UNIVERSITY OF PADOVA

Physics and Astronomy Department "Galileo Galilei"

Master's Degree in Physics

Master thesis

Sequential quantum measurements

Supervisor

Giuseppe Vallone

Co-supervisor

Lorenzo Coccia, Matteo Padovan

Candidate

Alessandro Rezzi

Academic Year 2023/2024

Contents

Chapter 1

Introduction

In the fast evolving world of quantum information, several key ideas are reshaping how we understand and utilize quantum systems. One of the most fascinating is quantum non-locality, the concept that the predictions of quantum theory fundamentally challenge the classical notion of local realism. Roughly speaking, quantum non-locality reveals that the behavior of entangled particles cannot be explained by any theory that assumes objects are only influenced by their immediate surroundings (a more precise definition will be given in chapter 2). Two main contributions that developed this idea are the EPR (Einstein–Podolsky–Rosen) paradox [7] and the work of J.S. Bell [1], both of which forced us to rethink the very nature of reality and had profound implications for the development of quantum technologies.

A fundamental framework for exploring quantum non-locality is the Bell's scenario. It involves a physical system, prepared in an entangled state, that is shared and measured by two spatially separated users, that have access to a set of quantum observables. Since the measurement process is intrinsically probabilistic, their outcomes can be used to generate random numbers. This task is crucial in many fields: for instance, in cryptography, the security of encryption keys and protocols relies on generating truly random values, ensuring that keys are unpredictable and resistant to attacks. Another example is in statistics, where Monte Carlo methods depend on the quality of the random number generator, as their accuracy is tied to the randomness of the samples. Furthermore, quantum mechanics is the only known way to generate true randomness: generators based on classical mechanics are deterministic and thus output pseudo-random numbers, which are, in principle, predictable.

The challenging part of generating random number with Bell's scenarios is making sure that everything is working as expected, for example it has to be verified that the system is actually entangled and the functionality of the complex measurement apparatus. To address these problems, the concept of self-testing has become an invaluable tool: it allows the verification of quantum devices without requiring knowledge of their internal workings. By analyzing the outcomes of a Bell's scenario, self-testing can confirm that all devices are performing the expected quantum operations, even if their internal mechanisms are not known or trusted. Hence, outcomes are used both to generate random bits and to ensure the integrity of the system.

A major limitation of the proposed scenario is the low extraction rate: since there are only two users, we can generate at most two numbers from each entangled state, but creating and preserving entanglement is particularly challenging. A natural solution is to sequentially increase the number of users, in the sense that each new pair will measure on the post-measurement state of the previous one. This works provided that each post-measurement state is still entangled, which is achieved by using generalized operators in place of projective ones.

In this thesis we will explore those topics explaining in more details what is non-locality, self-testing (chapters 2 and 3) and their applications to generate secure random numbers (chapter 5). In chapter 4 we will cover the NPA (Navascués-Pironio-Acín) hierarchy, a fundamental tool to perform numerical simulations and characterize non-locality. In chapter 6, we will talk about sequential Bell's scenario and generalized operators. Finally, we will propose an innovative approach to extend a large family of Bell's scenario to the sequential case with three users. Those results will be theoretically proved in chapter 7, and validated with numerical simulations techniques in chapter 8.

Chapter 2

Non-locality

2.1 Space of behaviors

Consider two spatially separated observers, Alice and Bob, that perform measurements on a shared physical system, generated by a source S. Each observer can choose from a set of m observables to measure their part of the system and each observable can yield Δ possible outcomes. In particular we label the inputs of Alice an Bob as $x, y \in \{1, 2, ..., m\}$ and their outputs as $a, b \in \{1, 2, ..., \Delta\}$ respectively, see figure 2.1. After measuring many pairs with different observables they can build the joint probability distribution

$$p(a, b|x, y),$$

which is the probability of having outcomes a and b from the observables x and y. The set with all probabilities

$$\mathbf{p} = \{p(a, b|x, y);\ a, b = 1, ..., \Delta;\ x, y = 1, ..., m\}$$

is called a behavior, and can be seen as a point of $\mathbb{R}^{\Delta^2 m^2}$ that satisfies the following constraints:

$$\begin{aligned} &p(a, b|x, y) \geq 0, \quad \text{positivity constraint;} \\ &\sum_{a,b=1}^{\Delta} p(a, b|x, y) = 1, \quad \text{normalization constraint.} \end{aligned}$$

Then, based on the actual physical model behind this experiment, we define different classes of joint probability distributions:

No-signaling behaviors A behavior is no-signaling if it satisfies the additional constraints

$$\begin{aligned} \sum_{a=1}^{\Delta} p(a, b|x, y) &= \sum_{a=1}^{\Delta} p(a, b|x', y), \quad \forall x, y, x', b; \\ \sum_{b=1}^{\Delta} p(a, b|x, y) &= \sum_{b=1}^{\Delta} p(a, b|x, y'), \quad \forall x, y, y', a. \end{aligned} \tag{2.1}$$

This is equivalent to requiring that the marginal probability of Alice

$$p(a|x) \equiv p(a|x, y) \equiv \sum_{b=1}^{\Delta} p(a, b|x, y)$$

doesn't depend on the choice of the Bob's observable y, and analogous property holds for Bob's marginal probability. Note that the assumption of Alice and Bob being spatially separated ensures the no-signaling condition: indeed if this was not the case, Bob's choice of the input would have a faster-than-light consequence on the probability distribution measured by Alice, thereby contradicting special relativity. Finally the subset of $\mathbb{R}^{\Delta^2 m^2}$ made by all no-signaling behaviors is called $\mathcal{NS}$.

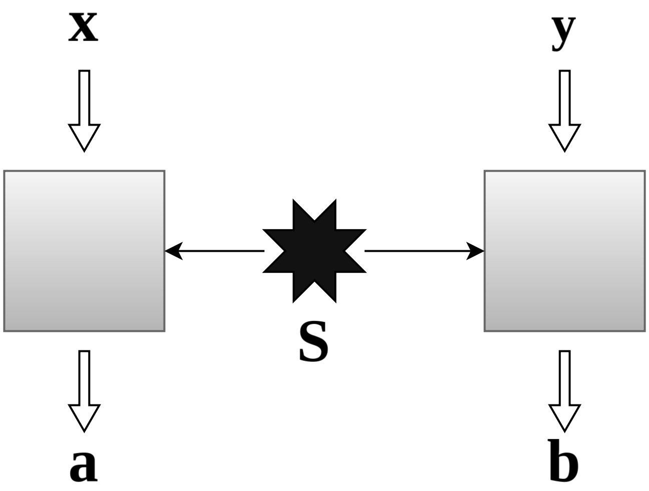

Figure 2.1: Representation of the experiment: A shared physical system is generated by a source S, and measured by Alice and Bob. They respectively choose an observable labeled by x, y, that outputs two numbers a,b.

Quantum behaviors A behavior is quantum if it can be written as:

$$p(a,b|x,y) = \mathrm{Tr}[\rho_{AB}\Lambda_a^x \otimes \Pi_b^y], \tag{2.2}$$

where ρ_{AB} is a quantum state belonging to a generic Hilbert state $\mathcal{H} \equiv \mathcal{H}_A \otimes \mathcal{H}_B$ and $\Lambda_a^x : \mathcal{H}_A \to \mathcal{H}_A$, $\Pi_b^y : \mathcal{H}_B \to \mathcal{H}_B$ are generic POVM operators, thus satisfying

$$\begin{aligned} &\Lambda_a^x \geq 0, \quad \Pi_b^y \geq 0; \\ &\sum_{a=1}^{\Delta} \Lambda_a^x = \mathbb{1}, \quad \sum_{b=1}^{\Delta} \Pi_b^x = \mathbb{1}. \end{aligned}$$

The subset of $\mathbb{R}^{\Delta^2 m^2}$ made by all quantum behaviors is called $\mathcal{Q}'$. There is also an alternative definition in which instead of using the tensor product of local operators Λ_a^x and Π_b^y we consider behaviors of the form

$$p(a,b|x,y) = \mathrm{Tr}[\rho_{AB}\widetilde{\Lambda}_a^x \widetilde{\Pi}_b^y], \tag{2.3}$$

where ρ_{AB} is again a quantum state belonging to a generic Hilbert state $\mathcal{H}$ and $\widetilde{\Lambda}_a^x : \mathcal{H} \to \mathcal{H}$, $\widetilde{\Pi}_a^x : \mathcal{H} \to \mathcal{H}$ are commuting POVM operators, thus satisfying the additional condition

$$[\widetilde{\Lambda}_a^x, \widetilde{\Pi}_b^y] = 0.$$

The subset of $\mathbb{R}^{\Delta^2 m^2}$ made by those behavior is called $\mathcal{Q}$. Clearly from the definition it follows that

$$\mathcal{Q}' \subseteq \mathcal{Q},$$

however is still an open question whether the two sets are equal. In the special case of finite dimensional Hilbert spaces, they turn out to be identical as proven in [8]. Practically we will use the subset $\mathcal{Q}'$ for proving theoretical result and $\mathcal{Q}$ for numerical simulations.

Local behaviors A behavior is local if it can be written as:

$$p(a,b|x,y) = \int d\lambda q(\lambda) p(a,b|x,y,\lambda) = \int d\lambda q(\lambda) p(a|x,\lambda) p(b|y,\lambda), \tag{2.4}$$

where λ is a random variable, with distribution $q(\lambda)$, that can be thought as a hidden variable that correlates the two parts of the shared system. Consequently, the outcome of a measurement now depends both on the observable and the value of λ. The idea behind this model is that, after marginalizing, there could be correlations between the outcomes of Alice and Bob, so in general

$$p(a,b|x,y) \neq p(a|x)p(b,y),$$

but those correlations depends only on the hidden variable. The measurement of Alice doesn't influence Bob and vice versa, indeed we assumed that for fixed x, y and λ the outcomes are completely uncorrelated

$$p(a,b|x,y,\lambda) = p(a|x,\lambda)p(b|y,\lambda).$$

The subset of $\mathbb{R}^{\Delta^2 m^2}$ made by all local behaviors is called $\mathcal{L}$. The sets $\mathcal{NS}$, $\mathcal{Q}$ and $\mathcal{L}$ are closed, bounded, convex [5] [11] and satisfy

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

see figure 2.2 for a pictionary representations of the three sets.

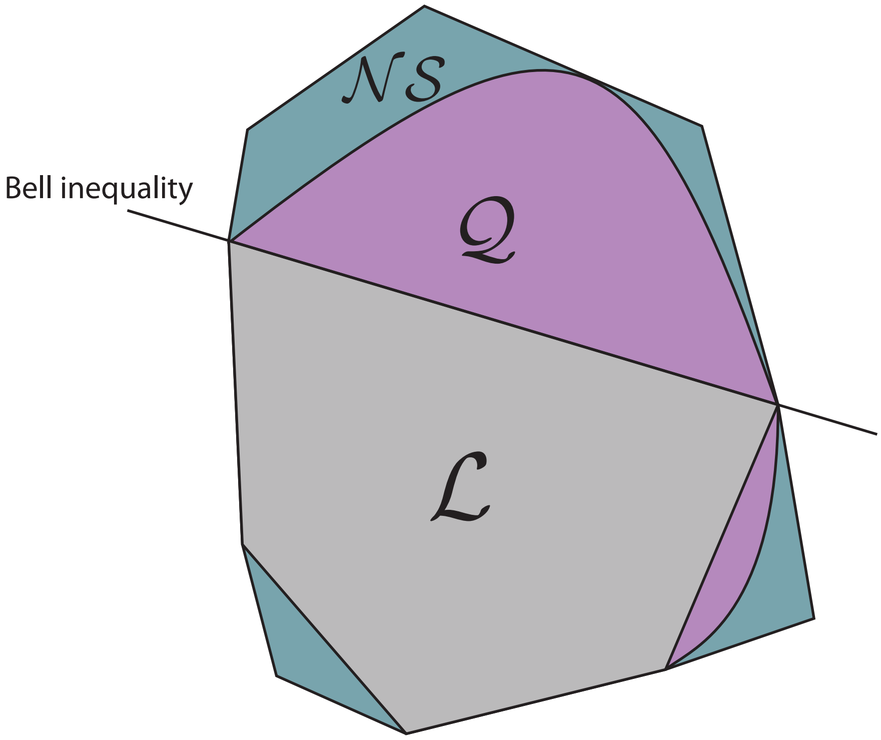

Figure 2.2: Sketch of the $\mathcal{L}, \mathcal{Q},$ and $\mathcal{NS}$ sets. Image taken from [5].

2.2 Bell's inequalities

Let $\mathbf{q} \in \mathbb{R}^{\Delta^2 m^2}$ be a behavior that doesn't belong to one of the sets $\mathcal{K} = \mathcal{NS}, \mathcal{Q}, \mathcal{L}$. Since those sets are convex, we can apply the hyperplane separation theorem that guarantees the existence of an hyperplane that separates $\mathbf{q}$ from $\mathcal{K}$. So there must exist an inequality of the form

$$\mathbf{s} \cdot \mathbf{p} \equiv \sum_{a,b,x,y} s^{a,b}_{x,y} p(a,b|x,y) \leq S_k, \qquad \mathbf{s} \in \mathbb{R}^{\Delta^2 m^2},$$

that is satisfied by all $\mathbf{p} \in \mathcal{K}$, but is violated by $\mathbf{q}$:

$$\mathbf{s} \cdot \mathbf{q} > S_k,$$

for $\mathcal{K} = \mathcal{L}$ those inequalities are called Bell's inequalities. Instead, for $\mathcal{K} = \mathcal{Q}$, they are referred to as quantum Bell's inequalities or Tsirelson inequalities.

Characterization of local and quantum set As proved in [5], the local set is a polytope, that is the convex hull of a finite numbers of points, and can be characterized by a finite number of Bell's inequality. The quantum set $\mathcal{Q}$ instead, is not a polytope and characterizing it is more difficult. A possible way is using the NPA hierarchy, a numerical algorithm that we will describe in chapter 4.

CHSH inequality As an example consider the particular case in which $m = \Delta = 2$ so that

$$x, y \in \{0, 1\}; \qquad a, b \in \{-1, +1\}.$$

Note that outcomes belongs to the set $\{-1, 1\}$ rather than $\{1, 2\}$, it is a convention that we will be using when $\Delta = 2$. Consider the following expression

$$S = \langle a_0 b_0 \rangle + \langle a_0 b_1 \rangle + \langle a_1 b_0 \rangle - \langle a_1 b_1 \rangle, \tag{2.5}$$

where the expectation values $\langle ab \rangle$ are defined as

$$\langle a_x b_y \rangle \equiv \sum_{a,b=\pm 1} abp(a, b|x, y).$$

If we constraint the theory to be local, then by applying equation (2.4) we derive that:

$$\langle a_x b_y \rangle = \int d\lambda q(\lambda) \langle a_x \rangle_\lambda \langle b_y \rangle_\lambda$$

where we introduced

$$\langle a_x \rangle_\lambda \equiv \sum_{a=\pm 1} ap(a|x, \lambda), \qquad \langle b_y \rangle_\lambda \equiv \sum_{b=\pm 1} bp(b|y, \lambda)$$

and S can be rewritten as

$$\begin{aligned} S = & \int d\lambda q(\lambda)(\langle a_0 \rangle_\lambda \langle b_0 \rangle_\lambda + \langle a_0 \rangle_\lambda \langle b_1 \rangle_\lambda + \langle a_1 \rangle_\lambda \langle b_0 \rangle_\lambda - \langle a_1 \rangle_\lambda \langle b_1 \rangle_\lambda) \leq \\ \leq & \int d\lambda q(\lambda)(|\langle b_0 \rangle_\lambda + \langle b_1 \rangle_\lambda| + |\langle b_0 \rangle_\lambda - \langle b_1 \rangle_\lambda|) \leq \int d\lambda q(\lambda) 2 = 2, \end{aligned} \tag{2.6}$$

where we used that both $\langle a_x \rangle_\lambda$ and $\langle b_y \rangle_\lambda$ take values in $[-1, 1]$. Equation (2.6) is an example of Bell's inequality for the local set $\mathcal{L}$ and is known as CHSH inequality. It's also easy to see that there is a quantum strategy that violates it: consider the 2-qubits pure state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \equiv |\phi_+\rangle$$

and let Alice and Bob measure the two observables

$$A_0 = \sigma_x, \quad A_1 = \sigma_z, \quad B_0 = \frac{\sigma_x + \sigma_z}{\sqrt{2}}, \quad B_1 = \frac{\sigma_x - \sigma_z}{\sqrt{2}}, \tag{2.7}$$

in such a way that the expectation values $\langle ab \rangle$ can be computed as

$$\langle a_x b_y \rangle = \langle \psi | A_x \otimes B_y | \psi \rangle$$

and a simple calculation shows that

$$\langle a_0 b_0 \rangle = \langle a_1 b_0 \rangle = \langle a_0 b_1 \rangle = \frac{1}{\sqrt{2}}, \langle a_1 b_1 \rangle = -\frac{1}{\sqrt{2}} \implies S = 2\sqrt{2} > 2.$$

Hence violating the bound (2.6). As we will see later $S_q = 2\sqrt{2}$ is also the maximal violation of the CHSH inequality achievable on the quantum set.

Chapter 3

Self-testing

So far we have treated our system as a black box: we didn't specify what is the actual physical system being measured (for example it could be photons, electrons, atoms...) and we didn't specify the, usually very complex, experimental apparatus used to measure, indeed we only focused on the final joint probability distribution $p(a, b|x, y)$. This limited knowledge of the system is usually called device-independent scenario. Then, from the sole knowledge of the joint probability distribution $p(a, b|x, y)$ we derived a sufficient condition to determine if the system is entangled (i.e. there is non-locality): we compute the CHSH value with equation (2.5) and check if the result is greater than 2. In general we can compare the probability distribution against all known Bell's inequalities and if at least one is violated we can conclude that the system is entangled. Such procedure is called device-independent certification of entanglement and is an example of self-testing. In this chapter we will formalize this concept and show that from the maximal violation of a Bell's inequality we might even certify the quantum state and the measurements done by Alice and Bob.

Self-testing scenario Let us begin by formally defining the device-independent scenario to perform self-testing. As already said, we want to consider an abstraction of the system that hides the internal mechanics of the measurement devices and the source. However, we regardless need to make some basic physical assumptions, that are:

- The experiment admits a quantum description: there exists a quantum state and measurement operators that lead to the observed outcomes;
- Alice and Bob are very distant and there is no communication between them. In particular Alice cannot communicate her choice of the observable to Bob and vice versa;
- Each round of the experiment is independent and physically equivalent to each other. With round we mean the choice of the observables and measurement process.

These conditions are necessary to write behaviors by using the quantum formalism

$$p(a, b|x, y) = \mathrm{Tr}[\rho_{AB}\Lambda_a^x \otimes \Pi_b^y],$$

and in particular imply that the set of operators Λ_a^x, Π_b^y and the density matrix ρ_{AB} are the same at every round.

3.1 Purified quantum behaviors

In the previous chapter we characterized the set of quantum behaviors, see equation (2.2), by using density matrices and generalized POVM operators. As proved in [6], without loss of generality we can always assume to live in a larger Hilbert space in which the state is pure and all operators are projective, more precisely:

- The Hilbert space becomes $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_P$, where $\mathcal{H}_A$ is the Hilbert space of Alice, $\mathcal{H}_B$ is the Hilbert space of Bob and $\mathcal{H}_P$ is the purification space.
- In this expanded space the wave function $|\psi\rangle_{ABP}$ is pure and we can recover the original one by tracing on the purification space:
$$\rho_{AB} = \text{Tr}_P[|\psi\rangle\langle\psi|];$$
- In this expanded space Alice and Bob observables are projective
$$\begin{aligned} &\Lambda_a^x \geq 0, \quad \Lambda_a^x \Lambda_{a'}^x = \delta_{a,a'}, \quad \Lambda_a^x : \mathcal{H}_A \to \mathcal{H}_A; \\ &\Pi_b^y \geq 0, \quad \Pi_b^y \Pi_{b'}^x = \delta_{b,b'}, \quad \Pi_b^x : \mathcal{H}_B \to \mathcal{H}_B. \end{aligned} \tag{3.1}$$

This procedure is called Stinespring dilation and it is a powerful theoretical tools, because working with projectors is much easier than considering generic POVM. Finally the joint probability distribution can be written as
$$p(a,b|x,y) = \langle\psi_{ABP}|\, \Lambda_a^x \otimes \Pi_b^y \otimes \mathbb{I}_P \,|\psi_{ABP}\rangle\,, \tag{3.2}$$
where $\mathbb{I}_P$ is the identity on the purification space.

3.2 Local isometries

As anticipated the aim of self-testing is to infer the state and the measurements by knowing the joint probability distribution. In other words given a behavior $p(a,b|x,y)$ we want to prove that there is a unique choice of state $|\psi\rangle$ and projectors Λ_a^x, Π_b^y that satisfy equation (3.2). However, notice that if we find a state and projectors compatible with $p(a,b|x,y)$ we can build infinitely many more:

- We can apply the local unitary transformation
$$|\psi\rangle \to U \otimes V\,|\psi\rangle\,, \quad \Lambda_a^x \to U\Lambda_a^x U^\dagger, \quad \Pi_b^y \to V\Pi_b^y V^\dagger,$$
where
$$V : \mathcal{H}_A \to \mathcal{H}_A, \quad UU^\dagger = U^\dagger U = \mathbb{I}_A, \quad V : \mathcal{H}_B \to \mathcal{H}_B, \quad VV^\dagger = V^\dagger V = \mathbb{I}_B,$$
and this transformation clearly leaves equation (3.2) unchanged
- We can expand the space and let the operators act trivially on the extra degrees of freedom
$$|\psi\rangle \to |\psi\rangle \otimes |\xi\rangle\,, \quad \Lambda_a^x \to \Lambda_a^x \otimes \mathbb{I}, \quad \Pi_b^y \to \Pi_b^y \otimes \mathbb{I}.$$
This transformation also leaves the probability distribution $p(a,b|x,y)$ unchanged.

So there is no way to infer unique state and measurements from a behavior $p(a,b|x,y)$, and we need to take in account those additional degrees of freedom. This is done by using a particular class of transformations: local isometries. In general an isometry
$$\Phi : \mathcal{H}_A \to \mathcal{H}_A \otimes \mathcal{H}_{A'}$$
is a linear transformation that preserve the inner product. The action of an isometry on a vector $|\psi\rangle_A$ can always be decomposed as the composition of a dilatation, in which the state is embedded in a larger Hilbert space, and a unitary transformation $U_{AA'}$
$$|\psi\rangle_A \xrightarrow{\text{dilatation}} |\psi\rangle_A \otimes |0\rangle_{A'} \xrightarrow{\text{unitary map}} U_{AA'}(|\psi\rangle_A \otimes |0\rangle_{A'}).$$
Then, a local isometry on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$
$$\Phi = \Phi_A \otimes \Phi_B : \mathcal{H}_A \otimes \mathcal{H}_B \to (\mathcal{H}_A \otimes \mathcal{H}_{A'}) \otimes (\mathcal{H}_B \otimes \mathcal{H}_{B'})$$
is a tensor product of isometries acting locally on $\mathcal{H}_A$ and $\mathcal{H}_B$. This is implemented by embedding the initial state in a larger Hilbert space $(\mathcal{H}_A \otimes \mathcal{H}_{A'}) \otimes (\mathcal{H}_B \otimes \mathcal{H}_{B'})$ and then performing a local unitary transformation on $(\mathcal{H}_A \otimes \mathcal{H}_{A'})$ and $(\mathcal{H}_B \otimes \mathcal{H}_{B'})$, see figure ??:
$$|\psi\rangle_{AB} \xrightarrow{\text{dilatation}} |\psi\rangle_{AB}\,|00\rangle_{A'B'} \xrightarrow{\text{unitary map}} U_{AA'} \otimes V_{BB'}(|\psi\rangle_{AB}\,|00\rangle_{A'B'}).$$
The generic state $|00\rangle_{A'B'}$ is called ancilla state.

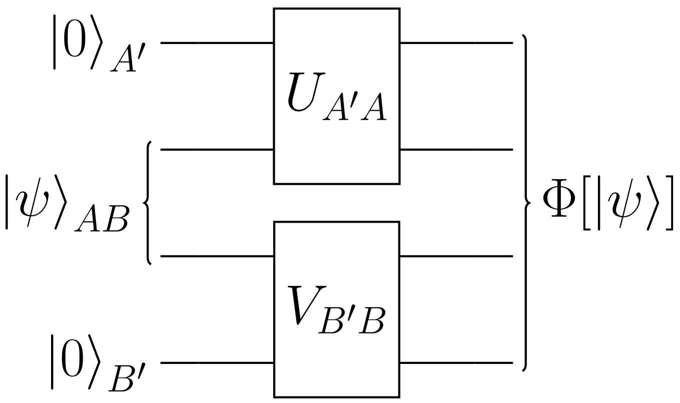

Figure 3.1: Representation of a local isometry

3.3 Self-testing of states

We are finally ready to formalize the concept of self-testing: the behavior $p(a,b|x,y)$ self-tests the state $|\psi'\rangle_{A'B'}$ if for any state ρ_{AB} compatible with $p(a,b|x,y)$ (for some choice of local measurements) and for any purification $|\psi\rangle_{ABP}$ of ρ_{AB} there exists a local isometry

$$\Phi_A \otimes \Phi_B : \mathcal{H}_A \otimes \mathcal{H}_B \to \mathcal{H}_{A'} \otimes \mathcal{H}_{\bar{A}} \otimes \mathcal{H}_{B'} \otimes \mathcal{H}_{\bar{B}}$$

such that

$$\Phi_A \otimes \Phi_B \otimes \mathbb{I}_P \left|\psi\right\rangle_{ABP} = \left|\psi'\right\rangle_{A'B'} \otimes \left|\xi\right\rangle_{\bar{A}\bar{B}P} \tag{3.3}$$

for some state $|\xi\rangle_{\bar{A}\bar{B}P}$ which is called junk state. This definition is consistent with what we said in the previous section: self-test is achieved when the infinite number of states, compatible with the behavior, are all connected by local isometries. Finally is also possible to get rid of the purification space by performing the partial trace on equation (3.3):

$$\Phi_A \otimes \Phi_B \,\mathrm{Tr}_P[|\psi\rangle\langle\psi|] = \Phi_A \otimes \Phi_B \rho_{AB} = |\psi'\rangle\langle\psi'| \otimes Tr_P[\left|\xi\right\rangle \left\langle\xi\right|_{\bar{A}\bar{B}P}],$$

and this shows how, with the local isometry, we can extract the state $|\psi'\rangle_{A'B'}$ from the density matrix ρ_{AB}.

