/2, \pi/3)$ on the Bell state $(1/\sqrt{2}, 0, 0, 1/\sqrt{2})$. This b is the probability distribution from Test 1. The x shown here is the optimal solution.

$$\begin{pmatrix} 1&1&0&0&1&1&0&0&0&0&0&0&0&0&0&0 \\ 1&0&1&0&1&0&1&0&0&0&0&0&0&0&0&0 \\ 1&1&0&0&0&0&0&0&1&1&0&0&0&0&0&0 \\ 1&0&1&0&0&0&0&0&1&0&1&0&0&0&0&0 \\ 0&0&1&1&0&0&1&1&0&0&0&0&0&0&0&0 \\ 0&1&0&1&0&1&0&1&0&0&0&0&0&0&0&0 \\ 0&0&1&1&0&0&0&0&0&0&1&1&0&0&0&0 \\ 0&1&0&1&0&0&0&0&0&1&0&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1&1&0&0&1&1&0&0 \\ 0&0&0&0&0&0&0&0&1&0&1&0&1&0&1&0 \\ 0&0&0&0&1&1&0&0&0&0&0&0&1&1&0&0 \\ 0&0&0&0&1&0&1&0&0&0&0&0&1&0&1&0 \\ 0&0&0&0&0&0&0&0&0&0&1&1&0&0&1&1 \\ 0&0&0&0&0&0&0&0&0&1&0&1&0&1&0&1 \\ 0&0&0&0&0&0&1&1&0&0&0&0&0&0&1&1 \\ 0&0&0&0&0&1&0&1&0&0&0&0&0&1&0&1 \end{pmatrix} \begin{pmatrix} 0.1704 \\ 0.1704 \\ -0.0171 \\ -0.0171 \\ 0.1704 \\ -0.0111 \\ 0.0514 \\ -0.0171 \\ -0.0171 \\ 0.0514 \\ -0.0111 \\ 0.1704 \\ -0.0171 \\ -0.0171 \\ 0.1704 \\ 0.1704 \end{pmatrix} = \begin{pmatrix} 0.5000 \\ 0.3750 \\ 0.3750 \\ 0.1250 \\ 0.0000 \\ 0.1250 \\ 0.1250 \\ 0.3750 \\ 0.0000 \\ 0.1250 \\ 0.1250 \\ 0.3750 \\ 0.5000 \\ 0.3750 \\ 0.3750 \\ 0.1250 \end{pmatrix} = \begin{pmatrix} P(11|11) \\ P(11|12) \\ P(11|21) \\ P(11|22) \\ P(12|11) \\ P(12|12) \\ P(12|21) \\ P(12|22) \\ P(21|11) \\ P(21|12) \\ P(21|21) \\ P(21|22) \\ P(22|11) \\ P(22|12) \\ P(22|21) \\ P(22|22) \end{pmatrix} \tag{23}$$

The corresponding inequality y found after renormalization is given in Table 2

3 Method

In order to analyse the effectiveness of this method in determining tight Bell Inequalities we applied it to the simplest (2,2,2) scenario, where we know there is a single tight Bell Inequality, the CHSH Inequality (13). We fix the local measurements of the two parties and vary the quantum state shared between them and calculate the corresponding probability distribution b for each of these states. The choice of measurements of the two parties are the same, they are $(Q_{1,1}, Q_{2,1}) = (Q_{1,2}, Q_{2,2}) = ((\frac{\pi}{2}, 0), (\frac{\pi}{2}, \frac{\pi}{3}))$. We solve the optimization problem for each of these states, exploring the whole of state space and calculate the facet dimension of each of these inequalities. Figure 3 shows a flowchart of this process in the file applyAlgorithm.m.

We first define the local measurements and solve the eigenvalue problem for the eigenvectors of these measurement operators. We use these eigenvectors to calculate the projectors of each of these measurements. We then calculate all of the possible tensor products of a projector from the first party with a projector of the second party. We then calculate the matrix A in the optimization problem that has columns that are the extremal points of the local polytope, for a given scenario this matrix is fixed. The calculation of all the extremal points of the local polytope is done during the facet dimension algorithm that we developed previously. We can just extract the part of the code that performs this calculation and use this to create a new function extremalPointLooper.m that takes a list of the maximum number of possible outcomes for each measurement of each party as input. This list defines the scenario and this function extremalPointLooper.m can be used for arbitrary scenarios.

