Self-test and nonlocality detection

Hugo Abreu†, Fanny Terrier† and Hugo Thomas†

†Master in Quantum Information, Sorbonne Université

December 1, 2022

Abstract

This report presents the work carried out on self-testing and nonlocality detection in the context of the *Advanced Quantum Information Project* (PQIA). Nonlocality is studied through the prism of semi-definite programming - used herein for the numerical simulation and characterization of specific Bell scenarios - and its application to self-testing is examined. Particular attention is given to the retrieval of Bell inequalities, and a novel approach for the study of local and nonlocal behaviours using nonlinear optimization is described.

1 Introduction

Before the advent of quantum mechanics in the beginning of the 20^{th} century, physicists held three fundamental beliefs about the nature of the universe: *determinism* (or *causal determinism*), i.e. events in a given paradigm are bound by causality in such a way that any state is completely determined by prior states; *reality*, i.e. the universe exists independently of observation; and *locality* (or *local causality*), i.e. what happens in a given space-time region should not influence what happens in another, space-like separated region. Quantum mechanics challenged each one of these notions [NC10].

Non-locality in particular posed the most problems: it was the corner-stone of Einstein, Podolsky and Rosen's argument in 1935 for the incompleteness of quantum mechanics [EPR35]. Since it seemed to be possible for two spatially-separated parties sharing an entangled state to influence each other's measurements, resulting in *faster-than-light communications*, they argued that such a behaviour was paradoxical (this became known as the *EPR paradox*), and postulated that quantum mechanics should be compatible with a *hidden local variable description*, in which non-local effects would not be allowed.

For more than 30 years, *local-realist* description of quantum theory were not fully discarded. One had to wait until 1964, when Bell proposed an experiment which would definitively decide whether or not certain physical effects of quantum entanglement could be reproduced by local hidden variables: the *Bell test* [Bel64]. He proposed a condition, known as *Bell's inequality*, that any physical experiment has to satisfy if nature could be faithfully described by a classical local hidden variable theory. Subsequently, several experiments [AGR82] have been designed that have violated Bell's inequality, proving that no hidden variable theory can describe certain phenomenons predicted by quantum mechanics.

Since the 1960s, the field of *Bell nonlocality* has grown quite considerably [Bru+14], especially with the establishment of *quantum information science*, where nonlocality plays a central role [NC10]. It is at the heart of *device-independent* (*d.i.*) protocols such as *self-testing* [SB20], *d.i. quantum key distribution* [MPA11], *d.i. randomness generation* [Liu+18], and *delegated blind quantum computing* [Aar+17].

Section 2 sets the stage for the rest of the report, establishing the relevant theoretical background and framework. Subsequently, deeper attention is given to the two main bodies of work: nonlocality detection (Section 3) and the retrieval of Bell inequalities (Section 4). In Section 5, the limits of locality detection in noisy systems are studied, and lastly, in Section 6, a nonlinear approach to study of nonlocality is described.

2 Theoretical Framework

Nonlocal behaviours were first predicted and formulated in the context of entangled spin one-half particles [EPR35]. With the establishment of Bell nonlocality as a standalone field, specific physical scenarios were replaced with more implementation-agnostic frameworks [Bru+14]. In this report, *Bell scenarios*, first introduced by Bell in 1964 [Bel64], are considered.

In a Bell scenario, two parties *Alice* and *Bob*, sharing a quantum state $|\psi\rangle$ on which they can perform valid measurements M and N, are considered. Each party has m choices of measurement and Δ possible outcomes.

The *outputs* are denoted (a, b) and the *inputs* (x, y), as is depicted in Figure 1. For the rest of the report, only Bell scenarios in which $\Delta = 2$ and $m \in \{2, 3\}$ are considered.

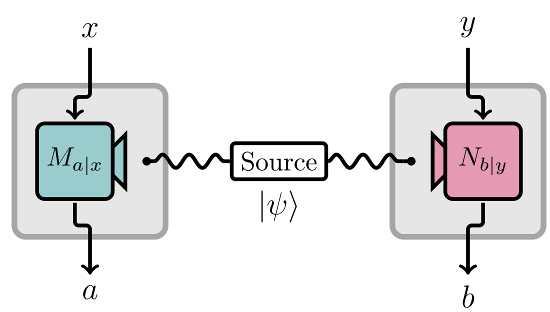

Figure 1: Bell scenario; a quantum state $|\psi\rangle$ is shared between two parties Alice and Bob, who receive x and y and output a and b according to measurement operators $M_{a|x}$ and $N_{b|y}$, respectively.

From one experiment to another, the outcomes a and b that are obtained may vary, even if Alice and Bob use the same measurement operators. As such, these outcomes can be described via a *probability distribution* [Bru+14].

This elements of this probability distribution are referred to as *joint probabilities*, and are denoted $p(a, b|x, y)$. The vector of all joint probabilities $\mathcal{P} = \{p(a, b|x, y)\}$ describes the operation of the Bell scenario and is referred to as the *behaviour* (Definition 2.1) or the *correlations* of the scenario.

Definition 2.1 (Behavior or correlations)**.** A behavior or correlations is a vector $\mathcal{P} = \{p(a, b|x, y)\}$ of all joint probabilities to obtain the output pair (a, b) given the input pair (x, y).

An interesting property of this definition, is that a behavior can be seen as a point $\mathcal{P}$ in the vector space $\mathbb{R}^{\Delta^2 m^2}$. This property will be useful for the expression of behaviours with semi-definite programming in Section 3.

To construct *valid behaviours*, one must ensure that it verifies certain conditions: essentially, the behaviour must constitute a *valid probability distribution* [Bru+14].

Proposition 2.2. *A behaviour $\mathcal{P}$ must satisfy two conditions: it must be* positive*, i.e.* $\forall a, b, x, y$*,*

$$p(a, b|x, y) \geq 0, \tag{1}$$

and it must be normalized*, i.e.* $\forall x, y \in \{1, m\}$*,*

$$\sum_{a,b=1}^{m} p(a, b|x, y) = 1. \tag{2}$$

Several types of behaviours exist: with this formalism, one can express both *local*, *quantum* and *no-signalling* behaviours, respectively referred to by the probability spaces to which they belong: $\mathcal{L}$, $\mathcal{Q}$ and $\mathcal{NS}$ (this will be addressed in more detail in Section 2.1). In the rest of the report, the space $\mathcal{K} \in \{\mathcal{L}, \mathcal{Q}, \mathcal{NS}\}$ will be used to refer to the probability space of an arbitrary valid behaviour.

2.1 Characterizing correlations

When considering two distant observers A and B performing measurements on a shared physical system, correlations arise as an important property of the system [Goh+18].

In this subsection, both local, quantum and no-signaling correlations are described.