3.4 Self-testing of states and measurements

We can extend the definition to cover also the certification of measurements. The behavior $p(a,b|x,y)$ self-test the state $|\psi'\rangle_{A'B'}$ and projective measurements $\widetilde{\Lambda}_a^x$, $\widetilde{\Pi}_b^y$ if for any state and projective measurements ρ_{AB}, Λ_a^x, Π_b^y compatible with $p(a,b|x,y)$ and for any purification $|\psi\rangle_{ABP}$ of ρ_{AB} there exist a local isometry

$$\Phi_A \otimes \Phi_B : \mathcal{H}_A \otimes \mathcal{H}_B \to \mathcal{H}_{A'} \otimes \mathcal{H}_{\bar{A}} \otimes \mathcal{H}_{B'} \otimes \mathcal{H}_{\bar{B}}$$

such that it satisfies equation (3.3) and

$$\Phi_A \otimes \Phi_B \otimes \mathbb{I}_P(\Lambda_a^x \otimes \Pi_b^y \otimes \mathbb{I}_P \left|\psi\right\rangle_{ABP}) = (\widetilde{\Lambda}_a^x \otimes \widetilde{\Pi}_b^y \left|\psi'\right\rangle_{A'B'}) \otimes \left|\xi\right\rangle_{\bar{A}\bar{B}P}$$

for all $x,y \in \{1,...m\}$, $a,b \in \{1,...,\Delta\}$ and for some junk state $|\xi\rangle_{ABP}$.

3.5 Complex conjugation

Let's say that we have a probability distribution $p(a,b|x,y)$ and a state and measurements $(\left|\psi\right\rangle; \Lambda_a^x; \Pi_b^y)$ compatible with it. Since $p(a,b|x,y)$ is real, equation (3.2) is invariant under the complex conjugation operator $*$ and therefore the state and measurements $(\left|\psi\right\rangle^*; (\Lambda_a^x)^*; (\Pi_b^y)^*)$ are also compatible with

$p(a, b|x, y)$. So we have to make sure that the $*$ operator is unitary, otherwise we would always have two sets of state and measurements not connected by an isometry, thus nullifying the definitions of self-testing. For the state $|\psi\rangle$ this is the case: indeed with a Schmidt decomposition we can always find two orthonormal sets

$$\{|u_1\rangle, ... |u_m\rangle\} \in \mathcal{H}_A, \qquad \{|v_1\rangle, ... |v_m\rangle\} \in \mathcal{H}_B$$

such that

$$|\psi\rangle = \sum_{j=1}^{m} \alpha_j |u_j\rangle \otimes |v_j\rangle, \qquad \alpha_j = \rho_j e^{i\theta_j} \in \mathbb{C},$$

where m is the so called Schmidt rank, which is smaller or equal than the smallest Hilbert space dimension:

$$m \leq \min(\dim(\mathcal{H}_A), \dim(\mathcal{H}_B)).$$

The action of complex conjugation operator is

$$|\psi\rangle \to |\psi\rangle^* = \sum_{j=1}^{m} \rho e^{-i\theta_j} |u_j\rangle \otimes |v_j\rangle$$

but this can always achieved with a unitary transformation, such as

$$|u_j\rangle \to e^{-i\theta_j} |u_j\rangle, \quad |v_j\rangle \to e^{-i\theta_j} |v_j\rangle.$$

Therefore $|\psi\rangle$ and $|\psi\rangle^*$ are always connected by a unitary transformation, and definition (3.3) remains well defined. Unfortunately the same result doesn't hold for observables: in general an operator $O \in \mathcal{H}$ is not unitary connected to O^*, and therefore the definition given in (3.4) becomes useless. It is however still applicable to some special cases, for example when all observables are real. In the most general case in which operators can be complex we need to find a new definition for self-testing that takes in account the additional degree of freedom given by the complex conjugation.

3.6 Self-testing of states and complex measurements

The generalization of self-testing to complex operators is the following: the behavior $p(a, b|x, y)$ self-test the state $|\psi'\rangle_{A'B'}$ and (complex-valued) projective measurements $\widetilde{\Lambda}^x_a$, $\widetilde{\Pi}^y_b$ if for any state and projective measurements ρ_{AB}, Λ^x_a, Π^y_b compatible with $p(a, b|x, y)$ and for any purification $|\psi\rangle_{ABP}$ of ρ_{AB} there exist a local isometry

$$\Phi_A \otimes \Phi_B : \mathcal{H}_A \otimes \mathcal{H}_B \to \mathcal{H}_{A'} \otimes \mathcal{H}_{\bar{A}} \otimes \mathcal{H}_{B'} \otimes \mathcal{H}_{\bar{B}}$$

such that it satisfies equation (3.3) and

$$\Phi_A \otimes \Phi_B \otimes \mathbb{I}_P(\Lambda^x_a \otimes \Pi^y_b \otimes \mathbb{I}_P |\psi\rangle_{ABP}) = M^x_a \otimes N^y_b \otimes \mathbb{I}_P(|\psi'\rangle_{A'B'} \otimes |\xi\rangle_{\bar{A}\bar{B}P})$$

for all $x, y \in \{1, ...m\}$, $a, b \in \{1, ..., \Delta\}$ and where

$$\begin{aligned} M^x_a &= \widetilde{\Lambda}^x_a \otimes S_0 + (\widetilde{\Lambda}^x_a)^* \otimes S_1; \\ N^y_b &= \widetilde{\Pi}^y_b \otimes T_0 + (\widetilde{\Pi}^y_b)^* \otimes T_1; \\ S_0 + S_1 &= \mathbb{I}_{\bar{A}}, \quad T_0 + T_1 = \mathbb{I}_{\bar{B}}; \\ \langle\xi| (S_0 &\otimes T_0 + S_1 \otimes T_1) \otimes \mathbb{I}_P |\xi\rangle = 1. \end{aligned}$$

The above definition can be understood by tracing out the ancilla and purification spaces $\mathcal{H}_{\bar{A}} \otimes \mathcal{H}_{\bar{B}} \otimes \mathcal{H}_P$

$$\Phi_A \otimes \Phi_B \otimes (\Lambda^x_a \otimes \Pi^y_b \rho_{AB}) = \langle\xi| S_0 \otimes T_0 |\xi\rangle \cdot \widetilde{\Lambda}^x_a \otimes \widetilde{\Pi}^y_b |\psi'\rangle\langle\psi'| + \langle\xi| S_1 \otimes T_1 |\xi\rangle \cdot (\widetilde{\Lambda}^x_a)^* \otimes (\widetilde{\Pi}^y_b)^* |\psi'\rangle\langle\psi'|$$

So we are effectively measuring a convex combination of $\widetilde{\Lambda}^x_a \otimes \widetilde{\Pi}^y_b$ and their complex conjugate. Notice that the probability of performing the conjugate depends on the junk state $|\xi\rangle$ and therefore is unknown.

3.7 SOS decomposition

We can rewrite the left hand side of any Bell's inequality

$$\mathcal{S} = \sum_{a,b,x,y} s_{x,y}^{a,b} p(a,b|x,y) \leq \beta$$

as the expectation value of an observable $\mathcal{B}$ called Bell operator

$$\mathcal{S} = \sum_{a,b,x,y} s_{x,y}^{a,b} p(a,b|x,y) = \langle\psi| \mathcal{B} |\psi\rangle ;$$
$$\mathcal{B} = \sum_{a,b,x,y} s_{x,y}^{a,b} \Lambda_a^x \otimes \Pi_b^y,$$

where $|\psi\rangle$ is a purification of the mixed state ρ shared by Alice and Bob. Furthermore, let β_q be the maximal violation of S achievable with quantum strategy (i.e. its quantum bound), so that

$$\beta_q - \mathcal{S} \geq 0 \implies \langle\psi| \beta_q \mathbb{I} - \mathcal{B} |\psi\rangle \geq 0, \quad \forall |\psi\rangle . \tag{3.4}$$

The operator $\beta_q \mathbb{I} - \mathcal{B}$ is called shifted Bell operator and equation (3.4) tells us that it is positive semidefinite. Now assume that the shifted Bell operator admits the following decomposition

$$\beta_q \mathbb{I} - \mathcal{B} = \sum_\lambda P_\lambda^\dagger P_\lambda$$

where each P_λ is a polynomial in Λ_a^x and Π_b^y. Such decomposition is called SOS (sum of squares). Now, if the state $|\psi\rangle$ maximally violates the Bell inequality, we can extract useful constraints on the system:

$$0 = \langle\psi| \beta_q \mathbb{I} - \mathcal{B} |\psi\rangle = \sum_\lambda \langle\psi| P_\lambda^\dagger P_\lambda |\psi\rangle = \sum_\lambda \langle P_\lambda \psi | P_\lambda \psi\rangle = \sum_\lambda \| P_\lambda |\psi\rangle \|^2,$$

and since the only vector with norm zero is the null vector, we conclude that

$$P_\lambda |\psi\rangle = 0, \qquad \forall \lambda. \tag{3.5}$$

Those relations often contains nontrivial statements about the system that can be used for self-testing.

3.8 An example of self-testing

In this section we will explicitly prove that from a maximal violation of the CHSH inequality, see section (2.2), we can self-test both state and measurements. For simplicity we will work with the observables, which are related to the projectors as

$$A_x = \sum_{a=\pm 1} a \Lambda_a^x;$$
$$B_y = \sum_{b=\pm 1} b \Pi_b^y.$$

From the properties of the projectors, equation (3.1), is easy to see that the observables satisfy:

$$\begin{aligned} &A_x^\dagger = A_x, \quad A_x^2 = \mathbb{I}; \\ &B_y^\dagger = B_y, \quad B_y^2 = \mathbb{I}; \\ &[A_x, B_y] = 0, \end{aligned} \tag{3.6}$$

the last property follows from the fact that Alice and Bob operators act on different Hilbert spaces. As usual, following the definition of self-testing, we will work with a purification $|\psi\rangle$ of the general mixed state ρ shared by Alice and Bob. Finally from equation (2.5) we see that the Bell operator corresponding to the CHSH inequality is

$$B = A_0 B_0 + A_1 B_0 + A_0 B_1 - A_1 B_1.$$

Quantum bound of the CHSH The quantum bound of the CHSH inequality is $\beta_q = 2\sqrt{2}$. Indeed the operator $2\sqrt{2}\mathbb{1} - B$ has the following SOS decomposition

$$2\sqrt{2}\mathbb{1} - B = \frac{1}{\sqrt{2}}\left[\left(\frac{A_0 + A_1}{\sqrt{2}} - B_0\right)^2 + \left(\frac{A_0 - A_1}{\sqrt{2}} - B_1\right)^2\right], \tag{3.7}$$

which follows from properties (3.6). Such decomposition ensures that $2\sqrt{2}\mathbb{1} - B$ is positive semidefinite, therefore

$$2\sqrt{2} \geq \langle\psi| B |\psi\rangle\,, \quad \forall |\psi\rangle$$

and in section (2.2) we found a realization that saturates the inequality. This concludes the proof.

Proof of self-testing Now assume that our state and measurements maximally violated the CHSH inequality. We will prove that, up to local isometries, Alice and Bob are measuring

$$\widetilde{A}_0 = \sigma_x, \quad \widetilde{A}_1 = \sigma_z, \quad \widetilde{B}_0 = \frac{\sigma_x + \sigma_z}{\sqrt{2}}, \quad \widetilde{B}_1 = \frac{\sigma_x - \sigma_z}{\sqrt{2}},$$

on the shared state

$$|\psi'\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \equiv |\phi_+\rangle\,.$$

The proof will be divided in two parts:

- Show that the observables anticommute on the support of the shared state

$$\{A_0, A_1\}|\psi\rangle = \{B_0, B_1\}|\psi\rangle = 0; \tag{3.8}$$

- Explicitly build the local isometry $|\Phi\rangle$, required by definition (3.3), with a quantum circuit.

Proof of anticommutativity To prove the anticommutativity we begin by applying equation (3.5) on the SOS decomposition (3.7), from which we derive:

$$\frac{A_0 + A_1}{\sqrt{2}}|\psi\rangle = B_0|\psi\rangle\,; \quad \frac{A_0 - A_1}{\sqrt{2}}|\psi\rangle = B_1|\psi\rangle\,, \tag{3.9}$$

Then consider the action of $\{B_0, B_1\}$ on $|\psi\rangle$:

$$\{B_0, B_1\}|\psi\rangle = (B_0B_1 + B_1B_0)|\psi\rangle = \frac{B_0(A_0 - A_1) + B_1(A_0 + A_1)}{\sqrt{2}}|\psi\rangle =$$
$$= \frac{(A_0 - A_1)(A_0 + A_1) + (A_0 + A_1)(A_0 - A_1)}{2}|\psi\rangle = 0$$

where we applied equations (3.9) and (3.6). Symmetrically, it is easy to see that the same results holds for $\{A_0, A_1\}$.

Building the local isometry At this point we are ready to explicitly build the local isometry required to perform self-testing. As anticipated we will use the swap gate circuit, as shown in figure 3.2. We choose

$$Z_A = \frac{1}{\sqrt{2}}(A_0 + A_1), \quad X_A = \frac{1}{\sqrt{2}}(A_0 - A_1), \quad Z_B = B_0, \quad X_B = B_1.$$

By applying equation (3.6) it is easy to see that Z_A and X_A anticommute:

$$\{Z_A, X_A\} = \frac{1}{2}\{A_0 + A_1, A_0 - A_1\} = A_0^2 - A_1^2 = 0 \tag{3.10}$$

and from equation (3.8) we derive that the same result applies for Z_B and X_B, on the support of the shared state

$$\{Z_B, X_B\}|\psi\rangle = \{B_0, B_1\}|\psi\rangle = 0. \tag{3.11}$$

Finally, equation (3.9) can be rewritten as

$$\begin{aligned} Z_A\left|\psi\right\rangle &= Z_B\left|\psi\right\rangle ; \\ X_A\left|\psi\right\rangle &= X_B\left|\psi\right\rangle . \end{aligned} \tag{3.12}$$

With those relations we are ready to apply the swap gate circuit. As you can see from figure 3.2 it is made by two components:

- The Hadamard gate. It is a one-qubit operator that acts on the basis $(\left|0\right\rangle, \left|1\right\rangle)$ as follows

$$\begin{aligned} \left|0\right\rangle &\to \frac{1}{\sqrt{2}}(\left|0\right\rangle + \left|1\right\rangle); \\ \left|1\right\rangle &\to \frac{1}{\sqrt{2}}(\left|0\right\rangle - \left|1\right\rangle). \end{aligned}$$

- The controlled-U gate. It is a (1+n)-qubits operator, defined for any n-qubits operator U. Its action is the following:

$$\begin{aligned} \left|0\right\rangle \otimes \left|\psi\right\rangle &\to \left|0\right\rangle \otimes \left|\psi\right\rangle ; \\ \left|1\right\rangle \otimes \left|\psi\right\rangle &\to \left|1\right\rangle \otimes U\left|\psi\right\rangle , \end{aligned}$$

  where $\left|\psi\right\rangle$ is a generic n-qubits vector. So basically it applies U on $\left|\psi\right\rangle$ only when the first qubit (usually called control qubit) is $\left|1\right\rangle$.

So let's compute the output of the swap and gate circuit: with reference to figure 3.2 we begin by applying the two Hadamard gates and the controlled Z_A and Z_B:

$$\begin{aligned} \left|00\right\rangle \otimes \left|\psi\right\rangle &\xrightarrow{\text{Hadamard}} \frac{1}{2}(\left|00\right\rangle + \left|01\right\rangle + \left|10\right\rangle + \left|11\right\rangle) \otimes \left|\psi\right\rangle \\ &\xrightarrow{\text{controlled } Z_A \text{ and } Z_B} \frac{1}{2}(\left|00\right\rangle + \left|01\right\rangle Z_B + \left|10\right\rangle Z_A + \left|11\right\rangle Z_A Z_B) \otimes \left|\psi\right\rangle . \end{aligned}$$

On this state we then apply again two Hadamard gates on the ancilla qubits

$$\frac{1}{4}\Big[(\left|00\right\rangle (1+Z_A)(1+Z_B) + \left|01\right\rangle (1+Z_A)(1-Z_B) + \left|10\right\rangle (1-Z_A)(1+Z_B) + \left|11\right\rangle (1-Z_A)(1-Z_B)\Big] \otimes \left|\psi\right\rangle ,$$

and eventually we apply the controlled X_A and X_B to get the final state

$$\begin{aligned} &\frac{1}{4}\Big[(\left|00\right\rangle \otimes (1+Z_A)(1+Z_B) + \left|01\right\rangle \otimes X_B(1+Z_A)(1-Z_B) + \\ &+ \left|10\right\rangle \otimes X_A(1-Z_A)(1+Z_B) + \left|11\right\rangle \otimes X_A(1-Z_A)X_B(1-Z_B)\Big]\left|\psi\right\rangle \equiv \left|\psi\right\rangle_F . \end{aligned} \tag{3.13}$$

This result can be simplified, for example the terms proportional to $\left|01\right\rangle$ and $\left|10\right\rangle$ vanish:

$$\begin{aligned} \left|01\right\rangle \otimes X_B(1+Z_A)(1-Z_B)\left|\psi\right\rangle &= \left|01\right\rangle \otimes X_B(1-Z_B^2)\left|\psi\right\rangle = 0; \\ \left|10\right\rangle \otimes X_A(1-Z_A)(1+Z_B)\left|\psi\right\rangle &= \left|10\right\rangle \otimes X_A(1-Z_B^2)\left|\psi\right\rangle = 0, \end{aligned}$$

where we used equation (3.12) and the unitarity of $Z_B = B_0$. Furthermore, notice that the terms proportional to $\left|11\right\rangle$ can be rewritten as:

$$\begin{aligned} &\left|11\right\rangle \otimes X_A(1-Z_A)X_B(1-Z_B)\left|\psi\right\rangle = \left|11\right\rangle \otimes (1+Z_A)X_A X_B(1-Z_B)\left|\psi\right\rangle = \\ &= \left|11\right\rangle \otimes (1+Z_A)(1+Z_B)X_A X_B\left|\psi\right\rangle = \left|11\right\rangle \otimes (1+Z_A)(1+Z_B)X_B^2\left|\psi\right\rangle = \\ &= \left|11\right\rangle \otimes (1+Z_A)(1+Z_B)\left|\psi\right\rangle . \end{aligned}$$

Where we used the two anti commutation relations (3.10), (3.11) and the unitarity of $X_B = B_1$. By plugging everything back in final state $\left|\psi\right\rangle_F$ we find the simplified expression

$$\left|\psi\right\rangle_F = \frac{\left|00\right\rangle + \left|11\right\rangle}{\sqrt{2}} \otimes \left(\frac{1}{4\sqrt{2}}(1+Z_A)(1+Z_B)\left|\psi\right\rangle\right) \equiv \frac{\left|00\right\rangle + \left|11\right\rangle}{\sqrt{2}} \otimes \left|\xi\right\rangle .$$

And this concludes the proof: with the swap gate circuit we managed to build an isometry that maps the initial, purified, state $\left|\psi\right\rangle$ to the maximal entangled state $\left|\phi_+\right\rangle$, with the addition of some junk state $\left|\xi\right\rangle$. Notice that the local isometry does not act on the purification space, as required by the definition given in section (3.3).

Self-testing of measurements with the same circuit it's also possible to self-test measurements. As an example we will do the explicit calculation for the observable B_0. So first of all we have to find the action of the swap gate circuit on the initial state $B_0 |\psi\rangle$. It is easy to see that the result is

$$\begin{aligned}&\frac{1}{4}\Big[(|00\rangle \otimes (1+Z_A)(1+Z_B) + |01\rangle \otimes X_B(1+Z_A)(1-Z_B)+ \\ &+|10\rangle \otimes X_A(1-Z_A)(1+Z_B) + |11\rangle \otimes X_A(1-Z_A)X_B(1-Z_B)\Big] Z_B |\psi\rangle \equiv |\psi\rangle_F\,,\end{aligned} \tag{3.14}$$

which is the same as equation (3.13) with the substitution

$$|\psi\rangle \to B_0 |\psi\rangle = Z_B |\psi\rangle\,.$$

Similarly to what happened in self-testing of the state, the terms proportional to $|01\rangle$ and $|10\rangle$ vanish:

$$\begin{aligned}&|10\rangle \otimes X_A(1-Z_A)(1+Z_B)Z_B |\psi\rangle = |10\rangle \otimes X_A(1-Z_A)(1+Z_B)|\psi\rangle = 0;\\ &|01\rangle \otimes X_B(1+Z_A)(1-Z_B)Z_B |\psi\rangle = -|01\rangle \otimes X_B(1+Z_A)(1-Z_B)|\psi\rangle = 0.\end{aligned}$$

Then, we rewrite the term proportional to $|11\rangle$ as

$$\begin{aligned}&|11\rangle \otimes X_A(1-Z_A)X_B(1-Z_B)Z_B |\psi\rangle = -|11\rangle \otimes (1+Z_A)X_AX_B(1-Z_B)|\psi\rangle = \\ &= -|11\rangle \otimes (1+Z_A)(1+Z_B)|\psi\rangle = .-|11\rangle \otimes (1+Z_A)(1+Z_B)Z_B|\psi\rangle\end{aligned}$$

Finally, by plugging everything back in equation (3.14), we find that the state after the application of the circuit is

$$|\psi_F\rangle = \frac{|00\rangle - |11\rangle}{2} \otimes |\xi\rangle = (\mathbb{1} \otimes \sigma_z)\frac{|00\rangle + |11\rangle}{2} \otimes |\xi\rangle\,,$$

where the junk state $|\xi\rangle$ is

$$|\xi\rangle = \frac{1}{4\sqrt{2}}(1+Z_A)(1+Z_B)Z_B|\psi\rangle$$

This concludes the proof: as required by definition (3.4), we showed that after the application of the isometry the initial state $B_0|\psi\rangle$ is mapped to $\sigma_z\phi_+$ with the addition of a junk state $|\xi\rangle$. The same proof can be easily adapted to the other three measurements A_0, A_1 and B_1.

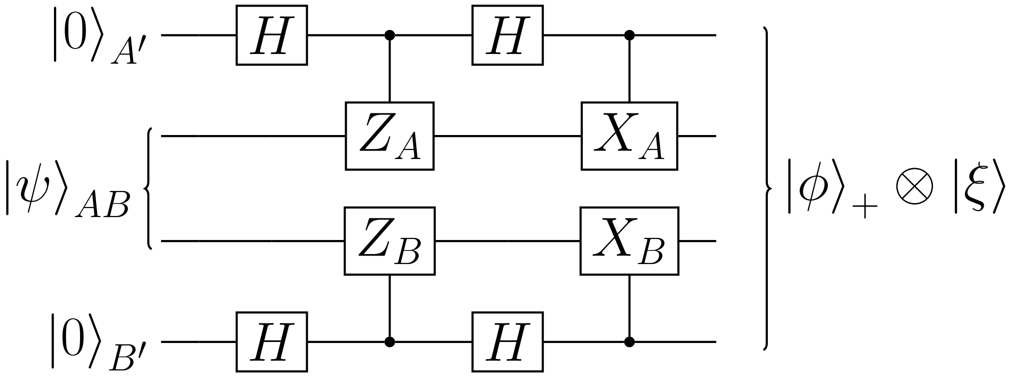

Figure 3.2: Swap gate circuit used to self test the CHSH inequality. It is made by H which is the Hadamard gate, and controlled-U gates, with $U = Z_A, Z_B, X_A, X_B$.

Operator regularization In the above proof of self-testing we overlooked an important detail: is the isometry made by unitary operators? To verify that, we need to check if all elements of the swap gate circuit are unitary. Clearly it is the case for both the Hadamard gate and for

$$Z_B = B_0, \quad X_B = B_1.$$

since we assumed the Bob's observables to be unitary, see equation (3.6). However, the two operators

$$Z_A = \frac{A_0 + A_1}{\sqrt{2}}, \quad X_a = \frac{A_0 - A_1}{\sqrt{2}} \tag{3.15}$$

are in general not unitary and, in particular, they might have some zero eigenvalues. To fix this issue we apply a regularization procedure where:

- We change all zero eigenvalues of Z_A and X_A to 1, resulting in two new operators $\widetilde{Z}_A$ and $\widetilde{X}_A$
- We re-normalize both operators by defining

  $$\hat{Z}_A = \frac{\widetilde{Z}_A}{|\widetilde{Z}_A|} \equiv \widetilde{Z}_A|\widetilde{Z}_A|^{-1}; \quad \hat{X}_A = \frac{\widetilde{X}_A}{|\widetilde{X}_A|} \equiv \widetilde{X}_A|\widetilde{X}_A|^{-1},$$

  where in general, for any operator O, we define $|O|$ to be

  $$|O| = \sqrt{OO^\dagger},$$

  so that $\hat{Z}_A$ and $\hat{X}_A$ are both unitary by construction;
- We prove these new operators act on the state $|\psi\rangle$ in the same way the original ones do:

  $$Z_A |\psi\rangle = \hat{Z}_A |\psi\rangle, \quad X_A |\psi\rangle = \hat{X}_A |\psi\rangle .$$

Let us prove the last step for Z_A (for X_A the proof is the same), it is sufficient to show that

$$\|(\hat{Z}_A - Z_A) |\psi\rangle\| = 0.$$

The left hand site of the equation can be written as

$$\|(\hat{Z}_A - Z_A) |\psi\rangle\| = \|\hat{Z}_A(\mathbb{I} - \hat{Z}_A^\dagger Z_A) |\psi\rangle\| = \|(\mathbb{I} - \hat{Z}_A^\dagger Z_A) |\psi\rangle\|, \tag{3.16}$$

where we used that $\hat{Z}_A$ is unitary. Then, notice that for any function $f(\widetilde{Z}_A)$ the following equality holds:

$$f(\widetilde{Z}_A)Z_A |\phi\rangle = f(Z_A)Z_A |\phi\rangle, \qquad \forall |\phi\rangle \in \mathcal{H}_A.$$

Indeed if $|\phi\rangle$ is in the kernel of Z_A we have that

$$f(\widetilde{Z}_A)Z_A |\phi\rangle = f(Z_A)Z_A |\phi\rangle = 0,$$

and in all other cases by definition the action of $\widetilde{Z}_A$ is the same of Z_A. With this property in mind and by choosing

$$f(\widetilde{Z}_A) = \frac{\widetilde{Z}_A^\dagger}{|\widetilde{Z}_A|},$$

we can rewrite the right hand side of equation (3.16) as

$$\|(\mathbb{I} - \hat{Z}_A^\dagger Z_A) |\psi\rangle\| = \left\| \left(\mathbb{I} - \frac{\widetilde{Z}_A^\dagger}{|\widetilde{Z}_A|} Z_A \right) |\psi\rangle \right\| = \left\| \left(\mathbb{I} - \frac{Z_A^\dagger}{|Z_A|} Z_A \right) |\psi\rangle \right\| = \|(\mathbb{I} - |Z_A|) |\psi\rangle\| .$$

Finally we have that

$$\|(\mathbb{I} - |Z_A|) |\psi\rangle\| = \|(\mathbb{I} - |Z_A Z_B|) |\psi\rangle\| \leq \|(\mathbb{I} - Z_A Z_B) |\psi\rangle\| = 0,$$

where we used the unitarity of Z_B, the operator inequality $|AB| \leq AB$ and equation (3.12). So we can conclude that

$$\|(\hat{Z}_A - Z_A) |\psi\rangle\| \leq 0 \implies \|(\hat{Z}_A - Z_A) |\psi\rangle\| = 0,$$

which is exactly what we wanted to prove.