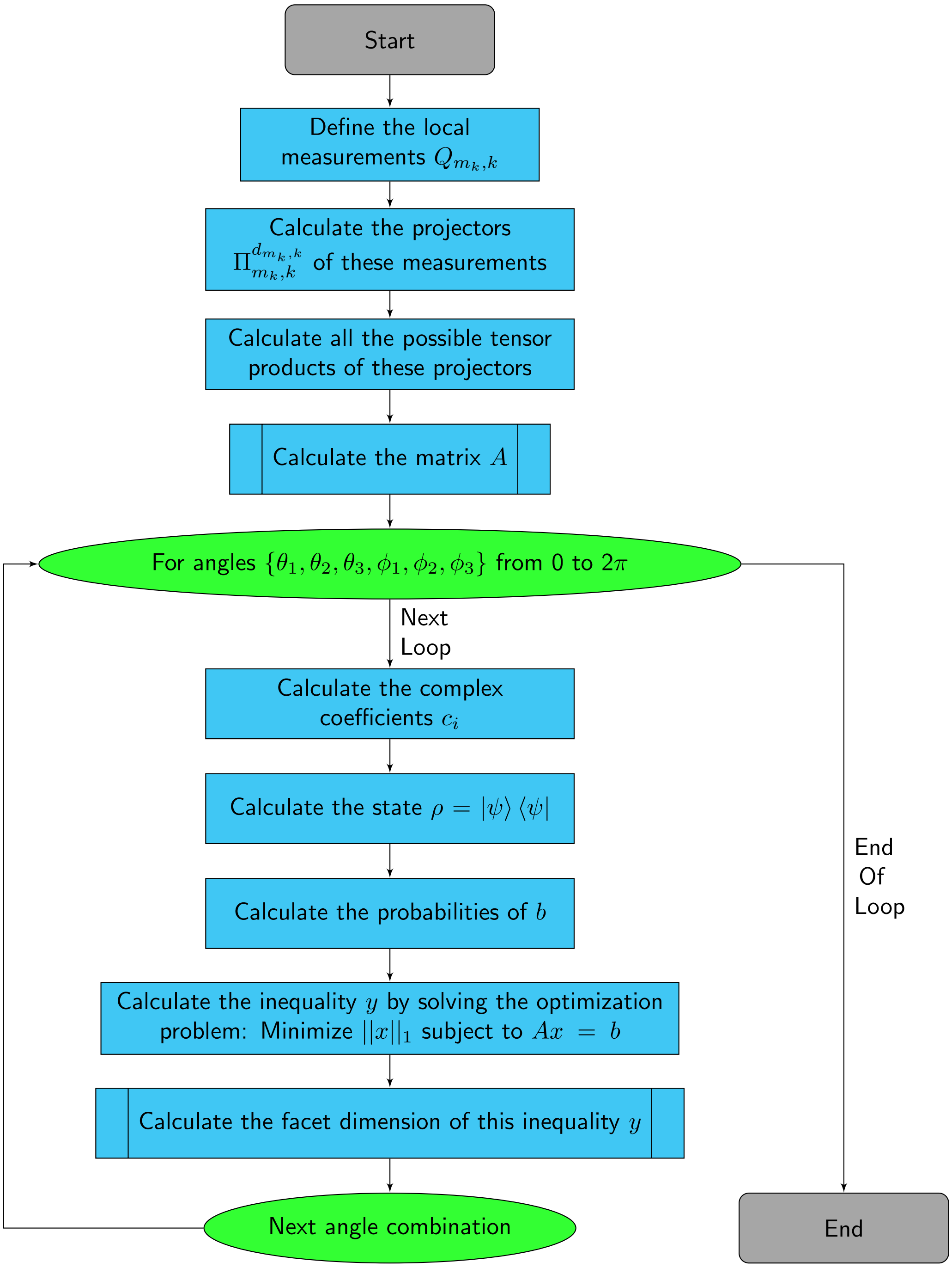

Figure 3: A flowchart of the file applyAlgorithm.m which calculates inequalities y by solving the optimization problem for a series of different states $|\psi\rangle$, for fixed measurements $Q_{m_k,k}$.