2.1.1 Local correlations

Local correlations correspond to behaviours that are allowed in classical physics: informally, they enforce the fact that the two observers A and B cannot communicate before outputting a and b. Definition 2.3 formalizes this notion:

Definition 2.3 (Local correlations)**.** A behaviour $\mathcal{P}$ is said to be local if its joint probabilities $p(a, b|x, y)$ can be written

$$p(a, b|x, y) = \int_{\Lambda} q(\lambda) p(a|x, \lambda) p(b|y, \lambda) \mathrm{d}\lambda \tag{3}$$

where $\lambda \in \Lambda$ are local hidden variables with a probability density distribution $q(\lambda)$. Otherwise, the correlations are non local.

Local correlations can also be expressed in a simpler form, in terms of deterministic local hidden-variable models - in what is referred to as a local deterministic behaviour (Definition 2.4). The equivalence of these two definitions follows from the fact any local randomness present in $p(a|x, \lambda)$ and $p(b|y, \lambda)$ can be incorporated in the shared random variable λ [Bru+14].

Definition 2.4 (Local deterministic behaviours)**.** Let $\lambda = (a_1, \ldots, a_m \,;\, b_1, \ldots, b_m)$ be the assignment of outputs a_x, b_y for each inputs possible. The corresponding deterministic behaviour is

$$\mathbf{d}_\lambda(a, b|x, y) = \begin{cases} 1 & \text{if } a = a_x \text{ and } b = b_y, \\ 0 & \textit{otherwise.} \end{cases} \tag{4}$$

There are Δ^{2m} possible deterministic behaviours.

One can also give an inductive definition for local behaviours, which will be useful in the linear programming formulation of Section 3:

Proposition 2.5 (Local correlations)**.** *A behaviour $\mathcal{P}$ is local if and only if it can be written as a convex combination of deterministic behaviours, i.e*

$$\mathcal{P} = \sum_\lambda \mu_\lambda \mathbf{d}_\lambda \ , \ \mu_\lambda \geq 0 \ , \ \sum_\lambda \mu_\lambda = 1. \tag{5}$$

2.1.2 Quantum correlations

Quantum correlations correspond to behaviours that are allowed in quantum mechanics: they can possess nonlocal characteristics [Goh+18].

Let $\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$ be the joint Hilbert space of Alice and Bob and ρ_{AB} be the quantum states representing their shared physical system. Let $\{M_{a|x}\}$ and $\{N_{b|y}\}$ be the sets of measurement operators respectively on $\mathcal{H}_\mathcal{A}$ and $\mathcal{H}_\mathcal{B}$. These measurement operators satisfy

$$\forall x, a, \ M_{a|x} \succcurlyeq 0 \quad \text{and} \quad \forall x, \ \sum_a M_{a|x} = \mathbb{1}_A \ ,$$
$$\forall y, b, \ M_{b|y} \succcurlyeq 0 \quad \text{and} \quad \forall y, \ \sum_b M_{b|y} = \mathbb{1}_B \ ,$$

therefore $\{M_{a|x}\}$ and $\{N_{b|y}\}$ characterize POVMs [NC10], and can thus be experimentally constructed.

The set of behaviours achievable in quantum mechanics $\mathcal{Q}$ corresponds to the set of behaviours who verify the following condition:

Definition 2.6 (Quantum correlations)**.** A behaviour $\mathcal{P}$ is quantum if its elements can be written as

$$p(a, b|x, y) = \mathrm{Tr}\left(\rho_{AB} M_{a|x} \otimes N_{b|y}\right) \tag{6}$$

The previous definition works for pure states, but one can always consider a purification for arbitrary quantum states (Proposition 2.7).

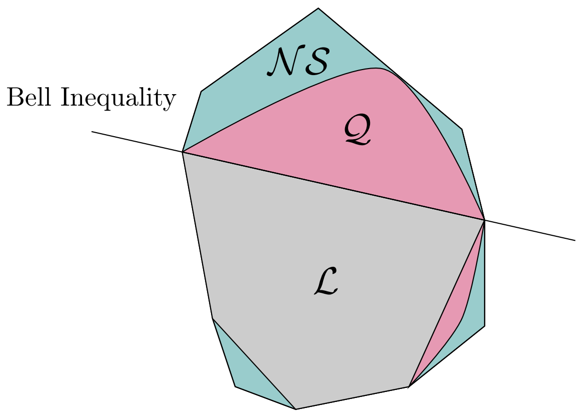

Figure 2: Geometric representation of local, quantum and no-signaling convex correlation spaces, with a Bell Inequality separating the quantum and local set.

Proposition 2.7. *Without loss of generality, one can take a purification Φ of ρ_{AB} and consider $M_{a|x}, N_{b|y}$ to be orthogonal projectors. Hence, quantum behaviour elements can be written*

$$p(a,b|x,y) = \langle\Phi| \, M_{a|x} \otimes N_{b|y} \, |\Phi\rangle \, . \tag{7}$$

2.1.3 No-signaling correlations

No-signalling constraints state that the marginal probabilities $p(a|x)$ (respectively $p(b|y)$) are independent of Bob's (respectively Alice's) measurement, i.e *Bob and Alice cannot signal instantaneously their outputs to each other by their choice of input* [Bru+14].

Definition 2.8 (No-signalling correlations)**.** A behaviour $\mathcal{P}$ is no-signalling if its elements fulfill the following constraints

$$\begin{aligned} \forall\ a, x, y, y' \quad \sum_{b=1}^{\Delta} p(a,b|x,y) &= \sum_{b=1}^{\Delta} p(a,b|x,y') \\ \forall\ b, y, x, x' \quad \sum_{a=1}^{\Delta} p(a,b|x,y) &= \sum_{a=1}^{\Delta} p(a,b|x',y) \end{aligned} \tag{8}$$

2.1.4 Correlation spaces and facet Bell inequalities

The relationship between the three types of correlations can be visualized in Figure 2. While the no-signaling and local sets can be easily characterized, the set of quantum correlations is more difficult to describe [Goh+18], but the property in Proposition 2.9 (introduced by Popescu and Rohrlich in [PR94]) always holds.

Proposition 2.9. *Let $\mathcal{NS}$, $\mathcal{Q}$ and $\mathcal{L}$ be the set of no-signalling, of quantum and of local correlations. They verify*

$$\mathcal{L} \subset \mathcal{Q} \subset \mathcal{NS},$$

and

$$\dim \mathcal{K} = (\Delta - 1)^2 m^2. \tag{9}$$

The following definitions are important for the characterisation of Facet Bell inequalities, which are important in the linear programming formulation.