Chapter 4

NPA hierarchy

Many problems of quantum information theory require solving an optimization problem on the quantum set $\mathcal{Q}$. Two examples, that we will discuss in more details in section 4.9 and chapter 5, are finding the quantum bound of a Bell's inequality and generating random numbers from the outcomes of measurements. To perform an optimization we need first of all a way to characterize the quantum set, that could be done by solving the following problem: given a behavior $p(a,b|x,y)$ does it belongs to the quantum set? In other words, we would like to determine the existence of state and measurements such that

$$p(a,b|x,y) = \mathrm{Tr}[\rho_{AB}\Lambda_a^x \otimes \Pi_b^y].$$

In this chapter we will present an algorithm, called NPA hierarchy, that allows us to asymptotically solve such problem. It has been originally developed in [9] and our discussion will be based on this paper. As we will see the NPA hierarchy consists in checking the existence of an infinite sequence of matrices

$$\Gamma^0, \Gamma^1, \Gamma^2, \Gamma^3, \ldots$$

These matrices, referred to as certificates, must satisfy the following conditions:

- Certain constraints dependent on $p(a,b|x,y)$;
- The positive semidefinite condition $\Gamma^n \succcurlyeq 0 \; \forall n \in \mathbb{N}$.

More details will be provided later, for now we anticipate only that checking their existence can be formulated as a semidefinite program (SDP). Therefore, we will begin by explaining what a SDP is.

4.1 Semidefinite programming

Semidefinite programming is a subfield of mathematical optimization concerned in solving the following problem: maximizing (or minimizing) a linear combination of the entries of an (unfixed) positive semidefinite matrix, with the optional addition of linear constraints. It is also known as the primal problem, and can be mathematically formulated as:

$$\begin{aligned}&\text{maximize}\,\mathrm{Tr}[GZ];\\&\text{subject to}\,\mathrm{Tr}[F_iZ] = c_i, \quad i = 1, \ldots, d;\\&\text{subject to } Z \succcurlyeq 0,\end{aligned} \tag{4.1}$$

where the variable of the problem is the $n \times n$ matrix Z, while G and F_i are fixed $n \times n$ matrices and c_i are fixed scalars. Due to the generality of equation (4.1) SDP have numerous applications and they can be efficiently solved with the interior point method [12]. To each primal problem there is an associated dual problem, which is a minimization of the form

$$\begin{aligned}&\text{minimize } c^T x;\\&\text{subject to } F(x) = G - \sum_{i=1}^{p} x_i F_i \succcurlyeq 0.\end{aligned}$$

This time the variable is the d-dimensional vector x. The dual problem is also semidefinite program, which means that it can be rewritten in the form (4.1). A vector x is said to be dual feasible when $F(x) \succcurlyeq 0$, and similarly a matrix Z that satisfy all constraints of (4.1) is said to be primal feasible. An important property of the dual program is that it yields a bound on the optimal value of the primal program. To see this let x and Z be respectively a dual and primal feasible points of the same problem, then

$$\mathrm{Tr}[GZ] - c^T x = \mathrm{Tr}[GZ] - \sum_{i=1}^{p} \mathrm{Tr}[F_i Z] x_i = \mathrm{Tr}[F(x)Z] \geq 0.$$

where the last inequality follows from the fact that both $F(x)$ and Z are positive semidefinite. Therefore the optimal primal values p^* and the optimal dual value d^* satisfy

$$d^* \leq p^*. \tag{4.2}$$

This result is called weak duality. Under certain conditions, for example the existence of a positive definite primal feasible solution $Z \succ 0$, it can be proven that

$$d^* = p^*,$$

this result is called strong duality. Concretely, there are many algorithms that allow to solve the primal and dual problem at the same time and output a pair (p, d). Then, duality theorems are a powerful tool to verify their optimality: for example if the strong duality holds, the condition $p = d$ (or a small gap between then), guarantees that the solution is indeed the optimal one. There are many numerical software packages that can be used to solve a SDP. In this work we have used SDPA and SDPA-DD, both open source and written in C++.

4.2 Characterizing the quantum set

At this point we go back to the original problem: given a behavior $p(a, b|x, y)$, can we write it in the form of equation (2.3)? In analogy to what we did in the previous chapter, we can perform a purification and the problem is simplified to finding two sets of operators $\Lambda = \{\Lambda_a^x\}$ and $\Pi = \{\Pi_b^y\}$ satisfying

$$\begin{aligned}
&(\Lambda_a^x)^\dagger = \Lambda_a^x, \quad (\Pi_b^y)^\dagger = \Pi_b^y, \quad \text{(hermiticity)};\\
&\Lambda_a^x \Lambda_{a'}^x = \delta_{a,a'} \Lambda_a^x, \quad \Pi_b^y \Pi_{b'}^y = \delta_{b,b'} \Pi_b^y, \quad \text{(orthogonality)};\\
&\sum_a \Lambda_a^x = \mathbb{I}, \quad \sum_b \Pi_b^y = \mathbb{I}, \quad \text{(completeness)};\\
&[\Lambda_a^x, \Pi_b^y] = 0, \quad \text{(commutativity)},
\end{aligned} \tag{4.3}$$

and a pure state $|\psi\rangle$ such that there is compatibility with the behavior:

$$p(a, b|x, y) = \langle\psi| \Lambda_a^x \Pi_b^y |\psi\rangle .$$

The completeness condition is there to ensure that the marginal distributions

$$\begin{aligned}
p(a|x) &\equiv \sum_b p(a, b|x, y);\\
p(b|y) &\equiv \sum_a p(a, b|x, y),
\end{aligned}$$

are well defined, more precisely the no-signaling condition is satisfied, see equation (2.1). It also generates some redundancy: not all operators are independent, for each x we can pick an observable $\Lambda_{a_x}^x$ and rewrite it in function of the others

$$\Lambda_{a_x}^x = \mathbb{I} - \sum_{a \neq a_x} \Lambda_a^x, \quad \forall x$$

and same for Bob's operators

$$\Pi^y_{b_y} = \mathbb{I} - \sum_{b \neq b_y} \Pi^y_b, \quad \forall y.$$

So it's not restrictive defining two reduced sets

$$\begin{aligned} \widetilde{\Lambda} &= \{\Lambda^x_a,\ \forall a, x\} - \{\Lambda^x_{a_x},\ \forall x\}; \\ \widetilde{\Pi} &= \{\Pi^y_b,\ \forall b, y\} - \{\Pi^y_{b_y},\ \forall y\}, \end{aligned} \tag{4.4}$$

on which we build an alternative definition: a behavior $p(a, b|x, y)$ is quantum if we can find a pure state $|\psi\rangle$ and two sets of operators $\widetilde{\Lambda}$ and $\widetilde{\Pi}$ such that for each $\Lambda^x_a \in \widetilde{\Lambda}$ and $\Pi^y_b \in \widetilde{\Pi}$:

$$\begin{aligned} &p(a, b|x, y) = \langle\psi|\, \Lambda^x_a \Pi^y_b \,|\psi\rangle\,; \\ &p(a|x) = \langle\psi|\, \Lambda^x_a \,|\psi\rangle\,; \\ &p(b|y) = \langle\psi|\, \Pi^y_b \,|\psi\rangle\,, \end{aligned}$$

and the operators satisfy

$$\begin{aligned} &(\Lambda^x_a)^\dagger = \Lambda^x_a, \quad (\Pi^y_b)^\dagger = \Pi^y_b, \quad \text{(hermiticity)}; \\ &\Lambda^x_a \Lambda^x_{a'} = \delta_{a,a'} \Lambda^x_a, \quad \Pi^y_b \Pi^y_{b'} = \delta_{b,b'} \Pi^y_b, \quad \text{(orthogonality)}; \\ &[\Lambda^x_a, \Pi^y_b] = 0, \quad \text{(commutativity)}, \end{aligned} \tag{4.5}$$

The two definitions are of course equivalent but for practical reasons we will use the second, as it requires less operators and therefore a smaller running time of the NPA algorithm.

4.3 Sets of sequences

Before presenting the NPA algorithm we need a few more definitions. Let $\mathcal{E}$ be the complete set of reduced projectors with the addition of the identity

$$\mathcal{E} = \widetilde{\Lambda} \cup \widetilde{\Pi} \cup \mathbb{I}$$

and let $\mathcal{O}$ be a set of n operators

$$\mathcal{O} = \{O_1, O_2, ..., O_n\},$$

where each O_i is a linear combination of products of projectors in $\mathcal{E}$. Then define $\mathcal{F}(\mathcal{O})$ as the set of all independent equalities of the form

$$\sum_{i,j=1}^{n} F_{i,j} \,\langle\psi|\, O_i^\dagger O_j \,|\psi\rangle = g(p), \tag{4.6}$$

where $F_{i,j}$ are real numbers and $g(p)$ is a linear function of the probabilities

$$g(p) = g_0 + \sum_{a,b,x,y} g^{x,y}_{a,b}\, p(a, b|x, y), \tag{4.7}$$

with coefficients chosen in such a way that equation (4.6) is satisfied. Concretely the set $\mathcal{F}(\mathcal{O})$ depends on the choice of $\mathcal{O}$ and on the properties of the projectors given in equation (4.5). As an example assume that $\mathcal{O}$ is made by

$$\mathcal{O} = \{\Lambda^x_a \Pi^y_b,\ \Lambda^x_a \in \widetilde{\Lambda},\ \Pi^y_b \in \widetilde{\Pi}\}.$$

In this case simple case $\mathcal{F}(\mathcal{O})$ is made by the following equalities

$$\langle\psi|\, (\Lambda^x_{a_1} \Pi^y_{b_1})^\dagger (\Lambda^x_{a_2} \Pi^y_{b_2}) \,|\psi\rangle = \langle\psi|\, \Lambda^x_{a_1} \Lambda^x_{a_2} \Pi^y_{b_1} \Pi^y_{b_2} \,|\psi\rangle = \delta_{a_1,a_2} \delta_{b_1,b_2}\, p(a_1, b_1|x, y).$$

Let a sequence S be a non-null operator made as product of projectors in $\mathcal{E}$. The length $|S|$ is the minimum number of projectors needed to generate it. For example the length of $S = \Lambda^x_a \Pi^y_b \Lambda^x_a$ is

$$|S| = |\Lambda^x_a \Pi^y_b \Lambda^x_a| = |\Pi^y_b \Lambda^x_a \Lambda^x_a| = |\Pi^y_b \Lambda^x_a| = 2 \tag{4.8}$$

By convention, we say that the identity operator has length $|\mathbb{I}| = 0$. We define S_n as the sets of sequences with length smaller or equal to n:

$$\begin{aligned}
&S_0 = \{\mathbb{I}\} \\
&S_1 = S_0 \cup \{\Lambda_i \in \widetilde{\Lambda}\} \cup \{\Pi_i \in \widetilde{\Pi}\} \\
&S_2 = S_0 \cup S_1 \cup \{\Lambda_i\Lambda_j \in \widetilde{\Lambda}\} \cup \{\Pi_i\Pi_j \in \widetilde{\Lambda}\} \cup \{\Lambda_i\Pi_j,\ \Lambda_i \in \widetilde{\Lambda},\ \Pi_j \in \Pi\} \\
&S_3 = ...
\end{aligned} \tag{4.9}$$

It follows that $S_0 \subseteq S_1 \subseteq S_2 ...$ and that each $O_i \in \mathcal{O}$ can be written as linear combination of elements of S_n for n big enough. Finally notice that we are considering only non-null operators, so for example

$$S = \Lambda^x_{a_1}\Lambda^x_{a_2} = 0, \quad a_1 \neq a_2$$

doesn't belong to any S_n.

4.4 Certification of quantum behaviors

With the formalism defined in the previous section we can begin finding conditions that restricts quantum behaviors.

Proposition, a necessary condition for quantum behaviors: Let $\mathcal{O}$ be a set of n operators and let $\mathcal{F}(\mathcal{O})$ be the set of independent equalities defined in (4.6). Then, a necessary condition for the behavior $p(a, b|x, y)$ to be quantum is that there exist a $n \times n$ complex hermitian positive semidefinite matrix Γ, that satisfy

$$\sum_{i,j=1}^{n} F_{i,j}\Gamma_{i,j} = g(p), \quad \forall F, g \in \mathcal{F}(\mathcal{O}). \tag{4.10}$$

Furthermore, if all $g(p)$ and $F_{i,j}$ are real Γ can be chosen real as well.

The proof of this proposition is simple: as Γ choose

$$\Gamma_{i,j} = \langle\psi| O_i^\dagger O_j |\psi\rangle ,$$

which by definition of $\mathcal{F}(\mathcal{O})$ satisfies equation (4.10). Then, Γ is positive semidefinite since for all $\in \mathbb{C}^n$

$$v^\dagger \Gamma v = \sum_{i,j} v_i^* \langle\psi| O_i^\dagger O_j |\psi\rangle v_j = \langle\psi| V^\dagger V |\psi\rangle ,$$

where $V = \sum_i O_i v_i$. If all $g(p)$ and F are real we can re-define

$$\Gamma \to \frac{\Gamma + \Gamma^*}{2},$$

which is positive semidefinite, real and satisfies equation (4.10). Such Γ is usually called a certificate associate to $\mathcal{O}$.

An example of certification To make an explicit example let's consider the most simple case in which both Alice and Bob have only 2-inputs and 2-outputs. Define the following average quantities

$$\begin{aligned}
C^A_x &= \sum_a p(a|x)a, \quad x \in \{0, 1\}; \\
C^B_y &= \sum_a p(b|y)b, \quad y \in \{0, 1\}; \\
C_{xy} &= \sum_{a,b} p(a, b|x, y)ab, \quad x, y \in \{0, 1\}.
\end{aligned}$$

There are 8 such numbers and they are fully equivalent to the probability distribution $p(a, b|x, y)$. For this case if the data measured by Alice and Bob corresponds to a quantum system, then there exists a real 5×5 positive semidefinite matrix Γ of the form

$$\Gamma = \begin{pmatrix} 1 & C_0^A & C_1^A & C_0^B & C_1^B \\ C_0^A & 1 & u & C_{00} & C_{01} \\ C_1^A & u & 1 & C_{10} & C_{11} \\ C_0^B & C_{00} & C_{10} & 1 & v \\ C_1^B & C_{01} & C_{11} & v & 1 \end{pmatrix} \tag{4.11}$$

where u and v are arbitrary real entries. Indeed, if the behavior is quantum we have the usual state $|\psi\rangle$ and projectors Λ_a^x and Π_b^y satisfying (4.5). Then, let $\mathcal{O}$ be

$$\begin{aligned} &\mathcal{O} = \{\sigma_0, \sigma_0^A, \sigma_1^A, \sigma_0^B, \sigma_1^B\}, \text{ where} \\ &\sigma_0 = \mathbb{I}; \\ &\sigma_x^A = \sum_a \Lambda_a^x a; \\ &\sigma_y^B = \sum_b \Pi_b^y b. \end{aligned}$$

The corresponding set $\mathcal{F}(\mathcal{O})$ is made by the following equalities

$$\begin{aligned} &\langle\psi| \sigma_0^\dagger \sigma_0 |\psi\rangle = 1; \\ &\langle\psi| \sigma_0^\dagger \sigma_x^A |\psi\rangle = C_x^A, \quad \langle\psi| \sigma_0^\dagger \sigma_y^B |\psi\rangle = C_y^B \\ &\langle\psi| \left(\sigma_x^A\right)^\dagger \sigma_x^A |\psi\rangle = \langle\psi| \left(\sigma_y^B\right)^\dagger \sigma_y^B |\psi\rangle = 1; \\ &\langle\psi| \left(\sigma_x^A\right)^\dagger \sigma_y^B |\psi\rangle = C_{x,y}. \end{aligned}$$

and therefore we can conclude that the associated certification $\Gamma_{i,j} = \langle\psi| O_i^\dagger O_j |\psi\rangle$ has the form of equation (4.11). So practically Alice and Bob will perform their measurements, estimate the probability distribution and build the Γ matrix (4.11), up to the two unknown coefficients u and v. Then, if they prove that for each value of u and v the matrix Γ is NOT positive semidefinite, they can conclude that surely their behavior is NOT quantum. As anticipated, this last part can be solved efficiently with a SDP.

4.5 Casting to SDP

In general checking the existence of a certificate Γ, see equation (4.10), is analogous to solving the following optimization problem:

$$\begin{aligned} &\text{maximize } \lambda; \\ &\text{subject to} \operatorname{Tr}[F^T\Gamma] = g(p), \quad \forall F, g \in \mathcal{F}(\mathcal{O}); \\ &\text{subject to } \Gamma - \lambda\mathbb{I} \succcurlyeq 0. \end{aligned} \tag{4.12}$$

Indeed a solution $\lambda \geq 0$ implies that

$$\Gamma \succcurlyeq \lambda\mathbb{I} \succcurlyeq 0,$$

and therefore a valid certificate exists. On the other hand a strictly negative solution $\lambda < 0$ implies that any solution Γ compatible with (4.10) is negative definite, and hence the behavior is not quantum. Equation (4.12) has the form of a primal SDP problem, and it can be proven that the corresponding dual is

$$\begin{aligned} &\text{minimize} \sum_{g_k(p) \in \mathcal{F}(\mathcal{O})} g_k(p) y_k \\ &\text{subject to } F(y) = \sum_{F_k \in \mathcal{F}(\mathcal{O})} F_k^T y_k \succcurlyeq 0 \\ &\text{subject to} \sum_{F_k \in \mathcal{F}(\mathcal{O})} y_k \operatorname{Tr}[F_k^T] = 1. \end{aligned}$$

From the dual problem we can also generate Bell inequalities: assume that a given behavior $q(a,b|x,y)$ returns a negative solution λ for the primal problem. This implies that $q(a,b|x,y)$ is not quantum and furthermore, from the weak duality theorem, see equation (4.2), we know that the solution of the dual problem is negative as well

$$S \equiv \sum_{g_k(q)\in\mathcal{F}(\mathcal{O})} g_k(q)y_k < 0,$$

and such expression is a Bell inequality violated by $q(a,b|x,y)$ indeed:

- The coefficients $g_k(q)$ are linear in $q(a,b|,y)$, as we showed in equation (4.7);
- All quantum behaviors yield $S \geq 0$, because

  $$S = \sum_{g_k(q)\in\mathcal{F}(\mathcal{O})} g_k(q)y_k = \mathrm{Tr}[F(y)\Gamma] \geq 0,$$

  and where the last inequality follows from $\Gamma, F(y) \succcurlyeq 0$.

4.6 Equivalence between certificates

Consider a given probability distribution $p(a,b|x,y)$. We saw that to each set of operators $\mathcal{O}$ corresponds a different certificate Γ. In this section we will see that not all certificates are linearly independent:

Proposition Let $\mathcal{O}$ and $\mathcal{O}'$ be two sets of operators such that every operator in $\mathcal{O}'$ is a linear combination of operators in $\mathcal{O}$. Then, the existence of a certificate Γ associated to $\mathcal{O}$ (for a given $p(a,b|x,y)$) implies the existence of a certificate Γ' associated to $\mathcal{O}'$.

To prove the proposition we begin by applying the hypothesis of linear dependence: each $O'_i \in \mathcal{O}'$ can be written as

$$O'_i = \sum_k C_{i,k}O_k, \quad O_k \in \mathcal{O}.$$

Therefore, the corresponding certificate for $\mathcal{O}'$ is related to the one of $\mathcal{O}$:

$$\Gamma'_{i,j} = \langle\psi|\,(O'_i)^\dagger O'_j\,|\psi\rangle = \sum_{k,h} C^*_{k,i}\,\langle\psi|\,O^\dagger_k O_h\,|\psi\rangle\,C_{j,h} = \sum_{k,h} C^*_{k,i}\Gamma_{k,h}C_{j,h}$$

from which we derive that

$$\Gamma' = C^\dagger\Gamma C.$$

By hypothesis Γ is positive semidefinite and this implies that Γ' has the same property. Therefore Γ' is a certificate associated to $\mathcal{O}'$ and this ends the proof.

So when looking for certificates we do not have to consider generic sets of operators $\mathcal{O}$. We can instead just focus on the sets of sequences S_n $\forall n \geq 0$, since their linear combination generate all other possible operators, see equation (4.9).

4.7 A hierarchy of necessary conditions

Following the results of the previous section, we define a certificate of order n, denoted as Γ^n, to be a certificate associated to the set S_n. Γ^n is a $|S_n|$x$|S_n|$ matrix and we will index its entries with this convention:

- Generic operators $S,T \in S_n$ are indexed with s and t respectively;
- The identity $\mathbb{I}$ is indexed with 1;
- Alice's projectors Λ^x_a are indexed with a_x and similarly Bob's projectors Π^y_b are indexed with b_y;

- A product of operators is indexed by concatenating the indexes of each operator. For example $S\Lambda_a^x$ is indexed with sa_x.

So with those rules Γ^n entries are

$$\Gamma^n_{s,t} = \langle\psi|\, S^\dagger T \,|\psi\rangle\,, \quad S, T \in S_n. \tag{4.13}$$

and in particular we can link some of them to the probability distribution $p(a,b|x,y)$:

$$\Gamma^n_{1,1} = 1, \quad \Gamma^n_{1,a_x} = p(a|x), \quad \Gamma^n_{1,b_y} = p(b|y), \quad \Gamma^n_{a_x,b_y} = p(a,b|x,y). \tag{4.14}$$

Not all entries are independent, indeed projectors properties (4.5) lead to constraints, for example:

$$\Gamma^n_{a_x b_y, a_x} = \Gamma^n_{a_x b_y, 1}, \quad \Gamma^n_{a_x, a'_x} = 0, \quad ... \tag{4.15}$$

and of course it is also symmetric $\Gamma^n_{s,t} = \Gamma^n_{t,s}$. The property $S_0 \subseteq S_1 \subseteq S_2...$ implies that the positive semidefiniteness of certificates $\Gamma^0, \Gamma^1, \Gamma^2, ...$ represents a hierarchy of conditions satisfied by quantum probabilities and where each condition $\Gamma^n \succcurlyeq 0$ is stronger than the previous one $\Gamma^{n-1} \succcurlyeq 0$. So for each behavior $p(a,b|x,y)$ we can build an algorithm to check its quantum properties:

1. Set $n = 1$;
2. Build the matrix Γ^n by using its relation with the behavior $p(a,b|x,y)$, equation (4.14), and the constraints between entries, equation (4.15);
3. Check if Γ^n is positive semidefinite. If the result is true, set $n = n+1$ and repeat from step 2, otherwise the behavior is non-quantum and the algorithm ends.

It is also natural to define, for any $n \geq 0$, a set Q^n, that contains all behaviors $p(a,b|x,y)$ such that the corresponding matrix Γ^n is positive semidefinite. Those sets satisfy

$$Q^1 \supseteq Q^2 \supseteq ... \supseteq \mathcal{Q} \supseteq \mathcal{Q}', \tag{4.16}$$

where we recall that $\mathcal{Q}'$ is the quantum set with the tensor product structure between Alice and Bob, while $\mathcal{Q}$ is the less constrained where we just require that Alice and Bob operators commute, see section 2.1. Equation (4.16) is geometrically represented in figure 4.1.

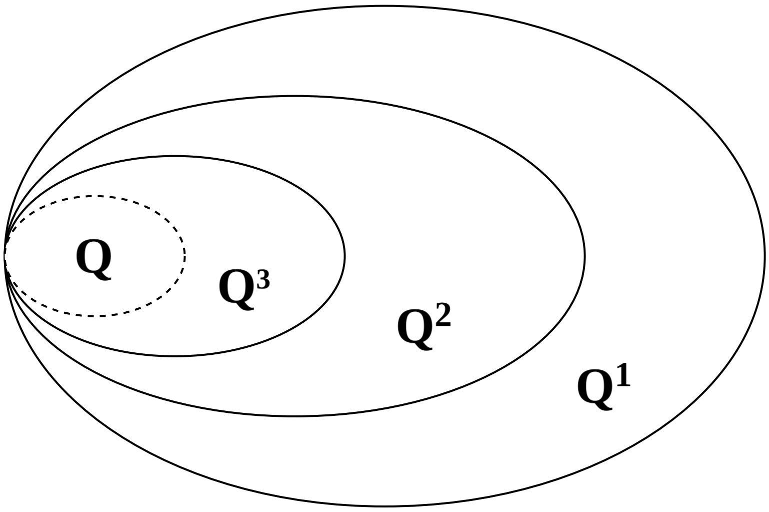

Figure 4.1: Geometric representation of the NPA hierarchy.

Furthermore, it can be proven, see [9], that in the limit $n \to \infty$ we recover exactly the quantum set

$$\lim_{n\to\infty} Q^n = \mathcal{Q}, \tag{4.17}$$

which means that, for all non-quantum behaviors $p(a,b|x,y)$, there exist a n such that Γ^n is not positive semidefinite.

Intermediate levels of the hierarchy Sometimes the running time needed to perform a simulation at the second order of the hierarchy is already too big. In this cases instead of stopping at the level Γ^1 it is useful to introduce an intermediate order:

$$S_{1+AB} = S_0 \cup S_1 \cup \{\Lambda_i \Pi_j,\ \Lambda_i \in \widetilde{\Lambda},\ \Pi_j \in \widetilde{\Pi}\}. \tag{4.18}$$

The name AB comes from the fact that we are considering the mixed term $\Lambda_i \Pi_j$, which is made by one Alice's projector (A) and one Bob's projector (B). Concretely, we are neglecting the AA and BB elements of the S_2 set:

$$\{\Lambda_i \Lambda_j\} \cup \{\Pi_i \Pi_j\}.$$

This has the overall effect of making the certificate matrix Γ smaller in size, hence checking if its positive semidefinite require less time. Similar intermediate levels can be defined for higher orders in the hierarchy.