Test	Measurements $\{Q_{1,1}, Q_{1,2}, Q_{2,1}, Q_{2,2}\}$	State	b	y	Ref	Result
1	$\{(\frac{\pi}{2},0),(\frac{\pi}{2},0),(\frac{\pi}{2},\frac{\pi}{3}),(\frac{\pi}{2},\frac{\pi}{3})\}$	$(\frac{1}{\sqrt{2}},0,0,\frac{1}{\sqrt{2}})$	Table 4	Unknown	Table 1 in [1]	Pass
2	$\{(\frac{\pi}{2},\frac{\pi}{8}),(\frac{\pi}{2},\frac{\pi}{8}),(\frac{\pi}{2},\frac{5\pi}{8}),(\frac{\pi}{2},\frac{5\pi}{8})\}$	$(\frac{1}{\sqrt{2}},0,0,\frac{1}{\sqrt{2}})$	Table 5	CHSH (13)	Table 3 in [1]	Pass
3	$\{(\frac{\pi}{2},\frac{7\pi}{8}),(\frac{\pi}{2},\frac{7\pi}{8}),(\frac{\pi}{2},\frac{3\pi}{8}),(\frac{\pi}{2},\frac{3\pi}{8})\}$	$(\frac{1}{\sqrt{2}},0,0,\frac{1}{\sqrt{2}})$	Table 5	CHSH (13)	Table 3 in [1]	Pass

Table 3: The tests applied to the algorithm.

We then use the parametrisation (21) to loop over all the possible two-qubit states. In this work we only consider local measurements on a pure state. We vary each of the angles $\{\theta_1, \theta_2, \theta_3, \phi_1, \phi_2, \phi_3\}$ in 4 equal steps from 0 to 2π for a total of $4^6 = 4096$ possible different states. From these real coefficients we calculate the complex coefficients c_i and therefore the state $|\psi\rangle$. We can then calculate the probability distribution corresponding to these measurements and state b using the trace (16). Now that we have both A and b we can start the optimization algorithm and calculate the corresponding optimal dual solution y, the inequality. We can then check the tightness of this inequality by applying the facet dimension algorithm we developed in earlier work. We then just repeat this process for the next state.

4 Tests

There is no simple test to check that the matrix A has the correct form because we do not know any pairs of x and b, but there are still some things we can do. We know that the facet dimension algorithm is correct and the code for the calculation of A comes from this, so we can be pretty sure its correct already. Each column represents an extremal point which each have four probabilities that are 1 and the rest are zero and each column does have four 1s and they therefore obey the normalization conditions. Not all of the extremal behaviours are linearly independent and we know that the matrix A is not full rank. The dimension of the space spanned by these extremal points is $\text{rank}\, A - 1 = 8$ and we calculated the rank of A to be 9 agreeing with this result. Another way to check A has the correct form is to check how many 1s we would expect in each row. We know that each probability of an extremal behaviour can be written as a product of two local deterministic probabilities. Through equation (7) we know that not all of these local deterministic probabilities are independent of each other. In the (2,2,2) scenario, with 2-outcome measurements, these constraints tell us that these local probabilities come in pairs, if one probability is 1 then the other must be 0. In total there are 4 pairs of variables, each of these pairs can take two different values. Suppose we fix one of these deterministic probabilities to be 1 i.e. $P(11|11) = 1 = D_1(1|1)D_2(1|1)$, then we fix two of these pairs of variables, the other two pairs can be freely chosen. Therefore there are $2 * 2$ different ways in which $P(11|11) = 1$ and this must be true for every probability and so there must be 4 1s in each row of the matrix A, which we can see in (23).

To check that the calculation of the probability distribution b is correct we can test the code against measurements and states for which we do know the form of the probability distribution. The three tests we applied are given in Table 3. The expected probability distributions b are given in Tables 4 and 5. In Test 1 the algorithm was tested against the probability distribution b in Table 1 of [1]. The choice of measurements of the two parties are the same and they perform measurements on the Bell state. In Tests 2 and 3 the algorithm was tested against the probability distribution b in Table 3 of [1]. The choice of measurements of the two parties are also the same and they perform measurements on the Bell state. Both of these sets of measurements on this Bell state should result in the same b. Furthermore, the inequality y produced with this probability distribution b should be the tight CHSH Bell Inequality, which we do find. The inequality y should be the CHSH inequality because performing these measurements on this quantum state produces the maximal quantum violation. As another quick check that the probability distributions are in the right form, we can also check that these are normalised properly. They are all normalised, the sum of the probabilities is 4. We can also check that the states are valid states, by calculating their trace, they do all have a trace of 1.