Proposition 2.10. *The sets $\mathcal{L}, \mathcal{Q}$ and $\mathcal{NS}$ are closed, bounded and convex.*

Proposition 2.11. *Let $\mathcal{K}$ be a closed, bounded and convex set. For all $\mathcal{P}_1, \mathcal{P}_2$ in $\mathcal{L}$, and all α in $[0,1]$ the following holds:*

$$\alpha\mathcal{P}_1 + (1-\alpha)\mathcal{P}_2 \in \mathcal{K} \tag{10}$$

Proposition 2.12. *For all behaviours $\mathcal{P} \in \mathcal{K}$*

$$\exists \mathbf{s} \in \mathbb{R}^t, \;\; S_k \in \mathbb{R}, \quad \mathbf{s} \cdot \mathcal{P} \leq S_k \tag{11}$$

Definition 2.13 (Violation of Bell inequalities)**.** For the local set $\mathcal{L}$, this type of inequalities are called *Bell inequalities* and it is said to be violated by a behaviour $\mathcal{P}' \notin \mathcal{L}$ whenever

$$\mathbf{s} \cdot \mathcal{P}' > S_l$$

Definition 2.14 (Facet Bell inequalities)**.** Let $\mathbf{s} \cdot \mathcal{P} \leq S_l$ be a valid Bell inequality for the polytope $\mathcal{L}$. Then $F = \{\mathcal{P} \in \mathcal{L} | \mathbf{s} \cdot \mathbf{p} = S_l\}$ is a face of $\mathcal{L}$. Besides, if $\dim F = \dim \mathcal{L} - 1$, F is called a facet of $\mathcal{L}$ and the corresponding inequalities are called facet Bell inequalities.

Hence, Bell inequalities characterize the local set $\mathcal{L}$, and thus help to determine whether a behaviour $\mathcal{P}$ is local or not. Moreover, Tsirelson's bound provides a characterization of the quantum set [Cir80]. It is important to stress that since the set of quantum behaviours is not a polytope, there exist infinitely many Tsirelson bounds.

2.2 Self-testing of quantum states

Self-testing is a method to recover information on the physics of a quantum experiment, in a black box scenario, i.e. without needing to have or to trust a model of the experiment [SB20].

A general method to perform self-testing is now described.

Definition 2.15 (Self-testing of pure states)**.** The correlations $p(a,b|x,y)$ self-test the state $|\Phi^+\rangle_{A'B'}$ if for any state ρ_{AB} compatible with $p(a,b|x,y)$ and for any purification $|\Psi\rangle_{ABP}$ of ρ_{AB}, there exists a local isometry

$$\mathbf{\Phi}_A \otimes \mathbf{\Phi}_B : \mathcal{H}_A \otimes \mathcal{H}_B \longrightarrow \mathcal{H}_{A'\bar{A}} \otimes \mathcal{H}_{B'\bar{B}} \ ,$$

such that

$$\mathbf{\Phi}_A \otimes \mathbf{\Phi}_B \otimes \mathbb{1}_P \left[|\Psi\rangle_{ABP}\right] = |\Psi'\rangle_{A'B'} \otimes |\xi\rangle_{\bar{A}\bar{B}P} \tag{12}$$

Proposition 2.16 (Self-test and maximum violation of a Bell inequality)**.** *The extremal points of $\mathcal{Q}$ can be used to self-test both a state and a measurement [Goh+18] and these extremal points are known to be candidates for observing a maximal violation of a Bell inequality.*

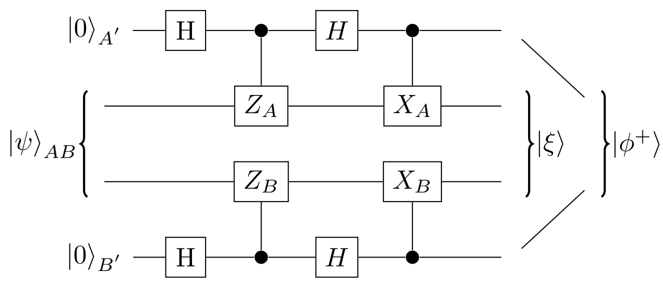

Figure 3: Swap gate isometry used to self-test the maximally entangled state of two qubits.

In most cases, it is possible to explicitly construct an isometry $\mathbf{\Phi}$ mapping the physical state $|\Psi\rangle$ to the reference state $|\Phi^+\rangle$ by using a partial Swap gate. For instance, let's focus on a Bell scenario with $x, y \in \{0, 1\}$. The main idea is that if the operators anticommute i.e., $\{A_0, A_1\} = \{B_0, B_1\} = 0$, we can construct operators Z_A, X_A, Z_B, X_B whose actions are analogous to those of the operators used in the 2-qubit Swap gate

$$\begin{aligned} Z_A &= \frac{1}{\sqrt{2}}(A_0 + A_1) \qquad Z_B = B_0 \\ X_A &= \frac{1}{\sqrt{2}}(A_0 + A_1) \qquad Z_B = B_0. \end{aligned} \tag{13}$$

Hence, if $|\Psi\rangle$ corresponds to the reference state up to an isometry, the Swap gate as pictured in Figure 3 enables the construction of the isometry

$$\mathbf{\Phi}\left[|\Psi\rangle_{ABP}\right] = |\Phi^+\rangle_{A'B'} \otimes |\xi\rangle_{ABP} \ , \tag{14}$$

where $|\Psi\rangle$ is a purification of the physical state, and A' and B' are the Hilbert spaces of the ancillary qubits. As such, the reference state is recovered in the ancillary space.

2.3 CHSH correlations

Consider outputs $a, b \in \{-1, 1\}$, and inputs $x, y \in \{0, 1\}$. Letting $\{A_x\}$ be the set of Alice's measurement observables (where $A_x = M_{1|x} - M_{-1|x}$ with $M_{a|x}$ the measurement operators), and $\{B_y\}$ the one of Bob, one can express the expected correlations as

$$\begin{aligned} \langle\Phi| A_0 \otimes B_0 |\Phi\rangle &= \frac{1}{\sqrt{2}} \quad , \langle\Phi| A_0 \otimes B_1 |\Phi\rangle &&= \frac{1}{\sqrt{2}} \\ \langle\Phi| A_1 \otimes B_0 |\Phi\rangle &= \frac{1}{\sqrt{2}} \quad , \langle\Phi| A_1 \otimes B_1 |\Phi\rangle &&= -\frac{1}{\sqrt{2}} \\ \langle\Phi| A_x \otimes \mathbb{1}_{\mathbb{B}} |\Phi\rangle &= 0 \ \forall x \quad , \langle\Phi| \mathbb{1}_{\mathbb{A}} \otimes B_y |\Phi\rangle &&= 0 \ \forall y. \end{aligned} \tag{15}$$

These correlations can be obtained if Alice and Bob share the quantum state

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big) \ , \tag{16}$$

and if Alice measures $A_0 = X$ and $A_1 = Z$, and Bob measures $B_0 = (Z + X)/\sqrt{2}$ and $B_1 = (X - Z)/\sqrt{2}$. These operators can be related to the measurement operators considered in Proposition 2.7 by

$$A_x = M_{1|x} - M_{-1|x} \text{ and } B_y = N_{1|y} - N_{-1|y} \tag{17}$$

and therefore, with Equation 7, we can write

$$\langle\Phi| A_x \otimes B_y |\Phi\rangle = \sum_{a,b} ab \ p(a, b|x, y) \tag{18}$$

and the vector $\mathcal{P} = \{p(a, b|x, y)\}$ can be derived by solving the systems of equations defined by Equations 15 and 18, and the normalization constraints in Equations 1 and 2.