4.8 Stopping criteria

So far we derived a hierarchy of conditions that allow us to asymptotically characterize the quantum set. While it is a good theoretical result, in practice it only allows us to test if a given behavior is non-quantum, more precisely whether it belongs to Q^n where n is the maximum value that we can numerically test with a viable running time. However, in some cases it's possible to prove, at a finite order n in the hierarchy, that a behavior belongs to $\mathcal{Q}$. To see how we first have to introduce the concept of rank loops.

Rank loops Let Γ^n be a certificate of order n associated to a given behavior $p(a,b|x,y)$. For two fixed inputs x and y consider the set J_{xy} made by all sequences of the form

$$J_{xy} = \{\Lambda_a^x \Pi_b^y T\ \forall a,b,T,\ \text{with } T \in S_n\} \cup S_{n-1},$$

and let Γ_{xy}^n be the sub matrix of Γ^n with entries

$$(\Gamma_{xy}^n)_{s,t} = \Gamma_{s,t}^n, \qquad \text{such that } S,T \in J_{xy} \cap S_n.$$

If for any choice of the inputs the rank of the two matrices is the same

$$\text{rank}(\Gamma_{xy}^n) = \text{rank}(\Gamma^n), \quad \forall x,y,$$

we say that the certificate Γ^n has a rank loop. With this definition we are ready to state the stopping criteria theorem.

Theorem, stopping criteria A behavior $p(a,b|x,y)$ has a quantum representation of finite dimension d if and only if $p(a,b|x,y)$ admits, for a finite n, a certificate Γ^n of order n, with a rank loop and $\text{rank}(\Gamma^n) \leq d$.

Note that with representation of finite dimension d, we mean that the behavior can be realized with states and measurements belonging to a Hilbert state $\mathcal{H}$ of dimension $\dim(\mathcal{H}) = d$. For a proof of the theorem see [9], here we will only focus on some practical applications: assume that the behavior $p(a,b|x,y)$ we are testing admits a quantum representation of dimension d and consider the series of certificates $\Gamma^1, \Gamma^2, \ldots$ From $S_n \subseteq S_{n+1}$ it follows that Γ^n is a sub matrix of Γ^{n+1}, and therefore

$$\text{rank}(\Gamma^n) \leq \text{rank}(\Gamma^{n+1}),$$

so the sequence of ranks is non-decreasing. At the same time the theorem tells us that

$$\text{rank}(\Gamma^n) \leq d, \quad \forall n,$$

and we can conclude that there exist a N such that

$$\text{rank}(\Gamma^N) = \text{rank}(\Gamma^{N+1}). \tag{4.19}$$

On the other hand, for all x, y

$$\text{rank}(\Gamma^N) \leq \text{rank}(\Gamma^{N+1}_{xy}) \leq \text{rank}(\Gamma^{N+1}),$$

and therefore we can conclude that Γ^{N+1} has a rank loop

$$\text{rank}(\Gamma^{N+1}) = \text{rank}(\Gamma^{N+1}_{xy}).$$

This means that we don't have to actually compute the sub matrices Γ^n_{xy}. We can just run the NPA hierarchy algorithm described above, compute order by order the ranks of the certificates, and in the lucky case in which we find a N such that equation (4.19) holds, we can conclude that the behavior belongs to the quantum set. In particular if $p(a, b|x, y)$ has a d-dimensional quantum representation, we will always find such $N \leq d$.

4.9 Applications of the NPA hierarchy

Apart from checking if a behavior belongs to the quantum set, the NPA hierarchy has many more applications. In this work we will focus on two of them:

- Estimating the quantum bound of Bell's inequalities;
- Finding a lower bound on the number of secure bits of randomness that can be generated from each round of a Bell scenario.

In this section we will focus on the first point, and analyze the generation of random bits in chapter 5.

Quantum bound of Bell's inequalities We saw in previous chapters that Bell's inequalities have the following form

$$S \leq I_q \implies \sum_{a,b,x,y} s^{a,b}_{x,y} p(a, b|x, y) \leq I_q,$$

where in this case I_q is the quantum bound, i.e. the maximum value of S obtainable with quantum strategies. Finding analytically the value of I_q (for fixed $s^{a,b}_{x,y}$) is in general a difficult problem: we saw that there are some techniques like the SOS decomposition, see section 3.7, but they do not always work. Luckily we can always find a numerical upper bound of I_q with the NPA hierarchy. This is possible for two main reasons:

- We want to find the maximum value of S, which is a function of the behavior;
- S is a linear function of the behavior.

In general the problem of maximizing (or minimizing) a linear function of the probability is always solvable with the NPA hierarchy. Indeed recall that at any order n, the certificate Γ^n has entries

$$\Gamma^n_{a_x, b_y} = p(a, b|x, y).$$

So, by defining the matrix

$$\beta^n_{i,j} = \begin{cases} s^{a,b}_{x,y} & \text{if } i = a_x, \quad j = b_y \\ 0 & \text{else} \end{cases}$$

we can view S, in the formalism of SDP, as

$$S = \text{Tr}[\beta^n \Gamma^n] = \sum_{a,b,x,y} s^{a,b}_{x,y} p(a, b|x, y).$$

Therefore solving the following SDP problem

$$\begin{aligned} &\text{maximize } \text{Tr}[\beta^n \Gamma^n]; \\ &\text{subject to } \text{Tr}[F^T \gamma] = g(p), \quad \forall F, g \in \mathcal{F}(\mathcal{O}); \\ &\text{subject to } \Gamma \succcurlyeq 0. \end{aligned}$$

outputs both a valid certificate Γ^n for the n-th order NPA hierarchy and an estimate I_n of I_q. Since we are finding the maximum in the set $Q_n \supseteq \mathcal{Q}$, see equation (4.16), I_n is also an upper bound of I_q, and hence by repeating at all orders we find the sequence

$$I_1 \geq I_2 \geq .... \geq I_q.$$

that also satisfies, see equation (4.17):

$$\lim_{n \to \infty} I_n = I_q.$$

Practically, due to the running time of the algorithm, it's usually only possible to compute the sequence up to I_3 and, by testing the algorithm with known Bell's inequalities, like the CHSH, it has been found that it is already a very good upper bound of I_q, even though there is no theoretical proof of it.

Chapter 5

Generating random numbers

In this chapter we will describe the possible ways to quantify the amount of randomness that can be generated in a Bell's scenario. To keep things simple, consider the particular case in which $\Delta = m = 2$. In this scenario Both Alice and Bob have two inputs that yield binary outputs

$$x, y \in \{0, 1\}; \quad a, b \in \{+1, -1\}.$$

As usual assume that their devices can be described by the theory of quantum mechanics and that, up to purification, their observables are projective and the state being measured is pure

$$\begin{aligned} &\Lambda_a^x \succcurlyeq 0, \quad \Lambda_a^x \Lambda_{a'}^x = \delta_{a,a'} \Lambda_a^x; \\ &\Pi_b^y \succcurlyeq 0, \quad \Pi_b^y \Pi_{b'}^y = \delta_{b,b'} \Pi_b^y; \\ &p(a, b|x, y) = \langle\psi| \Lambda_a^x \otimes \Pi_b^y |\psi\rangle \,, \end{aligned} \tag{5.1}$$

for more details, see chapter 3. Finally, without loss of generality, assume that Alice and Bob want to generate random numbers from the outcomes of the inputs corresponding to $x = y = 0$.

Number of bits generated per round The quantity of interest for extracting randomness is r, defined as the number of random bits generated per round of the Bell's experiment. It satisfies

$$0 \leq r \leq 2.$$

Indeed, in a single round we have two binary outcomes: the bit generated by Alice and the bit generated by Bob. If those numbers are random (as a coin flip) and completely uncorrelated we have the best case $r = 2$. We want to find an analytical expression of r in function of the probability distribution, in such a way that we can measure how good is a protocol, intended as the choice of Λ, Π and $|\psi\rangle$. We will begin by considering the trusted case.

5.1 Trusted case

This is the most simple case: trusted because we assume to know the explicit form of operators and wave function that produced the observed probability distribution (for example we trust the person that built the devices). They might not be projective, but as usual by purifying the space we find an explicit form of everything appearing in equation (5.1), that can be used for theoretical calculations. With that said, by direct generalization from classical information theory it can be proven, see [2], that a good definition for r is the Von Neumann entropy

$$r_{vn} = -\operatorname{Tr}[\rho \log_2(\rho)], \tag{5.2}$$

where ρ is the post-measurement state, defined as the quantum state after Alice and Bob measure the observables corresponding to $x = y = 0$:

$$\begin{aligned} &\rho = \sum_{a,b} |ab\rangle \langle ab| \langle\psi| \Lambda_a^0 \otimes \Pi_b^0 |\psi\rangle = \sum_{a,b} |ab\rangle \langle ab| \, p(a, b|0, 0); \\ &\text{where } |ab\rangle \equiv |a\rangle \otimes |b\rangle \,, \end{aligned} \tag{5.3}$$

and $|a\rangle$ and $|b\rangle$ are the eigenvectors of the operators being measured:

$$\begin{aligned} &A_0 = \sum_a a\Lambda_a^0, \quad B_0 = \sum_b b\Pi_b^0; \\ &A_0 |a\rangle = a |a\rangle ; \\ &B_0 |b\rangle = b |b\rangle . \end{aligned} \tag{5.4}$$

Notice in particular that ρ is a diagonal matrix, therefore the Von Neumann entropy reduces to the Shannon entropy:

$$r_{vn} = -\sum_{a,b} p(a,b|0,0)\log_2[p(a,b|0,0)]. \tag{5.5}$$

An alternative definition is the min-entropy:

$$r_{me} = -\log_2(\|\rho\|_{op}),$$

where ρ is the post-measurement state of equation (5.3) and the operator norm $\|\rho\|_{op}$ outputs the largest eigenvalue of ρ. Since in our case ρ is diagonal, the min-entropy reduces to

$$\begin{aligned} &r_{me} = -\log_2(p_{\max}); \\ &\text{where } p_{\max} \equiv \max_{a,b}[p(a,b|0,0)]. \end{aligned} \tag{5.6}$$

Notice that the Von Neumann entropy is never smaller than the min-entropy

$$r_{vn} \geq r_{me},$$

because the former averages among all probabilities $p(a,b|0,0)$, while the latter considers only the worst case p_{max}. So in the trusted case we should always stick to the Von Neumann entropy, however, as we will see later, the min-entropy has its advantages in the device-independent scenario:

- Computing the min-entropy numerically (NPA hierarchy) is always much faster;
- From a theoretical point of view, maximizing the guessing probability to compute r_{me} is usually easier than finding the inf of equation (5.9) to calculate r_{vn}.

5.2 Device-independent case

In the device independent case things are more complex because we cannot trust neither the devices nor the wave function. To extract randomness, in addition to what we assumed in chapter 3, we include the possibility of an adversary Eve, who has the role of trying to guess the outcomes of Alice and Bob. More precisely Eve has access to a part of the wavefunction $|\psi\rangle$, which now becomes a tripartite system

$$|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_E \otimes \mathcal{H}_P \tag{5.7}$$

where $\mathcal{H}_A \otimes \mathcal{H}_B$ is the Alice and Bob Hilbert space, $\mathcal{H}_E$ is the Eve space and $\mathcal{H}_P$ is the purification space. Actually, in this case it's convenient to trace out the purification space and work with the mixed density matrix:

$$\rho_0 = \mathrm{Tr}_P[|\psi\rangle\langle\psi|].$$

As before Alice and Bob can measure only on their sub-system (they do not even know of the presence of Eve)

$$\Lambda_a^x : \mathcal{H}_A \to \mathcal{H}_A, \quad \Pi_b^y : \mathcal{H}_B \to \mathcal{H}_B.$$

Similarly Eve has an hermitian operator E, that acts only on her subspace $\mathcal{H}_E$ and that can be written with projector formalism as

$$E = \sum_{e_1,e_2} e_1 e_2 E_{e_1 e_2},$$

where, as usual, the projectors satisfy

$$\begin{aligned}&E_{e_1e_2} : \mathcal{H}_E \to \mathcal{H}_E;\\ &E_{e_1e_2} \succcurlyeq 0;\\ &\sum_{e_1,e_2} E_{e_1e_2} = \mathbb{I}_E;\\ &E_{e_1e_2}E_{e_3e_4} = \delta_{e_1e_3}\delta_{e_2e_4}E_{e_1e_2}.\end{aligned} \tag{5.8}$$

So with the presence of Eve the Bell's scenario becomes

- Alice and Bob choose their input, perform their measurement and get outputs a, b;
- If their input was $x = y = 0$, then Eve perform her measurement and finds a two-bits output e_1, e_2. Her hope is that the probability of having

  $$e_1 = a; \quad e_2 = b,$$

  is very high, in such a way that she can likely guess the random numbers generated by Alice and Bob.

As in the trusted case, to define r we need an expression for the post-measurement state ρ (after Alice and Bob measured their input $x = y = 0$):

$$\rho = \sum_{a,b} |ab\rangle\langle ab| \operatorname{Tr}_{A,B}\left[\rho_0 \Lambda_a^0 \otimes \Pi_b^0 \otimes \mathbb{I}_E\right],$$

where $|ab\rangle$ is the same of equation (5.4) and $\operatorname{Tr}_{A,B}[]$ is the partial trace operator over Alice and Bob subspace $\mathcal{H}_A \otimes \mathcal{H}_B$.

Von Neumann entropy The number of bits generated per round with the Von Neumann entropy is

$$\begin{aligned}&r_{vn} = \inf_{\rho_0,\Lambda,\Pi} H(AB|E)_\rho;\\ &\text{inf among } \rho_0, \Lambda, \Pi \text{ satisfying } \operatorname{Tr}_{A,B,E}[\rho_0 \Lambda_a^x \otimes \Pi_b^y \otimes \mathbb{I}_E] = p(a, b|x, y),\end{aligned} \tag{5.9}$$

where $H(AB|E)_\rho$ is the conditional Von Neumann entropy (conditioned on Eve space) for the post-measurement state ρ. Such definition is a direct generalization of the corresponding one for the trusted case, see equation (5.2), indeed:

- We are in a device-independent scenario, so we find the minimum among all the state and measurements compatible with the probability distribution (we cannot assume their form as in the trusted case);
- Due to the presence of Eve, we have to use the conditional Von Neumann entropy instead of the unconditional one.

A special case Computing analytically equation (5.9) is, in general, very difficult. However, there are some cases in which it can be simplified: for example if we manage to prove that the Eve's part of the state is uncorrelated from Alice and Bob

$$\rho_0 = \rho_{AB} \otimes \rho_E,$$

then, the same result holds for the post-measurement state

$$\rho = \left(\sum_{a,b} |ab\rangle\langle ab| \operatorname{Tr}_{A,B}\left[\rho_{AB}\Lambda_a^0 \otimes \Pi_b^0\right]\right) \otimes \rho_E \equiv \widetilde{\rho}_{AB} \otimes \rho_E,$$

and in this case it can be proven that the conditional Von Neumann entropy reduces to the unconditional one:

$$r_{vn} = \inf_{\rho_0,\Lambda,\Pi} - \mathrm{Tr}[\widetilde{\rho}_{AB} \log_2 \widetilde{\rho}_{AB}];$$

$$\text{inf among } \rho_0, \Lambda, \Pi \text{ satisfying } \mathrm{Tr}_{A,B,E}[\rho_0 \Lambda_a^x \otimes \Pi_b^y \otimes \mathbb{I}_E] = p(a,b|x,y).$$

Finally notice that

$$\widetilde{\rho}_{AB} = \sum_{a,b} |ab\rangle \langle ab| \, \mathrm{Tr}_{A,B} \left[\rho_{AB} \Lambda_a^0 \otimes \Pi_b^0\right],$$

is a diagonal matrix. This implies that the Von Neumann entropy reduces to a Shannon entropy

$$r_{vn} = \inf_{\rho_0,\Lambda,\Pi} - \sum_{a,b} \mathrm{Tr}_{A,B} \left[\rho_{AB} \Lambda_a^0 \otimes \Pi_b^0\right] \log_2 \left\{\mathrm{Tr}_{A,B} \left[\rho_{AB} \Lambda_a^0 \otimes \Pi_b^0\right]\right\},$$

but at the same time

$$\mathrm{Tr}_{A,B} \left[\rho_{AB} \Lambda_a^0 \otimes \Pi_b^0\right] = p(a,b|0,0).$$

Hence the result doesn't depend on the choice of state and observables ρ_0, Λ, Π and we end up with:

$$r_{vn} = - \sum_{a,b} p(a,b|0,0) \log_2 [p(a,b|0,0)].$$

Min-entropy Similarly, equation (5.6) can be generalized to the device independent case. It is found that

$$r_{me} = - \log_2 (G),$$

where G is called guessing probability:

$$G \equiv \max_{\rho_0,\Lambda,\Pi,E} \left(\sum_{a,b} \mathrm{Tr}_{A,B,E}[\rho_0 \Lambda_a^0 \otimes \Pi_b^0 \otimes E_{a,b}] \right) \tag{5.10}$$
$$\text{max among } \rho_0, \Lambda, \Pi, E \text{ satisfying } \mathrm{Tr}_{A,B,E}[\rho_0 \Lambda_a^x \otimes \Pi_b^y \otimes \mathbb{I}_E] = p(a,b|x,y).$$

G represents the maximum probability that Eve has to guess Alice and Bob outcomes, among all possible states and operators compatible with $p(a,b|x,y)$.

A special case As for the Von Neumann entropy, in the special case in which the Eve's part of the state is uncorrelated from Alice and Bob

$$\rho_0 = \rho_{AB} \otimes \rho_E$$

we can find a simplified expression for the guessing probability:

$$\begin{aligned} G = \max_{\rho_0,\Lambda,\Pi,E} \left(\sum_{a,b} \mathrm{Tr}_{A,B}[\rho_{AB} \Lambda_a^0 \otimes \Pi_b^0] \, \mathrm{Tr}[\rho_E E_{a,b}] \right) = \max_{\rho_0,\Lambda,\Pi,E} \left(\sum_{a,b} p(a,b|0,0) \, \mathrm{Tr}[\rho_E E_{a,b}] \right) = \\ = \max_{E,\rho_E} \left(\sum_{a,b} p(a,b|0,0) \, \mathrm{Tr}[\rho_E E_{a,b}] \right). \end{aligned} \tag{5.11}$$

where we used that

$$\mathrm{Tr}_{A,B}[\rho_{AB} \Lambda_a^0] = p(a,b|0,0),$$

and therefore there is no need to maximize on state and measurements ρ_0, Π, Λ. We still need to find the maximum among all possible choices of Eve's state and observables ρ_E, E. To do this define

$$p_{max} \equiv \max_{a,b} p(a,b|x,y) = p(\widetilde{a}, \widetilde{b}, 0, 0)$$

and notice that equation (5.11) can be bounded from above

$$G = \max_{E,\rho_E}\left(\sum_{a,b} p(a,b|0,0)\,\mathrm{Tr}[\rho_E E_{a,b}]\right) \leq p_{max}\cdot\max_{E,\rho_E}\left(\mathrm{Tr}\left[\rho_E \sum_{a,b} E_{a,b}\right]\right) = p_{max}\cdot\mathrm{Tr}\left[\rho_E\right] = p_{max}$$

where we used equation (5.8), the linearity of the trace and the normalization of the density matrix. Finally note that the bound can actually be achieved by choosing E and ρ_E in such a way that

$$\mathrm{Tr}[\rho_E E_{a,b}] = \begin{cases} 1 & \text{if } a = \widetilde{a} \text{ and } b = \widetilde{b} \\ 0 & \text{else} \end{cases}$$

For such choice is easy to see that

$$G = p_{max},$$

and therefore the min entropy is

$$r_{me} = -\log_2(p_{max}).$$

Comparison between the two entropies The two definitions of r, with the Von Neumann entropy and the min entropy, are used in different contexts. Indeed, as discussed in [13], r_{vn} is the most suited in the limit of many repetitions of Bell's experiment, r_{me} instead is more appropriate in the "single shoot" case, in which the experiment is repeated a few times.

5.3 Numerical simulations

In this section we will see how the NPA hierarchy algorithm can be adapted to find numerical lower bounds on r_{me} and r_{vn} in the device-independent case. Having an algorithm to perform this task is very helpful for two main reasons:

- Studying the effects of the noise on the system: usually noised systems don't maximally violate any Bell's inequality, so performing self-testing and finding analytical values for r is often impossible. From numerical simulations instead, we get regardless reliable lower bounds.
- Help in theoretical proofs: let's say that we found a trusted protocol, with generic probability distribution $p(a,b|x,y)$, that generates many random bits per round $r_{\text{trust}} \approx 2$, and we want to determine if it's possible to extend the result to the device-independent case. In other words, by assuming only the observed behavior $p(a,b|x,y)$ and the possibility of an adversary Eve, can we still generate $r = r_{\text{trust}}$?. This is usually a difficult task, but if from the simulation we find a lower bound $r_{\text{sim}} \approx r_{\text{trust}}$ then the answer is likely yes.

Numerical algorithm for r_{me} Let us begin by the min-entropy r_{me}, defined as

$$r_{me} = -\log_2(G).$$

The guessing probability G is defined in function of the density matrix ρ, in equation (5.10). To translate the problem in the NPA formalism is convenient to work with the purified weave function (5.7), so that equation (5.10) can be rewritten as

$$\begin{aligned} &G \equiv \max_{|\psi\rangle,\Lambda,\Pi,E}\left(\sum_{a,b}\langle\psi|\,\Lambda_a^0\otimes\Pi_b^0\otimes E_{a,b}]\,|\psi\rangle\right)\\ &\text{such that } \langle\psi|\,\Lambda_a^x\otimes\Pi_b^y\otimes\mathbb{I}_E\,|\psi\rangle = p(a,b|x,y);\\ &\text{such that } E_{e_1e_2}\succcurlyeq 0;\quad \sum_{e_1,e_2}E_{e_1e_2}=\mathbb{I}_E;\quad E_{e_1e_2}E_{e_3e_4}=\delta_{e_1e_3}\delta_{e_2e_4}E_{e_1e_2};\quad E_{e_1e_2}^{\dagger}=E_{e_1e_2};\\ &\text{such that } \Lambda_a^x\succcurlyeq 0;\quad \sum_a\Lambda_a^x=\mathbb{I}_A;\quad \Lambda_{a_1}^x\Lambda_{a_2}^x=\delta_{a_1a_2}\Lambda_{a_1}^x;\quad (\Lambda_a^x)^{\dagger}=\Lambda_a^x;\\ &\text{such that } \Pi_b^y\succcurlyeq 0;\quad \sum_b\Pi_b^y=\mathbb{I}_B;\quad \Pi_{b_1}^y\Pi_{b_2}^y=\delta_{b_1b_2}\Pi_{b_1}^y;\quad (\Pi_b^y)^{\dagger}=\Pi_b^y, \end{aligned} \tag{5.12}$$

where we also added explicitly all the constraints on the operators and we omitted the identity on the purification space $\mathbb{I}_p$. We have three differences with the NPA formalism that we have to address:

1. In the original NPA hierarchy we didn't consider any Eve's operator E_{e_1,e_2}, see equation (4.9);
2. We are currently working on the quantum set $\mathcal{Q}'$: Alice and Bob (and Eve) projectors act on different Hilbert spaces and we consider their tensor product. In the NPA hierarchy instead we considered the set $\mathcal{Q}$, where Alice and Bob projectors satisfy the weaker condition of just commuting, see equation (4.5);
3. We are considering the completeness relations
$$\sum_{e_1,e_2} E_{e_1e_2} = \mathbb{I}_E; \quad \sum_a \Lambda_a^x = \mathbb{I}_A; \quad \sum_b \Pi_b^y = \mathbb{I}_B,$$
that, as we saw in section (4.2), generates some redundancy.