		Outcomes (d_1, d_2)			
		11	12	21	22
Measurements (m_1, m_2)	11	$1/2$	0	0	$1/2$
	12	$3/8$	$1/8$	$1/8$	$3/8$
	21	$3/8$	$1/8$	$1/8$	$3/8$
	22	$1/8$	$3/8$	$3/8$	$1/8$

Table 4: The probability distribution b_1

		Outcomes (d_1, d_2)			
		11	12	21	22
Measurements (m_1, m_2)	11	0.4268	0.0732	0.0732	0.4268
	12	0.0732	0.4268	0.4268	0.0732
	21	0.0732	0.4268	0.4268	0.0732
	22	0.0732	0.4268	0.4268	0.0732

Table 5: The probability distribution b_2

5 Results

Surprisingly, for rotations on the state of a 2-qubit, we only found Inequalities with facet dimensions of 7 or 8. The inequalities with a facet dimension of 8 are thought to be trivial Inequalities. If the facet dimension is 8, which is greater than the maximum 7 then this suggests we are dealing with Inequalities with a spatial dimension of 8 or greater. The dimension of the local polytope, and of the space spanned by these extremal points is 8. Upon further calculation of one of these trivial inequalities it was found that all of the extremal points lie on this 8 dimensional Inequality, which is not surprising, all the extremal points lie in this 8 dimensional space.

For 4096 different states, 3712 (90.6%) of them were trivial inequalities with a dimension of 8 and the other 384 (9.38%) were tight Inequalities with a dimension of 7. Of the trivial inequalities, only 330 (8.06%) of these were distinct. The maximum classical correlation of these trivial inequalities varied from 1.8146 to 4. Of the tight inequalities, only 4 (0.0977%) of these were distinct. The maximum classical correlation of these tight inequalities were all the same 1.7577.

Tables 6, 7, 8 and 9 show these four different tight Inequalities. At first glance none of these appear to be the CHSH Inequality, but upon further inspection we noticed they have a form which is reminiscent of tight Inequalities in a different space [2]. If we set $\alpha = 0$ in any of these inequalities we do recover the CHSH inequality. These inequalities appear to be equivalent inequalities, we can transform from one to the other by a relabeling of the measurements and outcomes.

If we calculate the inequality y from test 1 we end up with another Inequality, in Table 10, with a similar structure to these distinct tight Inequalities. In fact, it has exactly the same form as the third distinct tight Inequality in Table 8 but with a different value of $\alpha = 0.0614$. But why they have this form and not the usual form of the CHSH inequality we do not know. In [2], the Bell Inequalities have Bell Values that are functions of this parameter α, it appears this is also the case here $s = 1.7577 = 2 - 3(0.0808)$ and $s = 1.8159 = 2 - 3(0.0614)$. Therefore it appears we have found a family of tight inequalities $y(\alpha)$ for which $s(\alpha) = 2 - 3\alpha$.

		Outcomes (d_1, d_2)			
		11	12	21	22
Measurements (m_1, m_2)	11	-1	$1-3\alpha$	$1-3\alpha$	-1
	12	$-1+2\alpha$	$1-\alpha$	$1-\alpha$	$-1+2\alpha$
	21	$-1+2\alpha$	$1-\alpha$	$1-\alpha$	$-1+2\alpha$
	22	$1-\alpha$	$-1+2\alpha$	$-1+2\alpha$	$1-\alpha$

Table 6: The first distinct tight Inequality y with $\alpha = 0.0808$ $s = 1.7577$

		Outcomes (d_1, d_2)			
		11	12	21	22
Measurements (m_1, m_2)	11	$-1+2\alpha$	$1-\alpha$	$1-\alpha$	$-1+2\alpha$
	12	$1-\alpha$	$-1+2\alpha$	$-1+2\alpha$	$1-\alpha$
	21	$1-\alpha$	$-1+2\alpha$	$-1+2\alpha$	$1-\alpha$
	22	$1-3\alpha$	-1	-1	$1-3\alpha$

Table 7: The second distinct tight Inequality y with $\alpha = 0.0808$ $s = 1.7577$

		Outcomes (d_1, d_2)			
		11	12	21	22
Measurements (m_1, m_2)	11	$1-3\alpha$	-1	-1	$1-3\alpha$
	12	$1-\alpha$	$-1+2\alpha$	$-1+2\alpha$	$1-\alpha$
	21	$1-\alpha$	$-1+2\alpha$	$-1+2\alpha$	$1-\alpha$
	22	$-1+2\alpha$	$1-\alpha$	$1-\alpha$	$-1+2\alpha$