With local correlations, the following Bell inequality stands [Bel64]

$$\langle A_0B_0\rangle + \langle A_1B_0\rangle + \langle A_0B_1\rangle - \langle A_1B_1\rangle \leq 2, \tag{19}$$

which will be referred to as the CHSH inequality.

On the other hand, with a quantum behaviour the Tsirelson's inequality is always fulfilled, and [Cir80]

$$\langle A_0B_0\rangle + \langle A_1B_0\rangle + \langle A_0B_1\rangle - \langle A_1B_1\rangle \leq 2\sqrt{2}, \tag{20}$$

for which the upper bound, called the Tsirelson's bound, is achieved by the previous shared state and measurement.

Therefore, the quantum strategy mentioned above achieves a maximum violation of the Bell Inequality 19,

2.4 Mayers-Yao correlations

For Mayers-Yao's self-test, the inputs are $x, y \in \{0, 1, 2\}$, the possible outputs $a, b \in \{0, 1\}$ and the expected correlations

$$\begin{aligned} \langle\Phi| A_z \otimes B_z |\Phi\rangle \qquad &= 1 \ \forall z \in \{0, 1, 2\} \\ \langle\Phi| A_x \otimes \mathbb{1} |\Phi\rangle = \langle\Phi| \mathbb{1} \otimes B_y |\Phi\rangle \qquad &= 0 \ \forall x, y \\ \langle\Phi| A_0 \otimes B_1 |\Phi\rangle = \langle\Phi| A_1 \otimes B_0 |\Phi\rangle \qquad &= 0 \\ \langle\Phi| A_0 \otimes B_2 |\Phi\rangle = \langle\Phi| A_1 \otimes B_2 |\Phi\rangle \qquad &= \frac{1}{\sqrt{2}} \\ \langle\Phi| A_2 \otimes B_0 |\Phi\rangle = \langle\Phi| A_2 \otimes B_1 |\Phi\rangle \qquad &= \frac{1}{\sqrt{2}}. \end{aligned} \tag{21}$$

These correlations can be achieved whenever Alice and Bob share the state

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right),$$

and their measurement observables are

$$\begin{aligned} A_0 = X_A \ , \ A_1 = Z_A, \ A_2 = \frac{1}{\sqrt{2}}(X_A + Z_A), \\ B_0 = X_B, \ B_1 = Z_B \ , \ B_2 = \frac{1}{\sqrt{2}}(X_B + Z_B). \end{aligned}$$

3 Non-locality detection

According to Propositions 2.5 and 2.6, it is possible to detect a non-local behaviour using linear programming (see Appendix A).

3.1 Linear Programming formulation

Let $\mathcal{P}$ be the behaviour for which one want to learn whether it is local, and let $\mathbb{1}$ be the behaviour corresponding to the random outcome strategy. Observe that $\mathbb{1}$ is a local behaviour. One can test if a behaviour $\mathcal{P}$ is nonlocal by solving the following linear program

$$\min_{\alpha,\vec{\mu}} \ \alpha \\ s.t. \begin{cases} (1-\alpha)\mathcal{P} + \alpha\mathbb{1} & \leq \sum_\lambda \mu_\lambda \mathbf{d}_\lambda \\ \sum_\lambda \mu_\lambda & = 1 \\ \alpha & \leq 1 \\ \mu_\lambda, \ \alpha & \geq 0 \quad \forall\lambda \end{cases} \tag{22}$$

that expresses how much $\mathcal{P}$ can be mixed with a random behavior to become local. In the case of an optimal value of $\alpha^* = 0$, $\mathcal{P}$ is a local behaviour since it can be written as

$$\mathcal{P} = \sum_\lambda \mu^*_\lambda \mathbf{d}_\lambda \ ,$$

where the μ^*_λ are the coefficients found at the optimum; whereas an optimal value of $\alpha^* > 0$ means that $\mathcal{P}$ is non-local.

3.2 CHSH correlations

The behaviour obtained by solving the system of equations induced by the quantum correlations exposed in Equation 15 is

$\mathcal{P}$		(a,b)			
		(-1, -1)	(-1, 1)	(1, -1)	(1, 1)
(x,y)	(0,0)	$\frac{\cos^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\cos^2(\pi/8)}{2}$
	(0,1)	$\frac{\cos^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\cos^2(\pi/8)}{2}$
	(1,0)	$\frac{\cos^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\cos^2(\pi/8)}{2}$
	(1,1)	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\cos^2(\pi/8)}{2}$	$\frac{\cos^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$

Table 1: CHSH probability distribution

and solving Linear Program 22 with the aforementioned behaviour gives an optimal value of

$$\alpha^*_{CHSH} \approx 1 - \frac{1}{\sqrt{2}} \ . \tag{23}$$

$\mathcal{P}$		(a,b)			
		(-1,-1)	(-1,1)	(1,-1)	(1,1)
(x,y)	$(0,0)$	1/2	0	0	1/2
	$(0,1)$	1/4	1/4	1/4	1/4
	$(0,2)$	$\frac{\cos^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\cos^2(\pi/8)}{2}$
	$(1,0)$	1/4	1/4	1/4	1/4
	$(1,1)$	1/2	0	0	1/2
	$(1,2)$	$\frac{\cos^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\cos^2(\pi/8)}{2}$
	$(2,0)$	$\frac{\cos^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\cos^2(\pi/8)}{2}$
	$(2,1)$	$\frac{\cos^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\sin^2(\pi/8)}{2}$	$\frac{\cos^2(\pi/8)}{2}$
	$(2,2)$	1/2	0	0	1/2

Table 2: Mayers-Yao probability distribution

3.3 Mayers-Yao's correlations

Similarly, the behaviour obtained by solving the system of equations induced by the quantum correlations shown in Equation 21 is given in Table 2 and Linear Program 22 gives an optimal value of

$$\alpha^*_{MY} \approx 0.1715 \ . \tag{24}$$

4 Inequalities from duality

The linear program for non-locality detection also comes in a useful dual form

$$\max_{\gamma,\mathbf{y}} \ \mathcal{P}\cdot\mathbf{y} + \gamma - \omega \\ s.t. \begin{cases} (\mathcal{P}-\mathbb{1})\cdot\mathbf{y} - \omega & \leq 1 \\ \gamma + \mathbf{d}_\lambda \cdot \mathbf{y} & \leq 0 \quad \forall\lambda \\ \mathbf{y} \in \mathbb{R}^n_+, \gamma \in \mathbb{R}, \omega \geq 0 \end{cases} \tag{25}$$

that can be interpreted as a way to express Bell inequalities. Letting δ^* be the optimal objective value, we can write

$$\mathcal{P}\cdot\mathbf{y}^* \ \leq \ \delta^* + \omega^* - \gamma^* \ . \tag{26}$$