All those issues are easily addressed: first of all we get rid of the redundancy by defining the reduced sets, where for each x we remove an Alice's operator $\Lambda_a^{a_x}$, for each y we remove a Bob's operator $\Pi_b^{b_y}$, and we also remove an Eve's operator $E_{\widetilde{e}_1,\widetilde{e}_2}$.

$$\begin{aligned}\widetilde{\Lambda} &= \{\Lambda_a^x,\ \forall a,x\} - \{\Lambda_{a_x}^x,\ \forall x\};\\ \widetilde{\Pi} &= \{\Pi_b^y,\ \forall b,y\} - \{\Pi_{b_y}^y,\ \forall y\};\\ \widetilde{E} &= \{E_{e_1e_2},\ \forall e_1,e_2\} - \{E_{\widetilde{e}_1\widetilde{e}_2}\}.\end{aligned}$$

Since, analogously to what we did in section (4.2), they are a linear combination of the others

$$\begin{aligned}\Lambda_{a_x}^x &= \mathbb{I}_A - \sum_{a\neq a_x} \Lambda_a^x, \quad \forall x;\\ \Pi_{b_y}^y &= \mathbb{I}_B - \sum_{b\neq b_y} \Pi_b^y, \quad \forall y;\\ E_{\widetilde{e}_1\widetilde{e}_2} &= \mathbb{I}_E - \sum_{e_1\neq\widetilde{e}_1, e_2\neq\widetilde{e}_2} E_{e_1,e_2}.\end{aligned} \tag{5.13}$$

Equation (5.13) also needs to be substituted in equation (5.12), so that we have an expression of G that depends only on the reduced set of projectors. The next step is performing a relaxation by considering behaviors belonging to the quantum set $\mathcal{Q}$ instead of $\mathcal{Q}'$. Practically we only have to substitute the tensor product of local operators with product of commuting operators, in particular equation (5.12) becomes:

$$\begin{aligned}G \equiv \max_{|\psi\rangle,\Lambda,\Pi,E} &\left(\sum_{a,b} \langle\psi|\, \Lambda_a^0\Pi_b^0 E_{a,b}]\, |\psi\rangle\right)\\ \text{such that } &\langle\psi|\, \Lambda_a^x\Pi_b^y\, |\psi\rangle = p(a,b|x,y);\\ \text{such that } &[\Lambda_a^x,\Pi_b^y] = [\Lambda_a^x, E_{e_1e_2}] = [\Pi_b^y, E_{e_1e_2}] = 0;\\ &\ldots\end{aligned}$$

Performing this relaxation is not a problem: in chapter 2 we showed that

$$\mathcal{Q}' \subseteq \mathcal{Q},$$

therefore we are effectively finding G as the maximum in a larger set, which corresponds to a smaller lower bound on r_{me}. Finally, to take in account Eve's operator we simply add them to the set of sequences, equation (4.9):

$$\begin{aligned}S_0 &= \{\mathbb{I}\}\\ S_1 &= S_0 \cup \{\Lambda_i \in \widetilde{\Lambda}\} \cup \{\Pi_i \in \widetilde{\Pi}\} \cup \{E_i \in \widetilde{E}\}\\ S_2 &= S_0 \cup S_1 \cup \{\Lambda_i\Lambda_j \in \widetilde{\Lambda}\} \cup \{\Pi_i\Pi_j \in \widetilde{\Lambda}\} \cup \{E_iE_j \in \widetilde{E}\} \cup \{\Lambda_i\Pi_j,\ \Lambda_i \in \widetilde{\Lambda},\ \Pi_j \in \Pi\} \cup \ldots\\ S_3 &= \ldots\end{aligned} \tag{5.14}$$

this has the only effects of having larger certificates Γ^n at each level n. So overall, with these changes the maximization problems for G, equation (5.12), becomes:

$$\begin{aligned} &G \equiv \max_{|\psi\rangle,\widetilde{\Lambda},\widetilde{\Pi},\widetilde{E}} \left(\sum_{a,b} \langle\psi|\,\Lambda_a^0\Pi_b^0 E_{a,b}]\,|\psi\rangle \right) \\ &\text{such that } \langle\psi|\,\Lambda_a^x\Pi_b^y\,|\psi\rangle = p(a,b|x,y); \\ &\text{such that } E_{e_1e_2} \succcurlyeq 0; \quad E_{e_1e_2}E_{e_3e_4} = \delta_{e_1e_3}\delta_{e_2e_4}E_{e_1e_2}; \quad E^\dagger_{e_1e_2} = E_{e_1e_2}; \\ &\text{such that } \Lambda_a^x \succcurlyeq 0; \quad \Lambda_{a_1}^x\Lambda_{a_2}^x = \delta_{a_1a_2}\Lambda_{a_1}^x; \quad (\Lambda_a^x)^\dagger = \Lambda_a^x; \\ &\text{such that } \Pi_b^y \succcurlyeq 0; \quad \Pi_{b_1}^y\Pi_{b_2}^y = \delta_{b_1b_2}\Pi_{b_1}^y; \quad (\Pi_b^y)^\dagger = \Pi_b^y, \\ &\text{such that } [\Lambda_a^x,\Pi_b^y] = [\Lambda_a^x,E_{e_1e_2}] = [\Pi_b^y,E_{e_1e_2}] = 0. \end{aligned} \tag{5.15}$$

Which is solvable with the NPA hierarchy: indeed let Γ^n be a certificate of order $n > 1$ and let i_a, j_b, be its row and column indexes such that

$$(\Gamma^n)_{i_a,j_a} = \langle\psi|\,\Lambda_a^0\Pi_b^0 E_{a,b}]\,|\psi\rangle\,, \quad \forall a,b \in \{-1,+1\}. \tag{5.16}$$

If we define the matrix β^n as

$$\beta^n_{i,j} = \begin{cases} 1 & \text{if } i = i_a, \quad j = j_b, \quad \forall a,b \in \{-1,+1\} \\ 0 & \text{else} \end{cases}$$

Then clearly

$$\sum_{a,b} \langle\psi|\,\Lambda_a^0\Pi_b^0 E_{a,b}]\,|\psi\rangle = \mathrm{Tr}[\Gamma^n\beta^n],$$

and solving the following SDP problem

$$\begin{aligned} &\text{maximize } \mathrm{Tr}[\beta^n\Gamma^n]; \\ &\text{subject to } \mathrm{Tr}[F^T\Gamma^n] = g(p), \quad \forall F,g \in \mathcal{F}(\mathcal{O}); \\ &\text{subject to } \Gamma^n \succcurlyeq 0. \end{aligned}$$

outputs both a valid certificate Γ^n for the n-th order NPA hierarchy and an estimate G_n of G as defined in equation (5.15). As usual, solving the problem at any order of the NPA hierarchy will yield a sequence of better and better approximations

$$G_2 \geq G_3 \geq \ldots \geq G,$$

that also satisfies

$$\lim_{n\to\infty} G_n = G.$$

Furthermore, by testing the program with known protocols, it is found that most of the time G_2 is already a very good approximation of G. Finally, notice that we assumed $n > 1$, because the first order matrix Γ^1 doesn't contain the correlations of equation (5.16), therefore we cannot define G_1. In cases where solving the second order is already computationally infeasible, the best thing to do is considering the intermediate level $1 + AB$, as defined in equation (4.18).

Numerical algorithm for r_{vn} For the Von Neumann entropy things are more complex, and converting equation (5.2) to a NPA hierarchy problem is not easy. Here we will just show the final result, for a proof see [4]:

Let $m \in \mathbb{N}$ and let t_i, w_i be the nodes and weight of a m-point Gauss-Radau rule with $t_m = 1$. Then, in a device independent scenario, r_{vn} is bounded below by

$$\begin{aligned} & c_m + \inf \sum_{i=1}^{m-1} \frac{w_i}{t_i \ln 2} \sum_{a,b} \langle\psi| \Lambda_a^x \Pi_b^y (Z_{ab,i} + Z_{ab,i}^\dagger + (1-t_i) Z_{ab,i}^\dagger Z_{ab,i}) + t_i Z_{ab,i} Z_{ab,i}^\dagger |\psi\rangle \\ & \text{such that } \langle\psi| \Lambda_a^x \Pi_b^y |\psi\rangle = p(a,b|x,y); \\ & \text{such that } Z_{ab,i} Z_{ab,i}^\dagger \leq \alpha_i; \\ & \text{such that } \Lambda_a^x \succcurlyeq 0; \quad \Lambda_{a_1}^x \Lambda_{a_2}^x = \delta_{a_1 a_2} \Lambda_{a_1}^x; \quad (\Lambda_a^x)^\dagger = \Lambda_a^x; \\ & \text{such that } \Pi_b^y \succcurlyeq 0; \quad \Pi_{b_1}^y \Pi_{b_2}^y = \delta_{b_1 b_2} \Pi_{b_1}^y; \quad (\Pi_b^y)^\dagger = \Pi_b^y, \\ & \text{such that } [\Lambda_a^x, \Pi_b^y] = [\Lambda_a^x, Z_{cb,i}] = [\Pi_b^y, Z_{ac,i}] = [\Lambda_a^x, Z_{cb,i}^\dagger] = [\Pi_b^y, Z_{ac,i}^\dagger] = 0 \\ & \text{where } c_m = \sum_{i=1}^{m-1} \frac{w_i}{t_i \ln 2}. \end{aligned} \tag{5.17}$$

Overall it is quite similar to the result we found for the min-entropy, see equation (5.15). The main differences are:

- Eve's operator $Z_{ab,i}$ are no longer projective. They are only bounded linear operators (not even hermitian). To reduce the running time of simulations, the bounded condition

  $$Z_{ab,i} Z_{ab,i}^\dagger \leq \alpha_i \tag{5.18}$$

  can be neglected, with the trade off of computing a worse estimate of r_{vn};

- The goodness of the approximation will depend on two variables: the NPA hierarchy level n and the number of points m of the Gauss Radau quadrature. To converge to the real value of r_{vn} we have to consider the limit $(n, m) \to \infty$;

- The number of Eve's operator is

  $$2\Delta^2 (m-1) = 8(m-1),$$

  where the 2 factor is there because for each Z we have to consider its hermitian conjugate $Z^\dagger$. In the min-entropy case instead Eve's had only $\Delta^2 = 4$ operators, so computing r_{vn} has a much larger running time.

- We have the addition of t_i and w_i coefficients, but at the end of the day they are just real numbers that can be numerically computed with existing libraries.

The optimization problem of equation (5.17) can be converted to a SDP, solvable with the NPA hierarchy algorithm, by following the same steps we did for the min entropy. Notice that the computational complexity can be reduced by considering the following inequality

$$\begin{aligned} & \inf \sum_{i=1}^{m-1} \frac{w_i}{t_i \ln 2} \sum_{a,b} \langle\psi| \Lambda_a^x \Pi_b^y (Z_{ab,i} + Z_{ab,i}^\dagger + (1-t_i) Z_{ab,i}^\dagger Z_{ab,i}) + t_i Z_{ab,i} Z_{ab,i}^\dagger |\psi\rangle \geq \\ & \sum_{i=1}^{m-1} \inf \left(\frac{w_i}{t_i \ln 2} \sum_{a,b} \langle\psi| \Lambda_a^x \Pi_b^y (Z_{ab} + Z_{ab}^\dagger + (1-t_i) Z_{ab}^\dagger Z_{ab}) + t_i Z_{ab} Z_{ab}^\dagger |\psi\rangle \right), \end{aligned} \tag{5.19}$$

where we have swapped the outer sum with the inf. The main advantage of solving this new problem is that instead of running a single computationally heavy SDP, we run $(m-1)$ smaller SDP, where Eve has only $2\Delta^2 = 8$ operators, and sum their results. In this way the running time scales linearly in the number of nodes of the Gauss Radau quadrature and we still find a lower bound of r_{vn}. The only drawback is that there is no guarantee on the goodness of the result, which could be much worse than the one found by solving the original minimization (5.17).

Chapter 6

Sequential extensions

In the previous chapter we discussed the generation of random numbers from a Bell's scenario with two inputs and outputs, and we found that a big limitation is the impossibility of generating more than 2 bits per round

$$r < 2.$$

This limit follows from the fact that in a single round only two measurements happens and each of them yields a 1-bit result. A natural way to overcome this restriction is adding more users, that will sequentially measure the state. Protocols that follow such generalization are called ”sequential Bell's scenarios”.

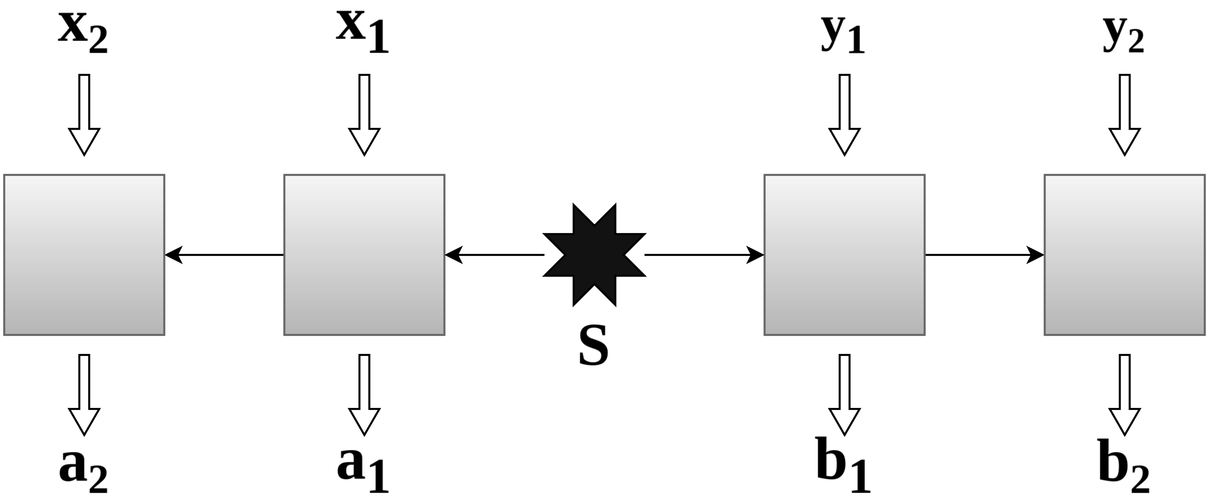

Figure 6.1: Sequential Bell's scenario

6.1 Sequential Bell's scenarios

Consider two groups of spatially separated observers that perform measurements on a shared physical system, generated by a source S. Each group is made respectively by N and M users that we'll label as Alice-1, Alice-2, ..., Alice-N and Bob-1, Bob-2,..., Bob-M. Then, we define round of sequential Bell's scenario to be made of the following steps:

- The physical system is generated and sent to Alice-1 and Bob-1;
- Alice-1 and Bob-1 respectively choose operators x_1, y_1, measure their part of the system, obtain an output a_1, b_1 and send the post-measurement state to Alice-2 and Bob-2;
- The process is iterated until all Alice and Bob performed a measurement.

See figure 6.1 for a representation of the case $N = M = 2$. After many rounds the probability distribution

$$\begin{aligned} & p(\vec{a},\vec{b}|\vec{x},\vec{y}), \\ & \text{where } \vec{a} \equiv (a_1,...,a_N); \quad \vec{b} \equiv (b_1,...,b_M); \\ & \text{where } \vec{x} \equiv (x_1,...,x_N); \quad \vec{y} \equiv (y_1,...,y_M); \end{aligned}$$

can be built, which represents the probability of measuring the $N+M$ outputs $\vec{a}$ and $\vec{b}$, for fixed choice of the inputs $\vec{x}$ and $\vec{y}$. As in the original Bell's scenario $p(\vec{a},\vec{b}|\vec{x},\vec{y})$ is also called behavior and satisfies the following constraints

$$\begin{aligned} & p(\vec{a},\vec{b}|\vec{x},\vec{y}) \geq 0, \quad \text{(positivity constraint);} \\ & \sum_{\vec{a},\vec{b}} p(\vec{a},\vec{b}|\vec{x},\vec{y}) = 1, \quad \text{(normalization constraint);} \\ & \sum_{b_k,...,b_M} p(\vec{a},\vec{b}|\vec{x},\vec{y}) = \sum_{b_k,...,b_M} p(\vec{a},\vec{b}|\vec{x},\vec{y}') \quad \begin{array}{l} \forall \vec{x},\vec{a}; \\ \forall k = 2,...,M; \\ \forall b_1,...b_{k-1}; \qquad \text{(sequential constraint).} \\ \forall \vec{y},\vec{y}', \text{ such that } y_j = y_j', \ j = 1,...,k-1, \end{array} \end{aligned} \tag{6.1}$$

The last condition is called sequential constraint, and it means that the input chosen by the last $(M-k+1)$ Bobs cannot influence the first $(k-1)$ Bobs. Clearly an analogous constraint holds on Alice's side.

6.2 Sequential Quantum Set

In the quantum case we can assume that, up to space purification, the physical system is described by a pure wavefunction $|\psi\rangle$ and each Alice and Bob are doing generalized measurements described by sets of Kraus operators

$$\begin{aligned} & \{\Lambda^{x_i}_{a_i,\mu_i}\}; \quad \{\Pi^{y_j}_{b_j,\mu_j}\}; \quad i = 1,...,N \text{ and } j = 1,...,M; \\ & \text{with } \sum_{a_i,\mu_i} (\Lambda^{x_i}_{a_i,\mu_i})^\dagger \Lambda^{x_i}_{a_i,\mu_i} = \mathbb{I}; \\ & \text{with } \sum_{b_j,\mu_j} (\Pi^{y_j}_{b_j,\mu_j})^\dagger \Pi^{y_j}_{b_j,\mu_j} = \mathbb{I}. \end{aligned}$$

Where the index μ considers the case in which there are multiple Kraus operators associated to the same outcome measurement a or b. Then, the (non-normalized) post-measurement state found after Alice-1 obtains outcome a_1 from input x_1 is

$$\rho_{a_1|x_1} = \sum_{\mu_1} \Lambda^{x_1}_{a_1,\mu_1} |\psi\rangle \langle\psi| (\Lambda^{x_1}_{a_1,\mu_1})^\dagger,$$

and the corresponding probability is

$$p(a_1|x_1) = \text{Tr}[\rho_{a_1|x_1}].$$

Continuing this process for the entire sequence with inputs $\vec{x}$ and outputs $\vec{a}$, we find that

$$p(\vec{a}|\vec{x}) = \langle\psi| \Lambda^{\vec{x}}_{\vec{a}} |\psi\rangle ,$$

where

$$\Lambda^{\vec{x}}_{\vec{a}} \equiv \sum_{\mu_1,...,\mu_n} (\Lambda^{x_1}_{a_1,\mu_1})^\dagger ... (\Lambda^{x_n}_{a_n,\mu_n})^\dagger \Lambda^{x_n}_{a_n,\mu_n} ... \Lambda^{x_1}_{a_1,\mu_1}. \tag{6.2}$$

The same results holds for the sequence of Bobs

$$p(\vec{b}|\vec{y}) = \langle\psi| \Pi^{\vec{y}}_{\vec{b}} |\psi\rangle ,$$

where

$$\Pi_{\vec{b}}^{\vec{y}} \equiv \sum_{\mu_1,...,\mu_n} (\Pi_{b_1,\mu_1}^{y_1})^{\dagger}...(\Pi_{b_n,\mu_n}^{y_n})^{\dagger}\Pi_{b_n,\mu_n}^{y_n}...\Pi_{b_1,\mu_1}^{y_1}. \tag{6.3}$$

Therefore, we can define the set of sequential quantum behaviors Q'_{SEQ}, as the set of probability distributions $p(\vec{a},\vec{b}|\vec{x},\vec{y})$ of the form

$$p(\vec{a},\vec{b}|\vec{x},\vec{y}) = \langle\psi|\, \Lambda_{\vec{a}}^{\vec{x}} \otimes \Pi_{\vec{b}}^{\vec{y}} \,|\psi\rangle \iff p(\vec{a},\vec{b}|\vec{x},\vec{y}) \in Q'_{\text{SEQ}},$$

where $\Lambda_{\vec{a}}^{\vec{x}}$ and $\Pi_{\vec{b}}^{\vec{y}}$ are operators that can be written as in equations (6.2) and (6.3). Furthermore, we can define a larger set Q_{SEQ} in which, instead of using the tensor product of local operators, we use global and commuting operators:

$$p(\vec{a},\vec{b}|\vec{x},\vec{y}) \in Q_{\text{SEQ}} \iff \begin{cases} p(\vec{a},\vec{b}|\vec{x},\vec{y}) = \langle\psi|\, \Lambda_{\vec{a}}^{\vec{x}}\Pi_{\vec{b}}^{\vec{y}} \,|\psi\rangle\,; \\ [\Lambda_{\vec{a}}^{\vec{x}}, \Pi_{\vec{b}}^{\vec{y}}] = 0. \end{cases},$$

As we have already seen, this definition is more natural for performing numerical simulations with the NPA hierarchy.

Why generalized measurements In a sequential protocol to fully exploit the power of quantum mechanics, we need a way to preserve the entanglement of the wave function throughout the whole measurement round. In other words, the post-measurements states received by Alice-1,...Alice-N and Bob-1,...,Bob-M have to be entangled. Without entanglement the protocol would reduce to a trivial classical system, from which extracting random bits is much more difficult. Unfortunately projective measurements (in the non-purified space) often lead to entanglement loss: for example if Bob-1 measures $\sigma_z \otimes \mathbb{I}$ on the Bell's state $|\phi_+\rangle$, it collapses to

$$|\phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \to \begin{cases} \text{either} & |00\rangle\,; \\ \text{or} & |11\rangle\,, \end{cases}$$

and the next user would receive a useless classical mixture

$$\rho = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|).$$

The only solution to this problem is to use generalized operator. Let's make some examples that we will use in protocols of the next chapter.

An example of Kraus operators As an example of generalized measurements consider the simple 1-qubit case with $N = \mu = 1$ and binary outputs. Such system is described only by two Kraus operators $\Lambda_{+1,1}^{1}$, $\Lambda_{-1,1}^{1}$, that we define as:

$$\Lambda_{+1,1}^{1} = \cos\theta\,|0\rangle\langle 0| + \sin\theta\,|1\rangle\langle 1|\,;$$
$$\Lambda_{-1,1}^{1} = \cos\theta\,|1\rangle\langle 1| + \sin\theta\,|0\rangle\langle 0|\,,$$

where $(|0\rangle, |1\rangle)$ is a basis of the 1-qubit Hilbert space and $\theta \in [0, \frac{\pi}{4}]$ is a fixed parameter. Then, following equation (6.2), we can compute the observables $\Lambda_{\pm 1}^{1}$:

$$\Lambda_{+1}^{1} = \cos^2(\theta)\,|0\rangle\langle 0| + \sin^2(\theta)\,|1\rangle\langle 1|\,;$$
$$\Lambda_{-1}^{1} = \cos^2(\theta)\,|1\rangle\langle 1| + \sin^2(\theta)\,|0\rangle\langle 0|\,.$$

Such generalized operator can be seen as a weak version of σ_z: indeed for $\theta = 0$ we recover the two projectors of σ_z

$$\Lambda_{+1}^{1} = |0\rangle\langle 0|\,; \quad \Lambda_{-1}^{1} = |1\rangle\langle 1|\,,$$

so we are performing a projective measurement. Instead, in the other limit $\theta = \frac{\pi}{4}$ it reduces to the non-interactive measurement

$$\Lambda^1_{+1} = \Lambda^1_{-1} = \frac{1}{2}\mathbb{1}.$$

More generally, given any 1-qubit unitary operator O, we define its weak version $w(O,\theta)$ as the operator made by the two Kraus

$$\begin{aligned}
\Lambda^1_{+1,1} &= \cos\theta\left(\frac{O+\mathbb{1}}{2}\right) + \sin\theta\left(\frac{-O+\mathbb{1}}{2}\right);\\
\Lambda^1_{-1,1} &= \cos\theta\left(\frac{-O+\mathbb{1}}{2}\right) + \sin\theta\left(\frac{O+\mathbb{1}}{2}\right),
\end{aligned}$$

and, again, from equation (6.2), we derive that

$$\begin{aligned}
\Lambda^1_{+1} &= \cos^2(\theta)\left(\frac{\mathbb{1}+O}{2}\right) + \sin^2(\theta)\left(\frac{\mathbb{1}-O}{2}\right);\\
\Lambda^1_{-1} &= \cos^2(\theta)\left(\frac{\mathbb{1}-O}{2}\right) + \sin^2(\theta)\left(\frac{\mathbb{1}+O}{2}\right).
\end{aligned}$$

6.3 Stinespring dilation

So far we considered a purified space in which the state $|\psi\rangle$ is pure but all operators are generalized. We managed to characterize any sequential quantum state

$$p(\vec{a},\vec{b}|\vec{x},\vec{y}) = \langle\psi|\,\Lambda^{\vec{x}}_{\vec{a}}\otimes\Pi^{\vec{y}}_{\vec{b}}\,|\psi\rangle\,,$$

but equations (6.2) and (6.3), that define $\Lambda^{\vec{x}}_{\vec{a}}$ and $\Pi^{\vec{y}}_{\vec{b}}$, are quite messy and it's not clear how to work with them. As we already discussed in section 3.1, the space can be further expanded with a technique called Stinespring dilation, and we can find a realization where $\Lambda^{\vec{x}}_{\vec{a}}$ and $\Pi^{\vec{y}}_{\vec{b}}$ are projective operators. More precisely:

A given behavior $p(\vec{a},\vec{b}|\vec{x},\vec{y})$ belongs to Q'_{SEQ} if and only if it can be realized as

$$p(\vec{a},\vec{b}|\vec{x},\vec{y}) = \langle\psi|\,\Lambda^{\vec{x}}_{\vec{a}}\otimes\Pi^{\vec{y}}_{\vec{b}}\,|\psi\rangle\,,$$

with the measurements operators being projective and satisfying the one-way 'no-signaling' and orthogonality condition. That is

$$\begin{aligned}
&\Lambda^{\vec{x}}_{\vec{a}}\Lambda^{\vec{x}}_{\vec{a}'} = \delta_{\vec{a},\vec{a}'}\Lambda^{\vec{x}}_{\vec{a}};\\
&\sum_{a_{k+1},\dots,a_n}\left(\Lambda^{\vec{x}}_{\vec{a}} - \Lambda^{\vec{x}'}_{\vec{a}}\right) = 0, \quad \begin{array}{l}\forall a_1,\dots,a_k\\ \forall \vec{x},\vec{x}' \text{ such that } x_i = x'_i(i\le k);\\ 1\le k\le n-1\end{array}\\
&\Lambda^{\vec{x}}_{\vec{a}}\Lambda^{\vec{x}'}_{\vec{a}'} = 0, \forall \vec{x},\vec{x}',\vec{a},\vec{a}' \text{ such that } \quad \begin{array}{l} x_i = x'_i, (i\le k)\\ (a_1,\dots,a_k)\neq(a'_1,\dots,a'_k);\\ 1\le k\le n\end{array}\\
&\sum_{a_1,\dots,a_n}\Lambda^{\vec{x}}_{\vec{a}} = \mathbb{1};\\
&(\Lambda^{\vec{x}}_{\vec{a}})^\dagger = \Lambda^{\vec{x}}_{\vec{a}},
\end{aligned}\tag{6.4}$$

and similarly for $\Pi^{\vec{y}}_{\vec{b}}$.

A proof of this result can be found in [3], here we will only discuss its applications to numerical simulations.

6.4 Sequential NPA hierarchy

Everything we said in chapter 4 about the NPA hierarchy can be easily adapted to the sequential case, and we can find an algorithm that solves the following problem: given a behavior $p(\vec{a}, \vec{b}|\vec{x}, \vec{y})$ does it belong to the sequential quantum set Q_{SEQ}? Indeed, the only differences between the two cases is that in the sequential one operators have more constraints, as it can be seen by comparing equation (6.4) with (4.3). To take them in account, we repeat the derivation of the NPA hierarchy done in chapter 4, with the following additions:

- When defining the reduced sets, see equation (4.4), we need to take in account the new additional redundancy added by the one-way no-signaling conditions
$$\sum_{a_{k+1},\dots,a_n} \left(\Lambda_{\vec{a}}^{\vec{x}} - \Lambda_{\vec{a}}^{\vec{x}'}\right) = 0, \quad \begin{array}{l} \forall a_1, ..., a_k \\ \forall \vec{x}, \vec{x}' \text{ such that } x_i = x'_i (i \le k); \\ 1 \le k \le n-1 \end{array}$$
Indeed, since they are linear in Λ, from each of them we can write a $\Lambda_{\vec{a}}^{\vec{x}}$ in function of the others, as we did with the completeness relation
$$\sum_{a_1,\dots,a_n} \Lambda_{\vec{a}}^{\vec{x}} = \mathbb{I};$$
- Then, the remaining new orthogonal constraints
$$\Lambda_{\vec{a}}^{\vec{x}} \Lambda_{\vec{a}'}^{\vec{x}'} = 0, \forall \vec{x}, \vec{x}', \vec{a}, \vec{a}' \text{ such that } \quad \begin{array}{l} x_i = x'_i, (i \le k) \\ (a_1, ..., a_k) \neq (a'_1, ..., a'_k) \\ 1 \le k \le n \end{array}$$
are simply included, at each order n of the NPA hierarchy, in the sets of constraints of the certificate Γ^n, similarly to what we did in (4.15).