Table 8: The third distinct tight Inequality y with $\alpha = 0.0808$ $s = 1.7577$

		Outcomes (d_1, d_2)			
		11	12	21	22
Measurements (m_1, m_2)	11	$1-\alpha$	$-1+2\alpha$	$-1+2\alpha$	$1-\alpha$
	12	$-1+2\alpha$	$1-\alpha$	$1-\alpha$	$-1+2\alpha$
	21	$-1+2\alpha$	$1-\alpha$	$1-\alpha$	$-1+2\alpha$
	22	-1	$1-3\alpha$	$1-3\alpha$	-1

Table 9: The fourth distinct tight Inequality y with $\alpha = 0.0808$ $s = 1.7577$

		Outcomes (d_1, d_2)			
		11	12	21	22
Measurements (m_1, m_2)	11	$1-3\alpha$	-1	-1	$1-3\alpha$
	12	$1-\alpha$	$-1+2\alpha$	$-1+2\alpha$	$1-\alpha$
	21	$1-\alpha$	$-1+2\alpha$	$-1+2\alpha$	$1-\alpha$
	22	$-1+2\alpha$	$1-\alpha$	$1-\alpha$	$-1+2\alpha$

Table 10: The Inequality y found from Test 1. $\alpha = 0.0614$ $s = 1.8159$ dimension $= 7$

6 Conclusions And Further Work

The algorithm successfully calculates a family of the tight CHSH Inequalities in the (2,2,2) scenario, each parametrised by a variable α. These reduce to the familiar form of the CHSH when $\alpha = 0$. The inequality $y(\alpha)$ and classical bound of these inequalities depend on the value of this parameter α. The dependence of the classical bound on this parameter is $s(\alpha) = 2 - 3\alpha$. The algorithm found this family of tight inequalities with a dimension of 7 and trivial inequalities with a facet dimension equal to the spatial dimension of the polytope 8. Most of the 4096 inequalities calculated are trivial inequalities but a substantial amount of the inequalities calculated $\approx 9\%$ are tight. All of these are the CHSH inequality, in some way or another, either by a relabeling of measurements or outcomes or both.

Further work is needed to confirm that this algorithm can be used to produce tight Inequalities in other scenarios. The next goal is to investigate whether this method could be used to produce new undiscovered inequalities. It would be interesting to investigate how the inequality produced varies with small changes in the probability distribution b to minimise the amount of calculation to produce the tight Inequalities. It would also be interesting to investigate why the inequalities produced have the form that they do and whether they also have this form in other scenarios. It would also be useful to see whether there are more inequalities with a dimension of 7 with a different classical bound s and confirm this relationship between the bound s and the parameter α and whether we find similar relationships in other scenarios. Here we only considered local measurements on pure states, It might be worth investigating whether measurements on mixed states could also lead to tight Inequalities.

The rows of the matrix A produced by the function extremalPointLooper.m is not in the correct form to be used in the optimization algorithm and first has to be transformed. It would be useful to ensure that the matrix A is output in the correct form to begin with. This code only works for the (2,2,2) scenario and the calculation of the probability distribution b and the transformation of y into the correct form for the facet dimension algorithm must be generalised to study other scenarios.

Overall the algorithm looks very promising and successfully calculates the tight CHSH inequality in the (2,2,2) scenario but further work is needed to verify that it can produce tight Inequalities in higher dimensions.

References

[1] Samson Abramsky, Rui Soares Barbosa, and Shane Mansfield. Contextual fraction as a measure of contextuality. *Phys. Rev. Lett.*, 119:050504, Aug 2017.

[2] Cédric Bamps and Stefano Pironio. Sum-of-squares decompositions for a family of clauser-horne-shimony-holt-like inequalities and their application to self-testing. *Phys. Rev. A*, 91:052111, May 2015.

[3] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964.

[4] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, 1969.

[5] Koon Tong Goh, J edrzej Kaniewski, Elie Wolfe, Tamás Vértesi, Xingyao Wu, Yu Cai, Yeong-Cherng Liang, and Valerio Scarani. Geometry of the set of quantum correlations. *Phys. Rev. A*, 97:022104, Feb 2018.

[6] S. Pironio. *Aspects of quantum non-locality*. PhD thesis, Université Libre de Bruxelles, 2004.

[7] G.M. Ziegler. *Lectures on Polytopes*. Graduate Texts in Mathematics. Springer New York, 2012.