Given that an inequality can be multiplied by a positive value, there exist many different inequalities we could derive, since

$$\begin{aligned} & \mathcal{P}\cdot\mathbf{y} && \leq \delta^* + \omega - \gamma \\ \Leftrightarrow \quad & M\left(\mathcal{P}\cdot\mathbf{y}\right) && \leq M\left(\delta^* + \omega - \gamma\right) \ \ \forall M \in \mathbb{R}^+ \ . \end{aligned}$$

4.1 CHSH correlations

The optimal solution is given by $\mathbf{y}^*_{\mathbf{CHSH}}$ in Table 3, $\gamma^*_{CHSH} \approx -2\sqrt{2}$ and $\omega^*_{CHSH} = 0$.
Therefore, we obtain the following Bell inequality

$$\mathcal{P}\cdot\mathbf{y}^*_{\mathbf{CHSH}} \ \lesssim \ 3.1213 \tag{27}$$

$\mathbf{y}^*_{\mathbf{CHSH}}$		(a,b)			
		$(-1,-1)$	$(-1,1)$	$(1,-1)$	$(1,1)$
(x,y)	$(0,0)$	1.414	0	1.414	1.414
	$(0,1)$	1.414	0	0	0
	$(1,0)$	0	0	0	1.414
	$(1,1)$	0	1.414	0	0

Table 3: Optimal solution of the dual

Multiplying Equation 27 by

$$M = \frac{\gamma^*_{CHSH}}{\delta^* + \omega^*_{CHSH} - \gamma^*_{CHSH}}$$

gives

$$\mathcal{P} \cdot \mathbf{z}^*_{\mathbf{CHSH}} \lesssim 2\sqrt{2}\ , \tag{28}$$

where

$$\mathbf{z}^*_{\mathbf{CHSH}} = M\, \mathbf{y}^*_{\mathbf{CHSH}}\ .$$

Using this dual with a local strategy such as

$\mathcal{P}_{loc}$		(a,b)			
		$(-1,-1)$	$(-1,1)$	$(1,-1)$	$(1,1)$
(x,y)	$(0,0)$	1	0	0	0
	$(0,1)$	0	0	1	0
	$(1,0)$	0	1/2	0	1/2
	$(1,1)$	0	1/2	1/2	0

Table 4: Example of a local behaviour

we obtain optimal values $\gamma^*_{loc} = -1$ and $\omega^*_{loc} = 0$. Hence, if $\mathcal{P}$ is local, we have

$$\mathcal{P}_{loc} \cdot \mathbf{y}^*_{\mathbf{loc}} \leq 2 \tag{29}$$

4.2 Mayers-Yao's correlation

The optimal solution of the Dual 25 is given by $\mathbf{y}^*_{\mathbf{MY}}$ in Table 5, $\gamma^*_{MY} \approx 3.313$ and $\omega^*_{MY} = 0$. Besides, we can notice that $\gamma^*_{MY} \approx 8\sqrt{\alpha^*_{MY}}$, where α^*_{MY} comes from Equation 24. Therefore, we obtain the following Bell inequality

$$\mathcal{P} \cdot \mathbf{y}^*_{\mathbf{MY}} \lesssim 3.49 \tag{30}$$

With a local strategy, we obtain again the following inequality

$$\mathcal{P}_{loc} \cdot \mathbf{y}^*_{\mathbf{loc}} \lesssim 2 \tag{31}$$

From the optimal solution $\mathbf{y}^*_{\mathbf{MY}}$ obtained with the quantum correlations shown in Table 5, one can notice that the input $x = 0$ on Alice's side is never used. Therefore, one could assume that Alice needs only two measurements instead of three to reproduce Mayers-Yao's correlations. Similarly, it can be noticed that the input $y = 1$ is never used by Bob in the optimal strategy.

$\mathbf{y}^*_{\mathbf{MY}}$		(a,b)			
		$(-1,-1)$	$(-1,1)$	$(1,-1)$	$(1,1)$
(x,y)	$(0,0)$	0	0	0	0
	$(0,1)$	0	0	0	0
	$(0,2)$	0	0	0	0
	$(1,0)$	0	1.657	0	0
	$(1,1)$	0	0	0	0
	$(1,2)$	0	0	0	1.657
	$(2,0)$	1.657	0	0	0
	$(2,1)$	0	0	0	0
	$(2,2)$	1.657	0	1.657	1.657

Table 5: Optimal solution of the dual for Mayers-Yao's correlations

4.3 Restricted inputs on Alice side for Mayers-Yao's self-test

In this section, Bob still has inputs $y \in \{0,1,2\}$ but Alice's inputs are restricted to $x \in \{1,2\}$, meaning that Alice only measures $A_0 = A_1 = Z$ and $A_2 = (X+Z)/\sqrt{2}$. Hence, we have less correlations to consider in the linear programs and the basis $\{p(a,b|x,y)\}$ will have a smaller dimension, but the primal and the dual keep the same form.

The primal gives $\alpha^*_{MY,\{1,2;0,1,2\}} \approx 0.1715 \approx \alpha^*_{MY}$ and the dual gives $\gamma^*_{MY,\{1,2;0,1,2\}} \approx 3.313$, $\omega^*_{MY,\{1,2;0,1,2\}} = 0$ and $\mathbf{y}^*_{\mathbf{MY},\{\mathbf{1,2;0,1,2}\}}$ in Table 6.

$\mathbf{y}^*_{\mathbf{MY},\{\mathbf{1,2;0,1,2}\}}$		(a,b)			
		$(-1,-1)$	$(-1,1)$	$(1,-1)$	$(1,1)$
(x,y)	$(1,0)$	0	1.657	1.657	1.657
	$(1,1)$	0	0	0	0
	$(1,2)$	1.657	0	0	0
	$(2,0)$	1.657	0	0	0
	$(2,1)$	0	0	0	0
	$(2,2)$	0	0	0	1.657

Table 6: Optimal solution of the dual for Mayers-Yao's correlations with the inputs on Alice's side restricted to $\{1,2\}$

The results are almost exactly the same as before, except that the strategy is modified: given an input (x,y), the outputs (a,b) are not the same as before. However, we still find the same objective value α^*_{MY} and the same Bell inequality.