6.5 Another characterization of Q'_{SEC}

Notation In this section we will use the following notation: given a sequence of input
$$\vec{x} = (x_1, ..., x_N)$$
we will denote with $\vec{x}_k$ its truncation to the k-th element:
$$\vec{x}_k = (x_1, ..., x_k).$$
Furthermore, we will say that $\vec{x}_l \geq \vec{x}_k$ if $l \geq k$ and the first k elements in $\vec{x}_l$ are the same of $\vec{x}_k$ (i.e. $\vec{x}_k$ is a truncation of $\vec{x}_l$).

Characterizing Q'_{SEQ} The set of constraints we found in equation (6.4), are particularly useful in numerical simulations since they are a superset of the non-sequential case, see (4.3). However, for theoretical calculations a better characterization can be found:

A given behavior $p(\vec{a}, \vec{b}|\vec{x}, \vec{y})$ belongs to the sequential quantum set Q'_{SEQ} if and only if it can be written as
$$p(\vec{a}, \vec{b}|\vec{x}, \vec{y}) = \langle\psi| \prod_k \Lambda_{a_k}^{\vec{x}_k} \otimes \prod_k \Pi_{b_k}^{y_k} |\psi\rangle\,,$$
where the operators satisfy
$$\begin{aligned} &\sum_{a_k} \Lambda_{a_k}^{\vec{x}_k} = \mathbb{I}, \quad \forall k, \vec{x}_k; \\ &(\Lambda_{a_k}^{\vec{x}_k})^\dagger = \Lambda_{a_k}^{\vec{x}_k}, \quad \forall k, \vec{x}_k, a_k; \\ &\Lambda_{a_k}^{\vec{x}_k}\Lambda_{a'_k}^{\vec{x}_k} = \delta_{a_k,a'_k}\Lambda_{a_k}^{\vec{x}_k}, \quad \forall k, \vec{x}_k, a_k, a'_k \\ &[\Lambda_{a_k}^{\vec{x}_k}, \Lambda_{a_l}^{\vec{x}_l}] = 0, \quad \forall k, l, a_k, a_l, \vec{x}_l \geq \vec{x}_k, \end{aligned} \tag{6.5}$$

and symmetrical results for Bob-1,...,Bob-M operators $\Pi^{\vec{y}_k}_{b_k}$. The proof of this result can be found in [10]. Moreover, in this work we will focus on the case with 1 Alice and 2 Bob, each of them with binary inputs $x_1, y_1, y_2 \in \{0,1\}$ and outputs $a_1, b_1, b_2 \in \{-1,+1\}$. In this special case Alice's side is non-sequential and equation (6.5) reduces to to

$$\begin{aligned}
&\sum_{a_1=\pm1} \Lambda^{x_1}_{a_1} = \mathbb{I}, \quad \forall x_1 \in \{0,1\};\\
&(\Lambda^{x_1}_{a_1})^\dagger = \Lambda^{x_1}_{a_1}, \quad \forall x_1 \in \{0,1\}, a_1 \in \{\pm1\};\\
&\Lambda^{x_1}_{a_1}\Lambda^{x_1}_{a'_1} = \delta_{a_1,a'_1}\Lambda^{x_1}_{a_1}, \quad \forall x_1 \in \{\pm1\}, a_1, a'_1 \in \{\pm1\},
\end{aligned} \tag{6.6}$$

which is indeed the correct set of non-sequential constraints. On Bob side instead, by explicitly expanding the vector $\vec{y} \equiv y_1y_2$, we can rewrite equation (6.5) as

$$\begin{aligned}
&\sum_{b_1=\pm1} \Pi^{y_1}_{b_1} = \mathbb{I}, \quad \forall y_1 \in \{0,1\};\\
&\sum_{b_2=\pm1} \Pi^{y_1y_2}_{b_2} = \mathbb{I}, \quad \forall y_1, y_2 \in \{0,1\};\\
&(\Pi^{y_1}_{b_1})^\dagger = \Pi^{y_1}_{b_1}, \quad \forall y_1 \in \{0,1\}, b_1 \in \{\pm1\};\\
&(\Pi^{y_1y_2}_{b_2})^\dagger = \Pi^{y_1y_2}_{b_2}, \quad \forall y_1, y_2 \in \{0,1\}, b_2 \in \{\pm1\};\\
&\Pi^{y_1}_{b_1}\Pi^{y_1}_{b'_1} = \delta_{b_1,b'_1}\Pi^{y_1}_{b_1}, \quad \forall y_1 \in \{\pm1\}, b_1, b'_1 \in \{\pm1\};\\
&\Pi^{y_1y_2}_{b_2}\Pi^{y_1y_2}_{b'_2} = \delta_{b_2,b'_2}\Pi^{y_1y_2}_{b_2}, \quad \forall y_1, y_2 \in \{\pm1\}, b_2, b'_2 \in \{\pm1\};\\
&[\Pi^{y_1}_{b_1}, \Pi^{y_1y_2}_{b_2}] = 0, \quad \forall b_1, b_2 \in \{\pm1\}, y_1, y_2 \in \{0,1\}.
\end{aligned} \tag{6.7}$$

The interpretation of those projectors is the following: let us assume that they are measuring on a pure state $|\psi\rangle$, then

$$p_1 = \langle\psi|\, \Pi^{y_1}_{b_1} \,|\psi\rangle$$

is the probability of Bob-1 having b_1 as output, given that he chose y_1 as input, instead

$$p_2 = \langle\psi|\, \Pi^{y_1y_2}_{b_2} \,|\psi\rangle$$

is the probability of Bob-2 having b_2 as output, given that Bob-1 and Bob-2 chose respectively y_1 and y_2 as inputs. In other words we found a representation in which Bob-2 projectors act directly on $|\psi\rangle$, in place of the post-measurement state of Bob-1, but as drawback they acquire an explicit dependence on the input chose by Bob-1. Finally, notice that the first 6 constraints of (6.7) trivially state that both $\Pi^{y_1}_{b_1}$ and $\Pi^{y_1y_2}_{b_2}$ are projectors. The only interesting addition of sequentiality is the commutation relation

$$[\Pi^{y_1}_{b_1}, \Pi^{y_1y_2}_{b_2}] = 0, \quad \forall b_1, b_2 \in \{\pm1\}, y_1, y_2 \in \{0,1\},$$

that we will exploit in next chapter for self-testing proofs.

Chapter 7

New protocols

With all the theoretical tools presented, we are ready to derive new results. The final aim is finding sequential protocols, with three users from which we can extract random numbers in a device-independent way. In the non-sequential case such task is usually done by considering behaviors $p(a, b|x, y)$ that are self-testable thanks to the maximal violation of one (or more) Bell's inequality. Then, by using self-testing results is possible to compute rate of extractions, for example by using the Von-Neumann entropy (5.9) or the min-entropy (5.11). In our, sequential, case we would like to do something similar: we want to find behaviors that maximally violate some Bell's inequalities, and exploit these violations to acquire knowledge about the system, that can be used to generate device-independent random bits. However, finding maximally violated Bell's inequalities is difficult, and therefore a natural idea is using the method of sequential extensions:

1. We start by considering a non-sequential protocol which has already been self-tested;
2. Then, we sequentially extend it by adding another user;
3. Finally we try to self-test the operators we added in the previous step. Note that, thanks to the initial assumption, everything else is already self-tested, thus simplifying the part of finding maximally violated Bell's inequalities.

7.1 Non-sequential starting point

So, let us begin with step 1 and consider all behaviors of the form

$$p(a, b|x, y) = \langle\psi| \widetilde{\Lambda}^x_a \otimes \widetilde{\Pi}^y_b |\psi\rangle$$

such that state and operators

$$\widetilde{A}_x = \sum_a a\widetilde{\Lambda}^x_a;$$
$$\widetilde{B}_y = \sum_b b\widetilde{\Pi}^y_b,$$

admit a self-testing of the following form

$$\begin{aligned}
&\widetilde{A}_0 = \cos(\alpha_0)\sigma_x + \sin(\alpha_0)\sigma_z; \quad \widetilde{A}_1 = \cos(\alpha_1)\sigma_x + \sin(\alpha_1)\sigma_z;\\
&\widetilde{B}_0 = \cos(\beta_0)\sigma_x + \sin(\beta_0)\sigma_z; \quad \widetilde{B}_1 = \cos(\beta_1)\sigma_x + \sin(\beta_1)\sigma_z;\\
&|\psi\rangle = |\phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).
\end{aligned} \tag{7.1}$$

where $\alpha_0, \alpha_1, \beta_0, \beta_1$ are four fixed parameters. So basically, up to local isometries, Alice and Bob are measuring two generic 1-qubit real operators, with binary outcomes $a, b \in \{-1, +1\}$. Indeed

$(\mathbb{I}, \sigma_x, \sigma_y, \sigma_z)$ is a basis for the 1-qubit operators, and therefore a generic element O can be expanded as

$$O = c_1\mathbb{I} + c_2\sigma_x + c_3\sigma_y + c_4\sigma_z.$$

Observables are Hermitian, and this add the constraint

$$c_1, c_2, c_3, c_4 \in \mathbb{R}.$$

The condition of being real is equivalent to requiring that operators have no component along σ_y, $c_3 = 0$. Finally having binary outcomes $a, b \in \{-1, +1\}$ is possible only if

$$c_1 = 0, \quad c_2^2 + c_3^2 = 1,$$

consequently

$$O = \cos(\theta)\sigma_x + \sin(\theta)\sigma_z$$

for some angle θ, consistent with equation (7.1). Notice that we are just requiring the existence of such self-testing, we don't really care how it is actually done, for example there could be one (or more) maximally violated Bell's inequalities. Some examples of protocol satisfying these conditions are the CHSH, described in section 3.8, or the ones proposed in [15].

7.2 Invariance under rotation

In equation (7.1) we parameterized a family of non-sequential protocols with 4 parameters $(\alpha_0, \alpha_1, \beta_0, \beta_1)$. In this section we will see that the system is symmetric and by ”rotating the x and z axis”, we can get rid of one angle, thus ending up with only 3 degrees of freedom. The symmetry can be seen by computing the only relevant physical quantity: the behavior $p(a, b|x, y)$

$$p(a, b|x, y) = \frac{1}{4}\langle\phi_+|\,(\mathbb{I} + a\widetilde{A}_x) \otimes (\mathbb{I} + b\widetilde{B}_y)\,|\phi_+\rangle = \frac{1}{4}\left[1 + (-1)^{ab}\cos(\alpha_x - \beta_y)\right].$$

It depends only on the difference

$$\alpha_x - \beta_y,$$

and therefore is invariant under translation of all angles

$$\alpha_x \to \alpha_x + c; \quad \beta_y \to \beta_y + c, \quad c \in \mathbb{R}.$$

So without loss of generality we can fix $\beta_1 = 0$.

Summary We found that the space of non-sequential protocol that satisfy the set of conditions described above, has 3 degrees of freedom and consists of Alice and Bob measuring

Person	Measurement 1	Measurement 2
Alice 1	$\cos(\alpha_0)\sigma_x + \sin(\alpha_0)\sigma_z$	$\cos(\alpha_1)\sigma_x + \sin(\alpha_1)\sigma_z$
Bob 1	$\cos(\beta_0)\sigma_x + \sin(\beta_0)\sigma_z$	σ_x

on the state $|\psi\rangle = |\phi_+\rangle$. The angles $\alpha_0, \alpha_1, \beta_0$ can take any values, provided that from the corresponding probability distribution $p(a, b|x, y)$ is possible to self-test both state and measurements.

7.3 Sequential extension of the protocol

At this point we want to find a sequential extension of the protocol. It can be done in many ways, our choice is to focus on the following points

- We will extend the protocol on the Bob's side, so there will be 3 users: 1-Alice and 2-Bob and a corresponding behavior of the form $p(a, b_1, b_2|x, y_1, y_2)$;

- All users will still have binary inputs and binary outputs
$$x, y_1, y_2 \in \{0,1\} \quad a, b_1, b_2 \in \{-1,1\}.$$
- Extension in the sense that the four original operators will still be measured and the shared state will be $|\psi\rangle = |\phi_+\rangle$. This is important because it will allow us to exploit the previous assumption about self-testing;
- Bob-1 will have to perform at least one generalized weak measurement. Indeed projective measurements remove the entanglement between the two qubits. Without entanglement the quantum system decays to a classical one, making the presence of Bob-2 (who would measure on the unentangled post-measurement state of Bob-1) useless.

Therefore, we propose the sequential extension of table 7.1, where all operators are measured on the shared state $|\phi\rangle = |\phi_+\rangle$. An important point: our results will not depend on the protocol, otherwise we would be in a trusted scenario. Indeed, results will depend only on the behavior $p(a, b_1, b_2|x, y_1, y_2)$ generated by it, so that from any other protocol that yields the same probability distribution we would be able to generate the same number of random bits (basically the definition of device-independence). The points of table 7.1 is just giving the intuition on how such probability distribution has been found. With that said let us analyze the table: Alice-1 is still measuring $\widetilde{A}_0$ and $\widetilde{A}_1$, Bob-1 is measuring $\widetilde{B}_0$

Person	Measurement 1	Measurement 2
Alice 1	$\cos(\alpha_0)\sigma_x + \sin(\alpha_0)\sigma_z$	$\cos(\alpha_1)\sigma_x + \sin(\alpha_1)\sigma_z$
Bob 1	$w(\cos(\delta_1)\sigma_x + \sin(\delta_1)\sigma_z, \theta)$	$\cos(\beta_0)\sigma_x + \sin(\beta_0)\sigma_z$
Bob 2	$\cos(\delta)\sigma_x + \sin(\delta)\sigma_z$	σ_x

Table 7.1: Operators for the sequential protocol

and the generalized operator
$$w(\cos(\delta_1)\sigma_x + \sin(\delta_1)\sigma_z, \theta),$$
where with $w(\hat{O}, \theta)$, as discussed in section (6.2), we denote the non-projective measurement of a 1-qubit operator $\hat{O}$, defined by the two Kraus
$$K_+(\theta) = \cos(\theta)\frac{\mathbb{I} + \hat{O}}{2} + \sin(\theta)\frac{\mathbb{I} - \hat{O}}{2};$$
$$K_-(\theta) = \sin(\theta)\frac{\mathbb{I} + \hat{O}}{2} + \cos(\theta)\frac{\mathbb{I} - \hat{O}}{2}.$$
Bob-2 instead is measuring $\widetilde{B}_1$ and a new operator
$$\cos(\delta)\sigma_x + \sin(\delta)\sigma_z.$$
Overall we have three new degrees of freedoms δ, δ_1 and θ. The last one in particular measure the "weakness" of the measurement, while δ and δ_1 should be chosen in such a way to maximize the device-independent number of bits generated per round. Again, we are always implicitly focusing on the behavior generated by the protocol (and not on the protocol itself, or we would be in a trusted scenario). Therefore a choice of the angles has to be thought has a choice of probability distribution.

Steps of the protocol In a single round the following steps are performed

- The state $|\psi\rangle$ is sent to Alice-1 and Bob-1, that randomly measure one of their operator.
- Then, if Bob-1 chose the projective measurement the protocol ends, otherwise the post-measurement state is sent to Bob-2 who randomly measure one of his operators.

This is slightly different from what described in chapter 6, because we are stopping the protocol if Bob-1 measures his projective operator. Such choice is an optimization because, as already said, after a projective measurement the system's entanglement is lost.

Effective set of operators Bob-2 operators defined on table (7.1) are not really useful, because they act on the post-measurement state of Bob-1. Instead we would like to define some effective operators of the form $B_{y_1 y_2}$ that act directly on $|\psi\rangle$ with the downside of acquiring an explicit dependence on the Bob-1 input y_1. We have already seen in equation (6.3) how to do such transformation and the result is

$$\begin{aligned}
B_{00}^{n_d} &= K_+(\theta,\delta_1)(\cos(\delta)\sigma_x + \sin(\delta)\sigma_z)K_+^\dagger(\theta,\delta_1) + K_-(\theta,\delta_1)(\cos(\delta)\sigma_x + \sin(\delta)\sigma_z)K_-^\dagger(\theta,\delta_1);\\
B_{01}^{n_d} &= K_+(\theta,\delta_1)\sigma_x K_+^\dagger(\theta,\delta_1) + K_-(\theta,\delta_1)\sigma_x K_-^\dagger(\theta,\delta_1),
\end{aligned}$$

where superscript n_d stands for "non-dilated", and its significance will become clear shortly. So $B_{00}^{n_d}$ and $B_{01}^{n_d}$ are respectively the first and second operators measured of Bob-2 on $|\psi\rangle$, given that Bob-1 chose $y_1 = 0$. Note that we don't have to define $B_{10}^{n_d}$ and $B_{11}^{n_d}$ since the protocol stops after the choice $y_1 = 1$. Therefore, by putting everything together, we conclude that the protocol consists of the following 6 effective operators

$$\begin{aligned}
&A_0^{n_d} = \widetilde{A}_0 = \cos(\alpha_0)\sigma_x + \sin(\alpha_0)\sigma_z, \quad A_1^{n_d} = \widetilde{A}_1 = \cos(\alpha_1)\sigma_x + \sin(\alpha_1)\sigma_z;\\
&B_0^{n_d} = w(\cos(\delta_1)\sigma_x + \sin(\delta_1)\sigma_z,\theta), \quad B_1^{n_d} = \widetilde{B}_0 = \cos(\beta_0)\sigma_x + \sin(\beta_0)\sigma_z;\\
&B_{01}^{n_d} = K_+(\theta,\delta_1)\sigma_x K_+^\dagger(\theta) + K_-(\theta,\delta_1)\sigma_x K_-^\dagger(\theta,\delta_1);\\
&B_{00}^{n_d} = K_+(\theta,\delta_1)(\cos(\delta)\sigma_x + \sin(\delta)\sigma_z)K_+^\dagger(\theta,\delta_1) + K_-(\theta,\delta_1)(\cos(\delta)\sigma_x + \sin(\delta)\sigma_z)K_-^\dagger(\theta,\delta_1),
\end{aligned} \tag{7.2}$$

This characterization is almost equivalent to the one defined in section 6.5, the only difference is that here we are using the Kraus operator formalism, in place of projectors. The discrepancy can be fixed by performing a Stinespring dilation that maps the generalized operator

$$B_0^{n_d} = w(\cos(\delta_1)\sigma_x + \sin(\delta_1)\sigma_z,\theta)$$

to a unitary one. To denote the dilated version we will just drop the n_d superscript

$$A_{x_1}^{n_d} \to A_{x_1}; \quad B_{y_1}^{n_d} \to B_{y_1}; \quad B_{0y_1}^{n_d} \to B_{0y_1}.$$

Those new operators exactly match the definition given in section 6.5, and in particular they satisfy the full set of constraints (6.6) and (6.7):

$$\begin{aligned}
&A_{x_1}^\dagger = A_{x_1}, \quad \forall x_1 \in \{0,1\};\\
&A_{x_1}^2 = \mathbb{I}, \quad \forall x_1 \in \{0,1\};\\
&B_{y_1}^\dagger = B_{y_1}, \quad \forall y_1 \in \{0,1\};\\
&B_{y_1}^2 = \mathbb{I}, \quad \forall y_1 \in \{0,1\};\\
&B_{0y_2}^2 = \mathbb{I} \quad \forall y_2, \in \{0,1\};\\
&(B_{0y_2})^\dagger = B_{0y_2}, \quad \forall y_2, \in \{0,1\};\\
&[B_0, B_{0y_2}] = 0, \quad \forall y_2 \in \{0,1\};\\
&[A_{x_1}, B_{y_1}] = [A_{x_1}, B_{0y_2}] = 0, \quad \forall x_1, y_1, y_2 \in \{0,1\},
\end{aligned} \tag{7.3}$$

which are extremely important and will be used for self-testing in next sections. Notice that we do not need the explicit form of the Stinespring dilation because it doesn't change expectation values, hence we can still compute all physical quantities with their generalized (weak) version of equation (7.2). So for example:

- To prove that on our protocol a given Bell's operator $\widetilde{S}$ has quantum bound $\widetilde{I}_s$ we would formally work on the Stinespring dilated space, therefore assuming constraints (7.3);
- To prove that our choice of state and measurements yield

  $$\langle \widetilde{S} \rangle = \widetilde{I}_s,$$

  can be done in the non-dilated space with generalized operators.

The explicit form of the Stinespring map is never needed, it's enough to know that it exists.

7.4 Bell inequalities

We assumed that from the original non-sequential behavior $p(a,b|x,y)$ is possible to self-test state and measurements

$$\widetilde{A}_0, \widetilde{A}_1, \widetilde{B}_0, \widetilde{B}_1, |\phi_+\rangle \,. \tag{7.4}$$

Clearly, in the sequential version we would like to exploit this result, and to achieve this we just need to make sure that we are still measuring somewhere those same operators (on the same wavefunction) in such a way to extract $p(a,b|x,y)$ from $p(a,b_1,b_2|x,y_1,y_2)$. Well, the wavefunction is still $|\phi_+\rangle$ and the results holds for Alice since

$$\widetilde{A}_0 = A_0, \quad \widetilde{A}_1 = A_1,$$

and for one of Bob's operator

$$\widetilde{B}_0 = B_1.$$

What about $\widetilde{B}_1 = \sigma_x$? Its corresponding in the sequential protocol is B_{01}, see equation (7.2), but it is a priori different, since it takes in consideration the previous measurement of Bob-1. Therefore we need to explicitly impose $\widetilde{B}_1 = B_{01}$:

$$B_{01} = \widetilde{B_1} \iff \sigma_x = K_+(\theta,\delta_1)\sigma_x K_+^\dagger(\theta,\delta_1) + K_-(\theta,\delta_1)\sigma_x K_-^\dagger(\theta,\delta_1),$$

and by doing the calculation it is found that the result holds if and only if

$$\delta_1 = 0.$$

So by fixing $\delta_1 = 0$, the joint probability distribution generated by

$$A_0, A_1, B_1, B_{01}, |\phi_+\rangle \tag{7.5}$$

is the same as the one generated by (7.1), and this allows us to perform self-testing on all elements of equation (7.5). As already said, we don't care how this self-testing is actually done, however for the sake of giving a name, we will symbolically say that it follows from the saturation of a given set of Tsirelson inequalities

$$\langle S_k \rangle \leq I_k, \quad k \in \mathbb{N}. \tag{7.6}$$

From the self-testing it follows that, up to local isometries:

$$\{A_0, A_1\}\,|\psi\rangle = 2(\cos\alpha_0 \cos\alpha_1 + \sin\alpha_0 \sin\alpha_1)\,|\psi\rangle \,. \tag{7.7}$$

For completeness we report the updated version of table 7.1, with the additional constraint of $\delta_1 = 0$.

Person	Measurement 1	Measurement 2
Alice 1	$\cos(\alpha_0)\sigma_x + \sin(\alpha_0)\sigma_z$	$\cos(\alpha_1)\sigma_x + \sin(\alpha_1)\sigma_z$
Bob 1	$w(\sigma_x, \theta)$	$\cos(\beta_0)\sigma_x + \sin(\beta_0)\sigma_z$
Bob 2	$\cos(\delta)\sigma_x + \sin(\delta)\sigma_z$	σ_x

Table 7.2: Updated operators for the sequential protocol: we have two parameters θ, that measure the weakness of Bob-1 first operator, and δ that should be chosen in such a way to maximize the RNG ratio.

At this point we have self-tested the wave function and 4 of the 6 operators. We still need to self-test B_0 and B_{00}. To perform such task we need another Bell's inequality, so consider the following operator:

$$\begin{aligned} S_{\theta,\delta} = {} & -\frac{1}{2}\mathbb{1} + \frac{\cos(\delta)\sin^2(2\theta) f_x(A_0,A_1) \otimes B_{00} + \cos(2\theta)\sin^2(\delta) f_x(A_0,A_1) \otimes B_0}{2(1-\cos^2(\delta)\cos^2(2\theta))} + \\ & + \frac{\sin(\delta)\sin(2\theta) f_z(A_0,A_1) \otimes B_{00} - \sin(\delta)\cos(\delta)\cos(2\theta)\sin(2\theta) f_z(A_0,A_1) \otimes B_0}{2(1-\cos^2(\delta)\cos^2(2\theta))}; \\ & \text{where} \quad \begin{pmatrix} f_x(A_0,A_1) \\ f_z(A_0,A_1) \end{pmatrix} = \begin{pmatrix} \cos\alpha_0 & \sin\alpha_0 \\ \cos\alpha_1 & \sin\alpha_1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} A_0 \\ A_1 \end{pmatrix}, \end{aligned} \tag{7.8}$$

well defined for

$$1 - \cos^2(\delta)\cos^2(2\theta) \neq 0.$$

The important result is that for any behavior that saturates all Tsirelson inequalities of (7.6), then

$$\langle S_{\theta,\delta} \rangle \leq 0,$$

and moreover our state and operators (7.2), with $\delta_1 = 0$, saturate the bound, i.e they yield

$$\langle S_{\theta,\delta} \rangle = 0.$$

To prove this results we need to use two properties

- All operators in $S_{\theta,\delta}$ are unitary and hermitian, in particular they satisfy (7.3);
- Relation (7.7) holds, since all $\langle S_k \rangle$ are maximally violated.

With those properties and a lot of algebra it is possible to show that

$$S^2_{\theta,\delta} |\psi\rangle = -S_{\theta,\delta} |\psi\rangle ,$$

then by using the hermiticity of $S_{\theta,\delta}$:

$$0 \leq \| S_{\theta,\delta} |\psi\rangle \|^2 = \langle \psi | S^2_{\theta,\delta} |\psi\rangle = -\langle S_{\theta,\delta} \rangle \implies 0 \geq \langle S_{\theta,\delta} \rangle, \tag{7.9}$$

ant this concludes the proof. In case of maximal violation of $\langle S_{\theta,\delta} \rangle$ we have that:

$$\begin{aligned} S_{\theta,\delta} |\psi\rangle = 0 \implies |\psi\rangle = C_1 f_x(A_0, A_1) \otimes B_{0,0} + C_2 f_x(A_0, A_1) \otimes B_0 + \\ + C_3 f_z(A_0, A_1) \otimes B_{0,0} + C_4 f_z(A_0, A_1) \otimes B_0. \end{aligned} \tag{7.10}$$

Where C_1, C_2, C_3, C_4 are four real coefficients defined as:

$$\begin{aligned} C_1 &= \frac{\cos(\delta)\sin^2(2\theta)}{1 - \cos^2(\delta)\cos^2(2\theta)}; \\ C_2 &= \frac{\cos(2\theta)\sin^2(\delta)}{1 - \cos^2(\delta)\cos^2(2\theta)}; \\ C_3 &= \frac{\sin(\delta)\sin(2\theta)}{1 - \cos^2(\delta)\cos^2(2\theta)}; \\ C_4 &= -\frac{\sin(\delta)\cos(\delta)\cos(2\theta)\sin(2\theta)}{1 - \cos^2(\delta)\cos^2(2\theta)}, \end{aligned} \tag{7.11}$$

that in particular satisfy

$$C_1 C_2 - C_3 C_4 = 0. \tag{7.12}$$

Finally let us define

$$\eta \equiv \frac{C_1}{C_1^2 + C_3^2}, \tag{7.13}$$

such combination will appear often in the following sections.