In Table 5, one can notice that Bob does not use the input $y = 1$, i.e the measurement $Y_1 = Z$. However, restricting also the inputs on Bob's side to $y \in \{0, 2\}$ does not give exactly the same result. We obtain $\alpha^*_{MY,\{1,2;0,2\}} \approx 0.1715 \approx \alpha^*_{MY}$ with the primal and the dual gives $\gamma^*_{MY,\{1,2;0,2\}} \approx 2.485$, $\omega^*_{MY,\{1,2;0,2\}} = 0$ and $\mathbf{y}^*_{\mathbf{MY},\{\mathbf{1},\mathbf{2};\mathbf{0},\mathbf{2}\}}$ in Table 7.

$\mathbf{y}^*_{\mathbf{MY},\{\mathbf{1},\mathbf{2};\mathbf{0},\mathbf{2}\}}$		(a, b)			
		$(-1, -1)$	$(-1, 1)$	$(1, -1)$	$(1, 1)$
(x, y)	$(1, 0)$	0	0.828	0.828	0
	$(1, 2)$	0.828	0	0	0.828
	$(2, 0)$	0.828	0	0	0.828
	$(2, 2)$	0.828	0	0	0.828

Table 7: Optimal solution of the dual for Mayers-Yao's correlations with the inputs on Alice's side restricted to $x \in \{1, 2\}$ and on Bob's side to $y \in \{0, 2\}$.

Therefore, even though the optimal strategy is changed, the correlations considered are still nonlocal for the restricted inputs $x \in \{1, 2\}$,$y \in \{0, 2\}$.

4.4 Alternative dual formulation

From the numerical results, we observed that the first constraint of the Primal 22 is always fulfilled and substituting the inequality by an equality yields the same results. However, this modified primal form provides an alternative dual for detecting non-locality

$$\max_{\gamma,\mathbf{w}} \; -\mathcal{P}\cdot\mathbf{w} + \gamma - \omega$$
$$s.t. \begin{cases} (\mathbb{1} - \mathcal{P})\cdot\mathbf{w} - \omega & \leq 1 \\ \gamma + \mathbf{d}_\lambda \cdot \mathbf{w} & \leq 0, \ \forall\lambda \\ \mathbf{w} \in \mathbb{R}^n, \ \gamma \in \mathbb{R}, \omega & \geq 0 \end{cases} \tag{32}$$

The results from CHSH and Mayers-Yao exhibit that one always has at the optimum

$$\mathcal{P}\cdot\mathbf{w}^* = -\alpha^*,$$
$$\mathbb{1}\cdot\mathbf{w}^* = 1 - \alpha^*$$

and thus, multiplying the first constraint of the primal by $\mathbf{w}^*$, one finds that

$$\sum_\lambda \mu^*_\lambda \mathbf{d}_\lambda \cdot \mathbf{w}^* = 0 \ , \tag{33}$$

which means that the solution $\mathbf{w}^*$ of the dual problem is orthogonal to the optimal convex sum of deterministic behaviours.

According to the Theorem A.3 of *strong duality*, the optimal objective value is the same as the one from the primal and the following inequality is derived

$$-\mathbf{w}^*\cdot\mathcal{P} \leq \alpha^* + \omega^* - \gamma^*$$

The main difference with Equation 26 is that the vector $\mathbf{w}^*$ belongs to $\mathbb{R}^n$.

For CHSH, we find $-\mathbf{w}^*_{\mathbf{CHSH}}\cdot\mathcal{P} \leq \alpha^*_{CHSH}$ and for Mayers-Yao's, $-\mathbf{w}^*_{\mathbf{MY}}\cdot\mathcal{P} \leq \alpha^*_{MY}$.

5 Robustness against white noise

The aim of this section is to study robustness against the addition of white noise in a communication channel. The noise of the quantum channel is denoted by α, and thus we define the accuracy of the channel by $\mathcal{V} = 1 - \alpha$, e.g. and accuracy of $\mathcal{V} = 1$ corresponds to a lossless quantum channel. In experiments using a singlet state generated by two entangled photons, the noise can originate from imperfect operations on the photon source and detectors. Hence generating a pure singlet state is very difficult in practice.

This noise effect can be modeled by the following mixed state, called a Werner state,

$$\rho = \mathcal{V}\left|\Phi\right\rangle\left\langle\Phi\right| + (1 - \mathcal{V})\mathbb{1} \ , \tag{34}$$

where $|\Phi\rangle$ is a maximally entangled state and $\mathbb{1}$ a fully randomized behaviour.

The aim is to check for which amount of noise this state produces nonlocal behaviour in CHSH and Mayers-Yao scenario. The study of the robustness against white noise can be done by two different manners :

1. Start with $\mathcal{V} = 1$ and a given precision δ. While the result of the primal is an $\alpha > 0$, i.e., the behaviour $\mathcal{P}$ generated by ρ and ideal measurements is non local, set $\mathcal{V} = \mathcal{V} - \delta$. This simple algorithm is useful to draw a curve as shown in Figure 4, but requires to solve $O(\delta^{-1})$ nonlinear problems, which is not time-wise negligible.

2. Knowing that the visibility $\mathcal{V}$ is in the range $[0, 1]$, and that there exists a minimal value $\mathcal{V}^*$ such that for all $\mathcal{V} > \mathcal{V}^*$, the behaviour $\mathcal{P}$ generated with state ρ (Equation 34) and ideal measurements is not local, one can use a binary search to find $\mathcal{V}^*$. This allows to proceed with a logarithmic number of iterations and an arbitrary large precision e.g., a twenty-digit accuracy with only 55 iterations.

The implementation of the first method yielded the curves represented in Figure 4.

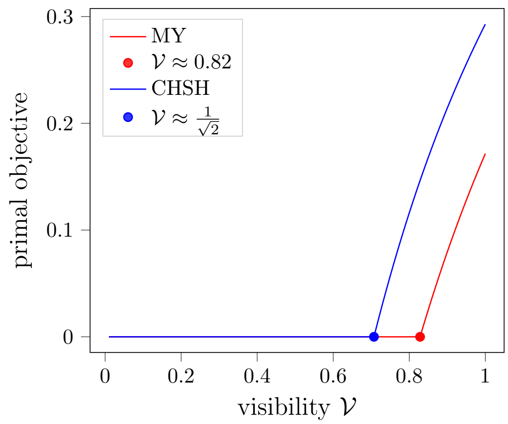

Figure 4: Optimal value of the linear program, depending on the set of correlations.