7.5 Partial self testing

By assuming that all $\langle S_k \rangle$ and $\langle S_{\theta,\delta} \rangle$ are saturated we want to characterize our state and measurements. From the former it follows that:

$$\begin{aligned} |\psi\rangle &= |\phi_+\rangle \otimes |\xi\rangle \\ A_0 |\psi\rangle &= (\cos(\alpha_0)\sigma_x + \sin(\alpha_0)\sigma_z) |\phi_+\rangle \otimes |\xi\rangle \\ A_1 |\psi\rangle &= (\cos(\alpha_1)\sigma_x + \sin(\alpha_1)\sigma_z) |\phi_+\rangle \otimes |\xi\rangle \\ B_1 |\psi\rangle &= (\cos(\beta_0)\sigma_x + \sin(\beta_0)\sigma_z) |\phi_+\rangle \otimes |\xi\rangle \\ B_{01} |\psi\rangle &= \sigma_x |\phi_+\rangle \otimes |\xi\rangle \end{aligned}$$

At this point we cannot fully determine B_0 and B_{00} but, up to local isometries, we can decompose them with Pauli matrices:

$$\begin{aligned} B_0 \left|\psi\right\rangle &= (\mathbb{1} \otimes \gamma_0 + \sigma_x \otimes \gamma_1 + \sigma_y \otimes \gamma_2 + \sigma_z \otimes \gamma_3) \left|\phi_+\right\rangle \otimes \left|\xi\right\rangle \\ B_{00} \left|\psi\right\rangle &= (\mathbb{1} \otimes \tau_0 + \sigma_x \otimes \tau_1 + \sigma_y \otimes \tau_2 + \sigma_z \otimes \tau_3) \left|\phi_+\right\rangle \otimes \left|\xi\right\rangle \end{aligned}$$

All γ and τ operators are unfixed, but we can extract additional information by imposing all constraints of equation (7.3): for example by requiring $[B_0, B_{01}] = 0$ we find

$$\gamma_2 \left|\xi\right\rangle = \gamma_3 \left|\xi\right\rangle = 0. \tag{7.14}$$

Then by imposing the maximal violation of $S_{\theta,\delta}$, see equation (7.10), we get 4 additional equations

$$C_1\tau_1 \left|\xi\right\rangle + C_2\gamma_1 \left|\xi\right\rangle + C_3\tau_3 \left|\xi\right\rangle = \left|\xi\right\rangle; \tag{7.15}$$

$$C_1\tau_0 \left|\xi\right\rangle + C_2\gamma_0 \left|\xi\right\rangle + iC_3\tau_2 \left|\xi\right\rangle = 0; \tag{7.16}$$

$$-iC_1\tau_2 \left|\xi\right\rangle + C_3\tau_0 \left|\xi\right\rangle + C_4\gamma_0 \left|\xi\right\rangle = 0; \tag{7.17}$$

$$-C_1\tau_3 \left|\xi\right\rangle + C_3\tau_1 \left|\xi\right\rangle + C_4\gamma_1 \left|\xi\right\rangle = 0. \tag{7.18}$$

In particular consider their linear combination:

$$\begin{aligned} &(7.16)\cdot C_1 + (7.17)\cdot C_3 \implies (C_1^2 + C_3^2)\tau_0 \left|\xi\right\rangle + (C_1C_2 + C_3C_4)\gamma_0 \left|\xi\right\rangle = 0 \implies \tau_0 \left|\xi\right\rangle = 0; \\ &(7.15)\cdot C_1 + (7.18)\cdot C_3 \implies (C_1^2 + C_3^2)\tau_1 \left|\xi\right\rangle + (C_1C_2 + C_3C_4)\gamma_1 \left|\xi\right\rangle = C_1 \left|\xi\right\rangle \implies \\ &\implies \tau_1 \left|\xi\right\rangle = \frac{C_1}{C_1^2 + C_3^2} \left|\xi\right\rangle = \eta\mathbb{I} \left|\xi\right\rangle, \end{aligned} \tag{7.19}$$

where equation (7.12) has been used. By substituting what we just found in equation (7.15) and (7.16) we obtain

$$C_2\gamma_1 \left|\xi\right\rangle + C_3\tau_3 \left|\xi\right\rangle = (1 - C_1\eta) \left|\xi\right\rangle; \tag{7.20}$$

$$C_2\gamma_0 \left|\xi\right\rangle + iC_3\tau_2 \left|\xi\right\rangle = 0. \tag{7.21}$$

Let us also impose $[B_0, B_{00}] = 0$

$$\begin{aligned} 0 = [B_0, B_{00}] \left|\psi\right\rangle &= [\mathbb{1} \otimes \gamma_0 + \sigma_x \otimes \gamma_1, \sigma_x \otimes \tau_1 + \sigma_y \otimes \tau_2 + \sigma_z \otimes \tau_3] \left|\phi_+\right\rangle \left|\xi\right\rangle = \\ \sigma_y \left|\phi_+\right\rangle &\otimes ([\gamma_0, \tau_2] - i\{\gamma_1, \tau_3\}) \left|\xi\right\rangle + \sigma_z \left|\phi_+\right\rangle \otimes ([\gamma_0, \tau_3] + i\{\gamma_1, \tau_2\}) \left|\xi\right\rangle, \end{aligned}$$

and the only possibility is that both addends are zero, in particular:

$$([\gamma_0, \tau_2] - i\{\gamma_1, \tau_3\}) \left|\xi\right\rangle = 0. \tag{7.22}$$

Similarly, imposing $B_0^2 = \mathbb{I}$ yields

$$\begin{aligned} \left|\psi\right\rangle = B_0^2 \left|\psi\right\rangle &\implies \left|\phi_+\right\rangle \left|\xi\right\rangle = (\mathbb{1} \otimes \gamma_0 + \sigma_x \otimes \gamma_1)(\mathbb{1} \otimes \gamma_0 + \sigma_x \otimes \gamma_1) \left|\phi_+\right\rangle \left|\xi\right\rangle = \\ &= \left|\phi_+\right\rangle \otimes (\gamma_0^2 + \gamma_1^2) \left|\xi\right\rangle + \sigma_x \left|\phi_+\right\rangle \otimes \{\gamma_0, \gamma_1\} \left|\xi\right\rangle, \end{aligned}$$

and we conclude that the two gamma operators anti-commute

$$\{\gamma_0, \gamma_1\} \left|\xi\right\rangle = 0; \tag{7.23}$$

$$(\gamma_0^2 + \gamma_1^2) \left|\xi\right\rangle = \mathbb{I} \left|\xi\right\rangle. \tag{7.24}$$

Finally, we can exploit that B_{00} is unitary

$$\begin{aligned} \left|\psi\right\rangle = B_{00}^2 \left|\psi\right\rangle &\implies \left|\phi_+\right\rangle \left|\xi\right\rangle = (\sigma_x \otimes \tau_1 + \sigma_y \otimes \tau_2 + \sigma_z \otimes \tau_3)^2 \left|\phi_+\right\rangle \left|\xi\right\rangle = \\ &= \left|\phi_+\right\rangle \otimes (\tau_1^2 + \tau_2^2 + \tau_3^2) \left|\xi\right\rangle + i\sigma_x \left|\phi_+\right\rangle \otimes [\tau_2, \tau_3] \left|\xi\right\rangle, \end{aligned}$$

and the only possibility is

$$(\tau_1^2 + \tau_2^2 + \tau_3^2) \left|\xi\right\rangle = \left|\xi\right\rangle \implies (\tau_2^2 + \tau_3^2) \left|\xi\right\rangle = (1 - \eta^2) \left|\xi\right\rangle. \tag{7.25}$$

7.6 Guessing probability with min entropy

By using the previous self-testing results, we have enough information to generate device-independent random numbers. We will extract them from the outcomes of the measurements where all users select the first input ($x_1 = y_1 = y_2 = 0$). Following what we did in chapter 5 we need to take in consideration an adversary, Eve, with her set of projectors E_{e_1,e_2,e_3} that has the role of trying to guess the outcomes of Alice-1 and all Bob. Recall also that the quantity to compute is the guessing probability G, see equation (5.10):

$$G = \max \sum_{a_1,b_1,b_2} p(e_1 = a_1, e_2 = b_2, e_3 = b_3 | x_1 = y_1 = y_2 = 0).$$

Which can also be written in functions of the projectors

$$G = \max \sum_{a_1,b_1,b_2} \langle\psi| \Lambda^0_{a_1} \otimes \Pi^0_{b_1} \Pi^{00}_{b_2} \otimes E_{e_1,e_2,e_3} |\psi\rangle\,.$$

At this point we can use what we found in the self-testing section:

$$\begin{aligned}
\Lambda^0_{a_1} |\psi\rangle &= \frac{1}{2}\left[\mathbb{I} + a_1(\cos(\alpha_0)\sigma_x + \sin(\alpha_0)\sigma_z) \otimes \mathbb{I} \otimes \mathbb{I}\right] |\phi_+\rangle |\xi\rangle\,; \\
\Pi^0_{b_1} |\psi\rangle &= \frac{1}{2}\left[\mathbb{I} + b_1\mathbb{I} \otimes (\mathbb{1} \otimes \gamma_0 + \sigma_x \otimes \gamma_1)\right] |\phi_+\rangle |\xi\rangle\,; \\
\Pi^{00}_{b_2} |\psi\rangle &= \frac{1}{2}\left[\mathbb{I} + b_2\mathbb{I} \otimes (\sigma_x \otimes \tau_1 + \sigma_y \otimes \tau_2 + \sigma_z \otimes \tau_3)\right] |\phi_+\rangle |\xi\rangle\,.
\end{aligned} \tag{7.26}$$

so that the guessing probability becomes:

$$\begin{aligned}
G = \frac{1}{8}\max\Bigg\{ &\sum_{a_1,b_1,b_2} \langle\phi_+| \langle\xi| \left[\mathbb{I} + a_1(\cos(\alpha_0)\sigma_x + \sin(\alpha_0)\sigma_z) \otimes \mathbb{I} \otimes \mathbb{I}\right] \cdot \\
&\cdot \left[\mathbb{I} + b_1\mathbb{I} \otimes (\mathbb{1} \otimes \gamma_0 + \sigma_x \otimes \gamma_1)\right] \left[\mathbb{I} + b_2\mathbb{I} \otimes (\sigma_x \otimes \tau_1 + \sigma_y \otimes \tau_2 + \sigma_z \otimes \tau_3)\right] \otimes E_{a_1,b_1,b_2} |\phi_+\rangle |\xi\rangle \Bigg\}.
\end{aligned} \tag{7.27}$$

The expression can be hugely simplified by noticing that for any Pauli matrix σ_i we have that:

$$\begin{aligned}
&\langle\phi_+| \mathbb{1} \otimes \sigma_i |\phi_+\rangle = \langle\phi_+| \sigma_i \otimes \mathbb{1} |\phi_+\rangle = 0; \\
&\langle\phi_+| \sigma_i \otimes \sigma_j |\phi_+\rangle = 0, \quad \text{If } i \neq j.
\end{aligned}$$

Eventually the terms that survive are:

$$\begin{aligned}
G = \frac{1}{8} + \frac{1}{8}\max\Bigg\{ &\sum_{a_1,b_1,b_2} \Big(b_1\langle\gamma_0 \otimes E_{a_1,b_1,b_2}\rangle + b_1 b_2\langle\gamma_1\tau_1 \otimes E_{a_1,b_1,b_2}\rangle\Big) + \\
&+ \cos(\alpha_0) \sum_{a_1,b_1,b_2} a_1\Big(b_1\langle\gamma_1 \otimes E_{a_1,b_1,b_2}\rangle + b_2\langle\tau_1 \otimes E_{a_1,b_1,b_2}\rangle + b_1 b_2\langle\gamma_0\tau_1 \otimes E_{a_1,b_1,b_2}\rangle\Big) + \\
&+ \sin(\alpha_0) \sum_{a_1,b_1,b_2} a_1\Big(b_2\langle\tau_3 \otimes E_{a_1,b_1,b_2}\rangle + b_1 b_2\langle\gamma_0\tau_3 \otimes E_{a_1,b_1,b_2}\rangle + i b_1 b_2\langle\gamma_1\tau_2 \otimes E_{a_1,b_1,b_2}\rangle\Big) \Bigg\}.
\end{aligned} \tag{7.28}$$

We can get rid of τ_1 by using equation (7.19). The guessing probability becomes:

$$\begin{aligned}
G = \frac{1}{8} + \frac{1}{8}\max\Bigg\{ &\sum_{a_1,b_1,b_2} (b_1\langle\gamma_0 \otimes E_{a_1,b_1,b_2}\rangle + \eta b_1 b_2\langle\gamma_1 \otimes E_{a_1,b_1,b_2}\rangle) + \\
&+ \cos(\alpha_0) \sum_{a_1,b_1,b_2} a_1 (b_1\langle\gamma_1 \otimes E_{a_1,b_1,b_2}\rangle + \eta b_2\langle E_{a_1,b_1,b_2}\rangle + b_1 b_2 \eta\langle\gamma_0 \otimes E_{a_1,b_1,b_2}\rangle) + \\
&+ \sin(\alpha_0) \sum_{a_1,b_1,b_2} a_1 (b_2\langle\tau_3 \otimes E_{a_1,b_1,b_2}\rangle + b_1 b_2\langle\gamma_0\tau_3 \otimes E_{a_1,b_1,b_2}\rangle + i b_1 b_2\langle\gamma_1\tau_2 \otimes E_{a_1,b_1,b_2}\rangle) \Bigg\}.
\end{aligned} \tag{7.29}$$

At this point we compute each expectation value individually.

Computation of $\langle\gamma_0 \otimes E_{a_1,b_1,b_2}\rangle$ We will prove that

$$\langle\gamma_0 \otimes E_{a_1,b_1,b_2}\rangle = 0. \tag{7.30}$$

Indeed, multiply both sides of equation (7.21) by $\langle\xi|\, E_{a_1,b_1,b_2}$:

$$C_2\langle\gamma_0 \otimes E_{a_1,b_1,b_2}\rangle = -iC_3\langle\tau_2 \otimes E_{a_1,b_1,b_2}\rangle. \tag{7.31}$$

Since

- The operators $\gamma_0 \otimes E_{a_1,b_1,b_2}$ and $\tau_2 \otimes E_{a_1,b_1,b_2}$ are Hermitian;
- The coefficients C_2, C_3 are real,

we can conclude that the left hand side of equation (7.31) is a real number, and the right hand side is imaginary. Therefore the only possibility is that

$$\langle\gamma_0 \otimes E_{a_1,b_1,b_2}\rangle = \langle\tau_2 \otimes E_{a_1,b_1,b_2}\rangle = 0,$$

provided that $C_2 \neq 0$ (the effects of these additional constraints will be discussed in section 7.8). This is exactly what we wanted to prove and with this result we can simplify equation (7.29):

$$\begin{aligned}
G = \frac{1}{8} + \frac{1}{8}\max\Bigg\{ & \sum_{a_1,b_1,b_2} \eta b_1 b_2 \langle\gamma_1 \otimes E_{a_1,b_1,b_2}\rangle + \\
& + \cos(\alpha_0) \sum_{a_1,b_1,b_2} a_1 \Big(b_1\langle\gamma_1 \otimes E_{a_1,b_1,b_2}\rangle + \eta b_2\langle E_{a_1,b_1,b_2}\rangle\Big) + \\
& + \sin(\alpha_0) \sum_{a_1,b_1,b_2} a_1 \Big(b_2\langle\tau_3 \otimes E_{a_1,b_1,b_2}\rangle + b_1 b_2 [\langle\gamma_0\tau_3 \otimes E_{a_1,b_1,b_2}\rangle + i\langle\gamma_1\tau_2 \otimes E_{a_1,b_1,b_2}\rangle]\Big)\Bigg\}.
\end{aligned}$$

Computation of $\langle\gamma_0\tau_3 \otimes E_{a_1,b_1,b_2}\rangle + i\langle\gamma_1\tau_2 \otimes E_{a_1,b_1,b_2}\rangle$ This expectation value is also zero

$$\langle\gamma_0\tau_3 \otimes E_{a_1,b_1,b_2}\rangle + i\langle\gamma_1\tau_2 \otimes E_{a_1,b_1,b_2}\rangle = \langle(\gamma_0\tau_3 + i\gamma_1\tau_2) \otimes E_{a_1,b_1,b_2}\rangle = 0.$$

To see that, let us rewrite τ_2 in terms of γ_0, by using equation (7.21) and assuming that $C_3 \neq 0$:

$$\langle(\gamma_0\tau_3 + i\gamma_1\tau_2) \otimes E_{a_1,b_1,b_2}\rangle = \left\langle \left(\gamma_0\tau_3 - \frac{C_2}{C_3}\gamma_1\gamma_0\right) \otimes E_{a_1,b_1,b_2} \right\rangle,$$

and τ_3 in terms of γ_0 by using equation (7.20):

$$\left\langle \left(\gamma_0\tau_3 - \frac{C_2}{C_3}\gamma_1\gamma_0\right) \otimes E_{a_1,b_1,b_2} \right\rangle = -\frac{C_2}{C_3}\langle\{\gamma_0,\gamma_1\} \otimes E_{a_1,b_1,b_2}\rangle + \frac{1 - C_1\eta}{C_3}\langle\gamma_0 \otimes E_{a_1,b_1,b_2}\rangle.$$

Finally equations (7.30) and (7.23) imply that both addends of the last expression are zero. This concludes the proof. With this relation the guessing probability can be further simplified:

$$\begin{aligned}
G = \frac{1}{8} + \frac{1}{8}\max\Bigg\{ & \sum_{a_1,b_1,b_2} \eta b_1 b_2 \langle\gamma_1 E_{a_1,b_1,b_2}\rangle + \\
& + \cos(\alpha_0) \sum_{a_1,b_1,b_2} a_1 \left(b_1\langle\gamma_1 E_{a_1,b_1,b_2}\rangle + \eta b_2\langle E_{a_1,b_1,b_2}\rangle\right) + \\
& + \sin(\alpha_0) \sum_{a_1,b_1,b_2} a_1 b_2 \langle\tau_3 E_{a_1,b_1,b_2}\rangle \Bigg\}.
\end{aligned}$$

At this point we can rewrite the expectation value $\langle \tau_3 \otimes E_{a_1,b_1,b_2} \rangle$ in terms of $\langle \gamma_1 \otimes E_{a_1,b_1,b_2} \rangle$ by using equation (7.20). The guessing probability becomes:

$$\begin{aligned} G = \frac{1}{8} + \frac{1}{8} \max \Bigg\{ & \sum_{a_1,b_1,b_2} \eta b_1 b_2 \langle \gamma_1 \otimes E_{a_1,b_1,b_2} \rangle + \\ & + \cos(\alpha_0) \sum_{a_1,b_1,b_2} a_1 \left(b_1 \langle \gamma_1 \otimes E_{a_1,b_1,b_2} \rangle + \eta b_2 \langle E_{a_1,b_1,b_2} \rangle \right) + \\ & + \sin(\alpha_0) \sum_{a_1,b_1,b_2} a_1 b_2 \left(\frac{(1 - C_1 \eta) \langle E_{a_1,b_1,b_2} \rangle - C_2 \langle \gamma_1 \otimes E_{a_1,b_1,b_2} \rangle}{C_3} \right) \Bigg\}. \end{aligned} \tag{7.32}$$

Computation of $\langle \gamma_1 E_{a_1,b_1,b_2} \rangle$ This is the last expectation value we need to compute, we will prove that

$$\langle \gamma_1 \otimes E_{a_1,b_1,b_2} \rangle = \frac{C_2}{1 - C_1 \eta} \langle E_{a_1,b_1,b_2} \rangle. \tag{7.33}$$

We begin by splitting it in two addends

$$\langle \gamma_1 \otimes E_{a_1,b_1,b_2} \rangle = \frac{1}{2} \langle (\gamma_1 + \gamma_1) \otimes E_{a_1,b_1,b_2} \rangle,$$

on the first addend we substitute equation (7.20) and on the second its adjoint form

$$\frac{1}{2} \langle (\gamma_1 + \gamma_1) \otimes E_{a_1,b_1,b_2} \rangle = \frac{C_2}{1 - C_1 \eta} \langle \gamma_1^2 \otimes E_{a_1,b_1,b_2} \rangle + \frac{C_3}{2(1 - C_1 \eta)} \langle \{ \gamma_1, \tau_3 \} \otimes E_{a_1,b_1,b_2} \rangle$$

Then we use equation (7.24) to write γ_1 in terms of γ_0

$$\begin{aligned} & \frac{C_2}{1 - C_1 \eta} \langle \gamma_1^2 \otimes E_{a_1,b_1,b_2} \rangle + \frac{C_3}{2(1 - C_1 \eta)} \langle \{ \gamma_1, \tau_3 \} \otimes E_{a_1,b_1,b_2} \rangle = \\ & = \frac{C_2}{1 - C_1 \eta} \langle E_{a_1,b_1,b_2} \rangle - \frac{C_2}{1 - C_1 \eta} \langle \gamma_0^2 \otimes E_{a_1,b_1,b_2} \rangle + \frac{C_3}{2(1 - C_1 \eta)} \langle \{ \gamma_1, \tau_3 \} \otimes E_{a_1,b_1,b_2} \rangle. \end{aligned}$$

Therefore, to get the final result, equation (7.33), we have to show that

$$- \frac{C_2}{1 - C_1 \eta} \langle \gamma_0^2 \otimes E_{a_1,b_1,b_2} \rangle + \frac{C_3}{2(1 - C_1 \eta)} \langle \{ \gamma_1, \tau_3 \} \otimes E_{a_1,b_1,b_2} \rangle = 0,$$

which is done by using equation (7.21) and its adjoint:

$$\begin{aligned} & - \frac{C_2}{1 - C_1 \eta} \langle \gamma_0^2 \otimes E_{a_1,b_1,b_2} \rangle + \frac{C_3}{2(1 - C_1 \eta)} \langle \{ \gamma_1, \tau_3 \} \otimes E_{a_1,b_1,b_2} \rangle = \\ & = - \frac{C_2}{2(1 - C_1 \eta)} \langle (\gamma_0^2 + \gamma_0^2) \otimes E_{a_1,b_1,b_2} \rangle + \frac{C_3}{2(1 - C_1 \eta)} \langle \{ \gamma_1, \tau_3 \} \otimes E_{a_1,b_1,b_2} \rangle = \\ & = -i \frac{C_3}{2(1 - C_1 \eta)} \langle [\tau_2, \gamma_0] \otimes E_{a_1,b_1,b_2} \rangle + \frac{C_3}{2(1 - C_1 \eta)} \langle \{ \gamma_1, \tau_3 \} \otimes E_{a_1,b_1,b_2} \rangle = \\ & = \frac{i C_3}{2(1 - C_1 \eta)} \langle ([\gamma_0, \tau_2] - i \{ \gamma_1, \tau_3 \}) \otimes E_{a_1,b_1,b_2} \rangle = 0. \end{aligned}$$

Where in the last line we used equation (7.22). This concludes the proof. By substituting the result in the guessing probability, equation (7.32), we find that

$$\begin{aligned} G = \frac{1}{8} \max \Bigg\{ 1 + \sum_{a_1,b_1,b_2} \langle E_{a_1,b_1,b_2} \rangle \Bigg[& b_1 b_2 \cdot \frac{\eta C_2}{1 - C_1 \eta} + \\ + \cos(\alpha_0) a_1 \left(b_1 \frac{C_2}{1 - C_1 \eta} + \eta b_2 \right) & + \sin(\alpha_0) a_1 b_2 \left(\frac{(1 - C_1 \eta)^2 - C_2^2}{C_3 (1 - C_1 \eta)} \right) \Bigg] \Bigg\}. \end{aligned}$$

Notice that we managed to express it in function of only one expectation value $\langle E_{a_1,b_1,b_2}\rangle$, that depends only on Eve's operator. Finally by using the definition of η, equation (7.13), we can find an alternative expression for G:

$$G = \frac{1}{8}\max\Bigg\{1 + \sum_{a_1,b_1,b_2} \langle E_{a_1,b_1,b_2}\rangle \Bigg[b_1 b_2 \cdot \frac{C_1 C_2}{C_3^2} + \cos(\alpha_0) a_1 \Bigg(b_1 \frac{C_2(C_1^2+C_3^2)}{C_3^2} + b_2 \frac{C_1}{C_1^2+C_3^2}\Bigg) + \\ + \sin(\alpha_0) a_1 b_2 \Bigg(\frac{C_3^4 - C_2^2(C_1^2+C_3^2)^2}{C_3^3(C_1^2+C_3^2)}\Bigg)\Bigg]\Bigg\}.$$

A nicer expression can be found by substituting equation (7.11), in such a way that it depends explicitly on δ and θ. It is found:

$$G = \frac{1}{8}\max\Bigg\{1 + \sum_{a_1,b_1,b_2} \langle E_{a_1,b_1,b_2}\rangle \Bigg[b_1 b_2 \cos(\delta)\cos(2\theta) + \\ + \cos(\alpha_0) a_1 \Bigg(b_1 \cos(2\theta) + b_2 \cos(\delta)\Bigg) + a_1 b_2 \sin(\alpha_0)\sin(2\theta)\sin(\delta)\Bigg]\Bigg\}.$$

We still have to compute the maximum among all possible Eve's operator E_{a_1,b_1,b_2} and states $|\xi\rangle$. This task is done by recalling that, see equation (5.8):

$$\langle E_{a_1,b_1,b_2}\rangle \geq 0; \qquad \sum_{a_1,b_1,b_2} \langle E_{a_1,b_1,b_2}\rangle = 1.$$

Therefore, the best strategy for Eve is finding the triplet $(\widetilde{a}_1, \widetilde{b}_1, \widetilde{b}_2) \in \{-1,+1\}^3$ that maximizes

$$p(a_1,b_1,b_2) \equiv 1 + b_1 b_2 \cos(\delta)\cos(2\theta) + \cos(\alpha_0) a_1 \Bigg(b_1 \cos(2\theta) + b_2\cos(\delta)\Bigg) + a_1 b_2 \sin(\alpha_0)\sin(2\theta)\sin(\delta)$$

and then picking her operators in such a way that

$$\langle E_{a_1,b_1,b_2}\rangle = \delta_{a_1\widetilde{a}_1}\delta_{b_1\widetilde{b}_1}\delta_{b_2\widetilde{b}_2}.$$

With this choice G assumes its maximum value

$$G = \frac{1}{8} p(\widetilde{a}_1, \widetilde{b}_1, \widetilde{b}_2).$$

From which we can find the number of bits of randomness generated per round

$$r_{me} = -\log_2(G) = -\log_2\left[\frac{1}{8} p(\widetilde{a}_1, \widetilde{b}_1, \widetilde{b}_2)\right] = 3 - \log_2\left[p(\widetilde{a}_1, \widetilde{b}_1, \widetilde{b}_2)\right] \tag{7.34}$$

The value for the triplet $(\widetilde{a}_1, \widetilde{b}_1, \widetilde{b}_2)$ will depend on the protocol, in particular on the angles θ, α_0, δ. Furthermore a simple calculation shows that the result is optimal, in the sense that it matches the number of bits generated in the trusted case from equation (5.6).