Concerning the CHSH's correlations, we obtained

$$\mathcal{V}^*_{CHSH} \approx 0.71 \approx \frac{1}{\sqrt{2}} ,$$

while comparatively, Mayer-Yao's correlations yield

$$\mathcal{V}^*_{MY} \approx 0.82 .$$

As a conclusion, the maximally entangled state is more robust to noise for CHSH correlations than for Mayers-Yao's correlations. The value $\mathcal{V}^*$ can be interpreted as the visibility lower bound, meaning that if the amount of noise is lower than $1-\mathcal{V}^*$, the violation of a Bell inequality remains detectable in an experiment.

Besides, experiments reproducing the maximum violation of a Bell inequality for the CHSH game have been realized without the need of calibrated devices [Sha+12], and the experimental visibility they observed matches with our result.

In each case, one can notice that $\mathcal{V}^* = 1 - \alpha^*$ where α^* was the optimal objective obtained in Section 3. The study of the robustness is equivalent to the primal detecting non-local behaviour since it was written using a convex combination of the behaviour and a fully randomized behaviour, which is actually a white noise. Therefore, the optimal value α^* of the primal corresponds to the maximal amount of noise tolerated for violating the Bell inequality. Still, this study would be necessary whenever one wants to study different kinds of noise or losses arising from experimental setups: specific noises can be modeled as specific local behaviours, instead of only white noise [Sha+12].

6 Non-linear programming

Considering Linear Program 22, one can transform the vector $\mathcal{P}$ into a vector of variables. This converts the initial program into a nonlinear program, more specifically a quadratic program, and provides as a result a behavior $\mathcal{P}$. This kind of transformation adds new difficulties for the resolution, such as non-convex constraints, but all of these are left over to the solver.

6.1 Finding the optimal probability distribution

Adding the constraints for $\mathcal{P}$ to be a probability vector yields the following nonlinear program

$$\max \ \beta_{CHSH}$$
$$s.t. \begin{cases} (1-q)\mathcal{P} + q\mathbb{1} & = \sum_\lambda \mu_\lambda \mathbf{d}_\lambda \\ \sum_\lambda \mu_\lambda & = 1 \\ \sum_{a,b} p(a,b|x,y) & = 1, \ \forall (x,y) \in \{0,1\}^2 \\ q & \leq 1 \\ \mu_\lambda \geq 0 \ \forall \lambda, \ q \geq 0 & \end{cases} \tag{35}$$

that one could solve to find the maximal value for β_{CHSH}, whose value is the given by the left-hand side of Equation 19. However, it must be remembered that the maximal value one can obtain doing this is $\beta_{CHSH} = 4$, where the resulting probability distribution describes a behavior belonging to $\mathcal{NS} \setminus \mathcal{Q}$.

Finding the boundaries between the sets $\mathcal{Q}$ and $\mathcal{NS}$ i.e., inequalities violated by behaviors in $\mathcal{NS} \setminus \mathcal{Q}$, is an open problem when considering arbitrary values of Δ and m. For bipartite scenarios with binary outcomes ($\Delta = 2$, $m = 2$), Equation 36

$$|\text{asin}\langle A_0B_0\rangle + \text{asin}\langle A_0B_1\rangle + \text{asin}\langle A_1B_0\rangle - \text{asin}\langle A_1B_1\rangle| \leq \pi, \tag{36}$$

introduced by Masanes in [Mas03], provides an explicit description of behaviours belonging to $\mathcal{Q}$. Moreover, the two sides are equal whenever the behavior maximally violates the CHSH Inequality 19.

While simple to understand, this constraint cannot be enforced by Gurobi, that cannot handle constraints described by trigonometric functions. Nonetheless, it will only be used to check whether a behavior stays in $\mathcal{Q}$.

Taking into consideration both Quadratic Program 35 and Equation 36 yields the following procedure to find the maximal value of β_{CHSH}

Algorithm 1 Numerical upper bound for CHSH inequality violation

Require: δ is the precision, roughly corresponds to the number of number of significant digits
Require: $\Pi = 2$ is the starting upper bound
while $|\sum_{x,y}(-1)^{xy}\text{asin}\langle A_xB_y\rangle| \leq \pi$ **do**
  $\Pi \leftarrow \Pi + \delta$
  solve linear program 35 with constraint $\beta_{CHSH} \leq \Pi$
return $\mathcal{P}$*: the optimal probability distribution*

The numerical result is, indeed, arbitrary close to $2\sqrt{2}$ depending on δ. One should note that not such equation as Equation 36 is known for the case $\Delta = 2$ and $m = 3$, but it is possible to use the generalized CHSH inequality as depicted by Wehner in [Weh06]. We could have use this for the case $\Delta = m = 2$ even though it will yield the same result, but we thought it was more interesting to try at least once another formulation. Considering the case

$\Delta = 2$ and $m = 3$, we indeed obtained results arbitrary close to $3\sqrt{3}$, and the associated probability distribution.

6.2 Making a behavior local

Looking for the optimal probability distribution with a non-linear approach led us to a side work, that is, from knowing how many elements of the behavior one can infer that the behavior is local. Taking into consideration that a behavior with $\Delta = m = 2$ is local if and only if $\beta_{CHSH} \leq 2$, it suffices to set k probabilities of, for instance, the behavior that maximally violates the CHSH inequality, and solving Nonlinear Program 35, where the maximization becomes a minimization. By doing this, if the optimal value is a $\beta^*_{CHSH} \leq 2$, then it is possible to find a full behavior that is local, and its elements are given in the same time by the solver. Otherwise, one can be sure that it is not possible to make the behavior local.

The results of the implementation with Gurobi yielded on one hand that with $k = 4$, where for each (x, y), a single $p(a, b|x, y)$ is fixed, it is not anymore possible to find values for the other probabilities such that the behavior is local, in addition, it only results on the behavior that gives $\beta_{CHSH} = 2\sqrt{2}$, since it removes all degrees of freedom for the other probabilities in the light of the normalization constraints (Equation 1) and (Equation 2).

Moreover, fixing only one expectation $\langle A_x B_y \rangle$ gives an optimal of 0 with $\beta_{CHSH} = 0$, while fixing an arbitrary number of probabilities among up to 3 expectations yields a local behavior, with an optimal 0, yet $\beta_{CHSH} \approx 1.1213$.

Besides, thanks to the implementation of Nonlinear Program 35, we were able to generate a behavior $\mathcal{P}_\Delta$ whose CHSH value is $\beta_{CHSH} = 2 + \Delta$, hence that violates the inequality with an arbitrary small precision Δ. It yielded for the both aforementioned cases exactly the same results.

For the case of the Mayers-Yao's correlations, the same result is obtained when fixing one probability per expectation i.e., the resulting probability distribution stays unchanged. One the other hand, fixing 3 arbitrary expectation also yields the same optimal objective as given in Equation 24, enforcing the behavior to be in $\mathcal{Q}$, while fixing only 2 arbitrary expectations causes the behavior to be local.

References

[Aar+17] Scott Aaronson et al. "Complexity-theoretic limitations on blind delegated quantum computation". In: (2017). DOI: `10.48550/ARXIV.1704.08482`.

[AGR82] Alain Aspect, Philippe Grangier, and Gérard Roger. "Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities". In: *Phys. Rev. Lett.* 49 (2 1982), pp. 91–94. DOI: `10.1103/PhysRevLett.49.91`.

[Bel64] J. S. Bell. "On the Einstein Podolsky Rosen paradox". en. In: *Physics Physique Fizika* 1.3 (Nov. 1964), pp. 195–200. DOI: `10.1103/PhysicsPhysiqueFizika.1.195`.

[Bru+14] Nicolas Brunner et al. "Bell nonlocality". In: *Reviews of Modern Physics* 86.2 (2014), pp. 419–478. DOI: `10.1103/revmodphys.86.419`.

[Cir80] B. S. Cirel'son. "Quantum generalizations of Bell's inequality". en. In: *Letters in Mathematical Physics* 4.2 (Mar. 1980), pp. 93–100. DOI: `10.1007/BF00417500`.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" en. In: *Physical Review* 47.10 (May 1935), pp. 777–780. DOI: `10.1103/PhysRev.47.777`.

[Goh+18] Koon Tong Goh et al. "Geometry of the set of quantum correlations". In: *Phys. Rev. A* 97 (2 2018), p. 022104. DOI: `10.1103/PhysRevA.97.022104`.

[Gur22] Gurobi Optimization, LLC. *Gurobi Optimizer Reference Manual*. 2022.

[Hoc07] Dorit S. Hochbaum. "Complexity and algorithms for nonlinear optimization problems". en. In: *Annals of Operations Research* 153.1 (June 2007), pp. 257–296. DOI: `10.1007/s10479-007-0172-6`.

[Kar84] N. Karmarkar. "A new polynomial-time algorithm for linear programming". en. In: *Combinatorica* 4.4 (Dec. 1984), pp. 373–395. DOI: `10.1007/BF02579150`.

[Liu+18] Yang Liu et al. "Device-independent quantum random-number generation". en. In: *Nature* 562.7728 (Oct. 2018), pp. 548–551. DOI: `10.1038/s41586-018-0559-3`.

[Mas03] Lluis Masanes. "Necessary and sufficient condition for quantum-generated correlations". In: *arXiv: Quantum Physics* (2003).

[MPA11] Lluís Masanes, Stefano Pironio, and Antonio Acín. "Secure device-independent quantum key distribution with causally independent measurement devices". en. In: *Nature Communications* 2.1 (Sept. 2011), p. 238. DOI: `10.1038/ncomms1244`.

[NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge: Cambridge University Press, 2010. ISBN: 9780511976667. DOI: `10.1017/CBO9780511976667`.

[PR94] Sandu Popescu and Daniel Rohrlich. “Quantum nonlocality as an axiom”. en. In: *Foundations of Physics* 24.3 (Mar. 1994), pp. 379–385. DOI: 10.1007/BF02058098.

[SB20] Ivan Supic and Joseph Bowles. “Self-testing of quantum systems: a review”. In: *Quantum* 4 (2020), p. 337. DOI: 10.22331/q-2020-09-30-337.

[Sha+12] Peter Shadbolt et al. “Guaranteed violation of a Bell inequality without aligned reference frames or calibrated devices”. en. In: *Scientific Reports* 2.1 (Dec. 2012), p. 470. DOI: 10.1038/srep00470.

[Weh06] Stephanie Wehner. “Tsirelson bounds for generalized Clauser-Horne-Shimony-Holt inequalities”. en. In: *Physical Review A* 73.2 (Feb. 2006), p. 022110. DOI: 10.1103/PhysRevA.73.022110.

A Mathematical optimization

Mathematical optimization aims to maximize (or minimize) the value of a function[1], under specific *constraints*. The simplest instance of mathematical optimization is *linear programming*, where both the objective and the constraints are *linear functions* of the variables.

Both linear optimization and nonlinear optimization will be examined.

For all the numerical simulations and the corresponding implementations, Gurobi [Gur22] (a solver that can handle both linear and quadratic optimization problems) was used.

A.1 Linear optimization

The field of linear optimization grew in tandem with the field of Computer Science, and the *"programmatic"* approach to linear optimization became very popular. As such, it will be explained via *linear programming* conventions and notations.

A *linear program* is written as

$$\begin{aligned} &\max \ \textit{or} \ \min z = c^T x \\ &s.t. \begin{cases} Ax & \leq b \\ x & \geq 0 \end{cases}, \end{aligned} \tag{37}$$

where $c^t = (c_1, \cdots, c_n)$ corresponds to the *objective function coefficients*, $x^t = (x_1, \cdots, x_n)$ to the *decision variables*, $b^t = (b_1, \cdots, b_m)$ to the *value of the constraints* and

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \dots & \ddots & \dots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

to the *constraint coefficients*. Here, z corresponds to the *objective function.*

The choice of inequality ($\geq$ or $\leq$) is arbitrary, since one can always multiply both sides by a negative coefficient and obtain the other. To obtain an equality, one can use two separate inequalities.

The problem formulated in Equation 37 is generally referred to as the *primal* problem, and its *dual* form is

$$\begin{aligned} &\min \ \textit{or} \ \max z' = v^T b \\ &s.t. \begin{cases} A^T v & \geq c \\ v & \geq 0 \end{cases}. \end{aligned} \tag{38}$$

In Section 4, the dual form is very important. To switch from one form to the other, Theorem A.3 is important since it states that the optimum of either objective form is the same. The following definitions are preliminaries for its introduction.

Definition A.1 (Feasible solution)**.** $x \in \mathbb{R}^n$ is a feasible solution as long as it satisfies every constraints of the problem.

Definition A.2 (Optimal solution)**.** $x \in \mathbb{R}^n$ is an optimal solution whenever it gives to z its optimal value.

Theorem A.3 (Strong duality)**.** *If there exists a feasible solution for the primal, then the primal optimal objective and the dual optimal objective are equal.*

Lastly, an important thing to keep in mind is that in linear optimization, the optimal solution is often an extremal point of a polytope belonging to $\mathbb{R}^n$, whose facets are described by the constraints.

A.2 Nonlinear optimization

Nonlinear optimization, unlike linear optimization, handles with problems where the constraints are nonlinear functions of the variables, such as exponentials, trigonometric, or polynomial functions. In Section 6, the formulated problems only deals with quadratic function, which are quite easy problem to solve.

While linear optimization can be solved in polynomial time [Kar84], nonlinear optimization is, except for some cases, an NP-Hard problem [Hoc07]. In addition, there are fundamental issues with nonlinear optimization inherent in digital computer e.g., floating point arithmetic precision. Consequently, the techniques for finding the optimal solutions are not exposed here.

[1] the *objective function*