7.7 Extension to Von Neumann entropy

In this section we will repeat the calculation by using the Von Neumann Entropy and verify that we can generate more than r_{me} bits of randomness. Following chapter 5 we have to compute

$$r_{vn} = \inf_{|\psi\rangle,\Lambda,\Pi} H(AB|E)_\rho$$

where ρ is the state after the measurement of Alice-1 Bob-1 and Bob-2

$$\rho = \sum_{a_1 b_1 b_2} |a_1 b_1 b_2\rangle\langle a_1 b_1 b_2| \operatorname{Tr}_{A,B}\left[|\psi\rangle\langle\psi| \Lambda^0_{a_1} \otimes \Pi^0_{b_1}\Pi^{0,0}_{b_2}\right] \tag{7.35}$$

and $|a_1 b_1 b_2\rangle$ are the eigenvalues of the chosen operators:

$$\begin{aligned}
&A_0 |a_1\rangle = a_1 |a\rangle ;\\
&B_0 |b_1\rangle = b_1 |b_1\rangle ;\\
&B_{00} |b_2\rangle = b_2 |b_2\rangle ;\\
&|a_1 b_1 b_2\rangle \equiv |a_1\rangle \otimes |b_1\rangle \otimes |b_2\rangle .
\end{aligned}$$

As mentioned in chapter 5 calculating r_{vn} is in general very difficult, but we know how to solve the simple case in which the Eve's part of ρ is uncorrelated from the rest

$$\rho \equiv \rho_{AB} \otimes \rho_E. \tag{7.36}$$

Luckily this is indeed the case, and we will now see how to factorize ρ. The first step is computing the trace of equation (7.35):

$$G_V \equiv \mathrm{Tr}_{A,B}\left[|\psi\rangle\langle\psi| \Lambda^0_{a_1} \otimes \Pi^0_{b_1} \Pi^{0,0}_{b_2}\right],$$

all projectors can be expanded by using the results about self-testing, see equation (7.26):

$$\begin{aligned}
G_V = \frac{1}{8}\,\mathrm{Tr}_{A,B}\Bigg\{ &|\phi_+\rangle\langle\phi_+| \otimes |\xi\rangle\langle\xi| \Big[\mathbb{I} + a_1(\cos(\alpha_0)\sigma_x + \sin(\alpha_0)\sigma_z) \otimes \mathbb{I} \otimes \mathbb{I}\Big]\cdot\\
&\Big[\mathbb{I} + b_1 \mathbb{I} \otimes (\mathbb{1} \otimes \gamma_0 + \sigma_x \otimes \gamma_1)\Big]\cdot\Big[\mathbb{I} + b_2 \mathbb{I} \otimes (\sigma_x \otimes \tau_1 + \sigma_y \otimes \tau_2 + \sigma_z \otimes \tau_3)\Big]\Bigg\}.
\end{aligned}$$

At this point it's convenient computing the trace over the 2-qubits subspace in which $|\phi_+\rangle$ lives, which is uncorrelated from the rest. Notice that for any Pauli matrix σ_i it holds that:

$$\begin{aligned}
&Tr[|\phi_+\rangle\langle\phi_+| \mathbb{I} \otimes \sigma_i] = 0;\\
&Tr[|\phi_+\rangle\langle\phi_+| \sigma_i \otimes \sigma_i] = 0, \quad \text{If } i \neq j.
\end{aligned}$$

This makes the computation easier and, eventually only a few terms survive

$$\begin{aligned}
G_V = \frac{1}{8}\,\mathrm{Tr}_B\Bigg\{ &|\xi\rangle\langle\xi| + b_1 |\xi\rangle\langle\xi| \gamma_0 + b_1 b_2 |\xi\rangle\langle\xi| \gamma_1\tau_1 +\\
&+ \cos(\alpha_0) a_1 \Big(b_1 |\xi\rangle\langle\xi| \gamma_1 + b_2 |\xi\rangle\langle\xi| \tau_1 + b_1 b_2 |\xi\rangle\langle\xi| \gamma_0\tau_1\Big) +\\
&+ \sin(\alpha_0) a_1 \Big(b_2 |\xi\rangle\langle\xi| \tau_3 + b_1 b_2 |\xi\rangle\langle\xi| \gamma_0\tau_3 + i b_1 b_2 |\xi\rangle\langle\xi| \gamma_1\tau_2\Big)\Bigg\}.
\end{aligned}$$

In particular notice the similarity with the corresponding equation for the min entropy (7.28). We can get rid of τ_1 by using equation (7.19) and the expression simplifies to

$$\begin{aligned}
G_V = \frac{1}{8}\,\mathrm{Tr}_B\Bigg\{ &|\xi\rangle\langle\xi| + b_1 |\xi\rangle\langle\xi| \gamma_0 + b_1 b_2 \eta |\xi\rangle\langle\xi| \gamma_1 +\\
&+ \cos(\alpha_0) a_1 \Big(b_1 |\xi\rangle\langle\xi| \gamma_1 + b_2 \eta |\xi\rangle\langle\xi| + b_1 b_2 \eta |\xi\rangle\langle\xi| \gamma_0\Big) +\\
&+ \sin(\alpha_0) a_1 \Big(b_2 |\xi\rangle\langle\xi| \tau_3 + b_1 b_2 |\xi\rangle\langle\xi| \gamma_0\tau_3 + i b_1 b_2 |\xi\rangle\langle\xi| \gamma_1\tau_2\Big)\Bigg\}.
\end{aligned}$$

Now we compute each partial trace individually:

Computation of $\mathrm{Tr}_B[|\xi\rangle\langle\xi|\gamma_0]$ This one is zero, indeed we can trivially write

$$\mathrm{Tr}_B[|\xi\rangle\langle\xi|\gamma_0] = \frac{1}{2}\mathrm{Tr}_B[|\xi\rangle\langle\xi|\gamma_0 + \gamma_0|\xi\rangle\langle\xi|],$$

where the cyclicity of the trace has been used. Then on the first addend we substitute equation (7.21) and on the second its adjoint:

$$\frac{1}{2}\mathrm{Tr}_B[|\xi\rangle\langle\xi|\gamma_0 + \gamma_0|\xi\rangle\langle\xi|] = \frac{iC_3}{2C_2}\mathrm{Tr}_B[|\xi\rangle\langle\xi|\tau_2 - \tau_2|\xi\rangle\langle\xi|] = 0.$$

Where in the last equality we used again the cyclicality of the trace. With this result the expression of G_V can be simplified to

$$\begin{aligned} G_V = \frac{1}{8}\mathrm{Tr}_B\Big\{ & |\xi\rangle\langle\xi| + b_1b_2\eta|\xi\rangle\langle\xi|\gamma_1 + \\ & + \cos(\alpha_0)a_1\Big(b_1|\xi\rangle\langle\xi|\gamma_1 + b_2\eta|\xi\rangle\langle\xi|\Big) + \\ & + \sin(\alpha_0)a_1\Big(b_2|\xi\rangle\langle\xi|\tau_3 + b_1b_2|\xi\rangle\langle\xi|\gamma_0\tau_3 + ib_1b_2|\xi\rangle\langle\xi|\gamma_1\tau_2\Big)\Big\}. \end{aligned}$$

Computation of $\mathrm{Tr}_B[|\xi\rangle\langle\xi|\gamma_0\tau_3 + i\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\gamma_1\tau_2]$ This partial trace is also zero

$$\mathrm{Tr}_B[|\xi\rangle\langle\xi|\gamma_0\tau_3 + i\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\gamma_1\tau_2] = 0.$$

To see that, let us rewrite τ_2 in terms of γ_0, by using equation (7.21):

$$\mathrm{Tr}_B[|\xi\rangle\langle\xi|(\gamma_0\tau_3 + i\gamma_1\tau_2)] = \mathrm{Tr}_B\left[|\xi\rangle\langle\xi|\left(\gamma_0\tau_3 - \frac{C_2}{C_3}\gamma_1\gamma_0\right)\right],$$

and τ_3 in terms of γ_0 by using equation (7.20):

$$\mathrm{Tr}_B\left[|\xi\rangle\langle\xi|\left(\gamma_0\tau_3 - \frac{C_2}{C_3}\gamma_1\gamma_0\right)\right] = -\frac{C_2}{C_3}\mathrm{Tr}_B[|\xi\rangle\langle\xi|\{\gamma_0,\gamma_1\}] + \frac{1-C_1\eta}{C_3}\mathrm{Tr}_B[|\xi\rangle\langle\xi|\gamma_0].$$

Finally equations (7.30) and (7.23) imply that both addends of the last expression are zero. This concludes the proof and we can further simplify the expression for G_V:

$$\begin{aligned} G_V = \frac{1}{8}\mathrm{Tr}_B\Big\{ & |\xi\rangle\langle\xi| + b_1b_2\eta|\xi\rangle\langle\xi|\gamma_1 + \\ & + \cos(\alpha_0)a_1\Big(b_1|\xi\rangle\langle\xi|\gamma_1 + b_2\eta|\xi\rangle\langle\xi|\Big) + \\ & + \sin(\alpha_0)a_1b_2|\xi\rangle\langle\xi|\tau_3\Big\}. \end{aligned}$$

Notice that $\mathrm{Tr}_B[|\xi\rangle\langle\xi|\tau_3]$ can be rewritten in terms of $\mathrm{Tr}_B[|\xi\rangle\langle\xi|\gamma_1]$ by using equation (7.20). G_V becomes:

$$\begin{aligned} G_V = \frac{1}{8}\mathrm{Tr}_B\Big\{ & |\xi\rangle\langle\xi| + b_1b_2\eta|\xi\rangle\langle\xi|\gamma_1 + \\ & + \cos(\alpha_0)a_1\Big(b_1|\xi\rangle\langle\xi|\gamma_1 + b_2\eta|\xi\rangle\langle\xi|\Big) + \\ & + \sin(\alpha_0)a_1b_2\Big(\frac{1-C_1\eta}{C_3}|\xi\rangle\langle\xi| - \frac{C_2}{C_3}|\xi\rangle\langle\xi|\gamma_1\Big)\Big\}. \end{aligned}$$

Computation of $\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\gamma_1]$ This is the last expectation value we need to compute, we will show that

$$\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\gamma_1] = \frac{C_2}{1-C_1\eta}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|]. \tag{7.37}$$

We begin with the trivial identity

$$\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\gamma_1] = \frac{1}{2}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\gamma_1 + |\xi\rangle\langle\xi|\,\gamma_1],$$

on the first addend we substitute equation (7.20) and on the second its adjoint form

$$\frac{1}{2}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\gamma_1 + |\xi\rangle\langle\xi|\,\gamma_1] = \frac{C_2}{1-C_1\eta}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\gamma_1^2] + \frac{C_3}{2(1-C_1\eta)}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\{\gamma_1,\tau_3\}]$$

Then we use equation (7.24) to write γ_1 in terms of γ_0

$$\begin{aligned}&\frac{C_2}{1-C_1\eta}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\gamma_1^2] + \frac{C_3}{2(1-C_1\eta)}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\{\gamma_1,\tau_3\}] =\\ &= \frac{C_2}{1-C_1\eta}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|] - \frac{C_2}{1-C_1\eta}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\gamma_0^2] + \frac{C_3}{2(1-C_1\eta)}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\{\gamma_1,\tau_3\}].\end{aligned}$$

Therefore, to match the result (7.37) we have to prove that

$$-\frac{C_2}{1-C_1\eta}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\gamma_0^2] + \frac{C_3}{2(1-C_1\eta)}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\{\gamma_1,\tau_3\}] = 0,$$

which is done by using equation (7.21) and its adjoint:

$$\begin{aligned}&-\frac{C_2}{1-C_1\eta}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\gamma_0^2] + \frac{C_3}{2(1-C_1\eta)}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\{\gamma_1,\tau_3\}] =\\ &= -\frac{C_2}{2(1-C_1\eta)}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,(\gamma_0^2+\gamma_0^2)] + \frac{C_3}{2(1-C_1\eta)}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\{\gamma_1,\tau_3\}] =\\ &= -i\frac{C_3}{2(1-C_1\eta)}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,[\tau_2,\gamma_0]] + \frac{C_3}{2(1-C_1\eta)}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,\{\gamma_1,\tau_3\}] =\\ &= \frac{iC_3}{2(1-C_1\eta)}\,\mathrm{Tr}_B[|\xi\rangle\langle\xi|\,([\gamma_0,\tau_2] - i\{\gamma_1,\tau_3\})] = 0.\end{aligned}$$

Where in the last line we used equation (7.22). This concludes the proof, and we can use the result to re-write G_V as

$$G_V = \frac{1}{8}\,\mathrm{Tr}_B\left[|\xi\rangle\langle\xi|\left(1 + b_1 b_2\cdot\frac{\eta C_2}{1-C_1\eta} + \cos(\alpha_0)a_1\left(b_1\frac{C_2}{1-C_1\eta} + \eta b_2\right) + \sin(\alpha_0)a_1 b_2\left(\frac{(1-C_1\eta)^2 - C_2^2}{C_3(1-C_1\eta)}\right)\right)\right].$$

Equivalently we can substitute the definition of η, equation (7.13):

$$\begin{aligned}G_V = \frac{1}{8}\Bigg\{&1 + b_1 b_2\cdot\frac{C_1C_2}{C_3^2} + \cos(\alpha_0)a_1\left(b_1\frac{C_2(C_1^2+C_3^2)}{C_3^2} + b_2\frac{C_1}{C_1^2+C_3^2}\right) +\\ &+ \sin(\alpha_0)a_1 b_2\left(\frac{C_3^4 - C_2^2(C_1^2+C_3^2)^2}{C_3^3(C_1^2+C_3^2)}\right)\Bigg\}\mathrm{Tr}_B[|\xi\rangle\langle\xi|],\end{aligned}$$

and finally by substituting equation (7.11), we can write it in function of δ and θ:

$$G_V = \frac{1}{8}\Bigg\{1 + b_1 b_2\cos(\delta)\cos(2\theta) + \cos(\alpha_0)a_1\Big(b_1\cos(2\theta) + b_2\cos(\delta)\Big) + a_1 b_2\sin(\alpha_0)\sin(2\theta)\sin(\delta)\Bigg\}\mathrm{Tr}_B[|\xi\rangle\langle\xi|].$$

By inserting this result back in equation (7.35) we obtain the post-measurement state

$$\rho = \sum_{a_1 b_1 b_2} |a_1 b_1 b_2\rangle \langle a_1 b_1 b_2| \, \frac{1}{8}\Bigg\{ 1 + b_1 b_2 \cos(\delta)\cos(2\theta) +$$
$$+ \cos(\alpha_0) a_1 \Bigg(b_1 \cos(2\theta) + b_2 \cos(\delta) \Bigg) + a_1 b_2 \sin(\alpha_0)\sin(2\theta)\sin(\delta) \Bigg\} \mathrm{Tr}_B[|\xi\rangle\langle\xi|],$$

which is separable and has the form of equation (7.36), up to defining

$$\rho_E \equiv \mathrm{Tr}_B[|\xi\rangle\langle\xi|]$$

and

$$\rho_{AB} = \sum_{a_1,b_1,b_2} \frac{1}{8}\Bigg[1 + b_1 b_2 \cos(\delta)\cos(2\theta) + \cos(\alpha_0) a_1 \left(b_1 \cos(2\theta) + b_2 \cos(\delta) \right) +$$
$$+ a_1 b_2 \sin(\alpha_0)\sin(2\theta)\sin(\delta) \Bigg] |a_1 b_1 b_2\rangle \langle a_1 b_1 b_2| \, .$$

As discussed in chapter 5, since the Eve's part of the state is uncorrelated from the rest, the conditioned Von Neumann entropy reduces to the unconditioned one

$$r_{vn} = \inf_{|\psi\rangle,\Lambda,\Pi} H(AB|E)_\rho = \inf_{|\psi\rangle,\Lambda,\Pi} H(AB)_{\rho_{AB}},$$

but ρ_{AB} is a diagonal matrix, and this implies that Von Neumann entropy reduces to the Shannon entropy:

$$\begin{aligned} r_{vn} &= 3 - \sum_{a_1,b_1,b_2} \frac{p(a_1,b_1,b_2)\log_2\big[p(a_1,b_1,b_2)\big]}{8} \\ &\text{Where } p(a_1,b_1,b_2) \equiv 1 + b_1 b_2 \cos(\delta)\cos(2\theta) + \\ &+ \cos(\alpha_0) a_1 \left(b_1 \cos(2\theta) + b_2 \cos(\delta) \right) + a_1 b_2 \sin(\alpha_0)\sin(2\theta)\sin(\delta), \end{aligned} \tag{7.38}$$

and this concludes the proof. This result is also optimal and matches the number of bits generated in the trusted case, as it can be seen by computing equation (5.5) and we can conclude that

$$r_{vn} \geq r_{me},$$

since we proved the inequality for the trusted case in section 5.1.

7.8 Constraints on the parameters

In the previous section we found a family of protocols, defined by the angles

$$\alpha_0, \alpha_1, \theta, \beta_0, \delta$$

from which we can safely generate random numbers in a device independent way. However, those parameters are not free because in the proofs we made the following assumptions:

1. The angles α_0, α_1 and β_0 are such that the operators and the state

$$\begin{aligned} &\widetilde{A}_0 = \cos(\alpha_0)\sigma_x + \sin(\alpha_0)\sigma_z; \quad \widetilde{A}_1 = \cos(\alpha_1)\sigma_x + \sin(\alpha_1)\sigma_z; \\ &\widetilde{B}_0 = \cos(\beta_0)\sigma_x + \sin(\beta_0)\sigma_z; \quad \widetilde{B}_1 = \sigma_x; \\ &|\psi\rangle = |\phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \end{aligned}$$

   are self-testable

2. The Bell's operator $S_{\theta,\delta}$ is finite, see equation (7.8) which implies the constraint

$$1 - \cos^2(\delta)\cos^2(2\theta) \neq 0.$$

3. We assumed $C_2 \neq 0$:
$$C_2 = \frac{\cos(2\theta)\sin^2(\delta)}{1-\cos^2(\delta)\cos^2(2\theta)} \neq 0 \iff \cos(2\theta)\sin^2(\delta) \neq 0$$

4. We assumed $C_3 \neq 0$:
$$C_3 = \frac{\sin(\delta)\sin(2\theta)}{1-\cos^2(\delta)\cos^2(2\theta)} \neq 0 \iff \sin(\delta)\sin(2\theta) \neq 0$$

5. The weak operator of Bob-1 $w(\sigma_x, \theta)$ is periodic with period $\frac{\pi}{4}$, therefore is not restrictive to consider
$$\theta \in \left[0, \frac{\pi}{4}\right].$$

From conditions 2. 3. 4. and 5. it follows that
$$\theta \neq 0, \quad \theta \neq \frac{\pi}{4} \implies \theta \in \left(0, \frac{\pi}{4}\right),$$
which implies that the weak measurement of Bob-1 $w(\sigma_x, \theta)$ must not be neither projective ($\theta = 0$) nor the trivial identity operator ($\theta = \frac{\pi}{4}$). From those conditions it also follows that
$$\delta \neq 0, \quad \delta \neq \pi,$$
which implies that Bob-2 operators cannot be both along σ_x, see table 7.2. Condition 1. instead, cannot be directly converted as a constraint on the three remaining angles $(\alpha_0, \alpha_1, \beta_0)$, we just have to make sure that the original non-sequential protocol is self-testable, some examples are the CHSH inequality of section (3.8) and the ones presented in [15].

7.9 Generating 3 bits of randomness

Our hope is finding a protocol, seen as choice of parameters $(\alpha_0, \delta, \theta)$, that maximizes the bits generated per round $r_{vn} = 3$. Three bits is the maximum, because from each round we have three 1-bit outcomes (the one of Bob-1, Bob-2 and Alice-1) that we could potentially use. From equation (7.38) we see that the only possibility for it to happen is that
$$\begin{aligned}&p(a_1, b_1, b_2) = 1 \quad \forall a_1, b_1, b_2 \in \{-1, +1\}^3 \iff b_1 b_2 \cos(\delta)\cos(2\theta) + \\ &+ \cos(\alpha_0) a_1 \left(b_1 \cos(2\theta) + b_2 \cos(\delta)\right) + a_1 b_2 \sin(\alpha_0)\sin(2\theta)\sin(\delta) = 0, \quad \forall a_1, b_1, b_2 \in \{-1, +1\}^3,\end{aligned} \tag{7.39}$$
which is a non linear system of 8 (dependent) equations in three variables. In particular by summing the case $a_1 = b_1 = b_2 = 1$ with $a_1 = -1, b_1 = b_2 = 1$ we obtain that
$$2\cos(\delta)\cos(2\theta) = 0 \implies \cos(\delta) = 0,$$
the other possibility $\cos(2\theta) = 0$ cannot happen since $\theta \in \left(0, \frac{\pi}{4}\right)$. Therefore, equation (7.39) reduces to
$$0 = a_1(b_1 \cos(\alpha_0)\cos(2\theta) \pm b_2 \sin(\alpha_0)\sin(2\theta)), \quad \forall a_1, b_1, b_2 \in \{0, 1\}^3$$
where the plus minus sign depends on the value of $\sin(\delta) = \pm 1$. By choosing $b_1 = \pm b_2$ and by summing and subtracting the two equations we find that
$$\begin{aligned}\cos(\alpha_0)\cos(2\theta) = 0;\\ \sin(\alpha_0)\sin(2\theta) = 0.\end{aligned}$$

But such system has no solution for $\theta \in \left(0, \frac{\pi}{4}\right)$, and this implies that generating 3 bits of randomness per round is impossible. An intuitive explanation is that we are extracting randomness from three measurements
$$\begin{aligned}&\cos(\alpha_0)\sigma_x + \sin(\alpha_0)\sigma_z;\\ &w(\sigma_x, \theta);\\ &\cos(\delta)\sigma_x + \sin(\delta)\sigma_z,\end{aligned}$$

that have components, in the 1-qubit basis of $(\mathbb{I}, \sigma_x, \sigma_y, \sigma_z)$, only along σ_x and σ_z. Hence, from the pigeonhole principle it follows that at least two measurements will have a common component (for any choice of α_0 and δ), and this is the reason for which we cannot generate 3 bits. A possible solution would be exploring a larger space in which operators can have components along σ_x,σ_y and σ_z, in such a way that we could pick the orthogonal triplet

$$\sigma_y; \quad w(\sigma_x, \theta); \quad \sigma_z,$$

from whose outcomes we could potentially extract 3 bits of randomness (we have no theoretical proofs, but numerical simulations seems to confirm).

7.10 Summary

In this chapter, we proposed a method to sequentially extend a large family of behaviors to the sequential case with three users

$$p(a, b|x, y) \xrightarrow{\text{sequential extension}} p(a, b_1, b_2|x, y_1, y_2).$$

These new behaviors, in particular, are guaranteed to saturate two sets of Tsirelson inequalities:

- The first set is symbolic, meaning that its explicit form depends on the specific initial non-sequential protocol used. It can be expressed as:

$$\langle S_k \rangle \leq I_k, \quad k \in \mathbb{N};$$

- The second set consists of a single element, previously defined in equation (7.8):

$$\langle S_{\theta,\delta} \rangle \leq 0.$$

  In particular, this inequality holds only if the first set is saturated, meaning:

$$\langle S_k \rangle = I_k, \quad k \in \mathbb{N},$$

  and the sequentiality conditions, defined in equation (6.1), is satisfied.

From the saturation of the two sets we managed to partially self-test the system and derive analytical results on device-independent generation of random bits, using either the min-entropy (7.34), or the Von Neumann entropy (7.38). Both results will be tested in the next chapter with numerical simulations. Finally we proved that each behavior can be realized in a relatively simple way, by using 1-qubit real operators and the maximally entangled state, as shown in table 7.2.

Chapter 8

Numerical simulations

In this chapter we will describe from a more practical point of view how numerical simulations are done. The final aim is developing a program that can simulate the number of device-independent bits generated per round given the initial conditions of a generic protocol, that are

1. The number of Alice and Bob;
2. The number of inputs and outputs of each user;
3. The observables from which we want to extract randomness;
4. The observed probability distribution $p(\vec{a},\vec{b}|\vec{x},\vec{y})$

Then, we could use such programs for many different applications: for example verifying numerically the results of the previous chapter, in particular equations (7.34) and (7.38) or we could even go beyond and check what happens if we add noise to the system. The standard way of doing so, is slightly perturbing the shared wavefunction

$$\rho = (1-p)\left|\phi_+\right\rangle\left\langle\phi_+\right| + p\mathbb{I}, \quad p \in [0,1]. \tag{8.1}$$

p is a parameter that quantify the noise: as it increases the system loses entanglement and therefore its quantum properties. Eventually, by increasing p we change the behavior $p(\vec{a},\vec{b}|\vec{x},\vec{y})$, and in particular we no longer maximally violate the Bell's inequalities on which the self-testing proof was based. Therefore, we surely expect the number of bits generated per round r_{me} to decrease, and this is the most we manage to say theoretically. The easiest way to actually quantify the function $r_{me}(p)$ is with numerical simulations.

8.1 Recap of the problem

For simplicity let us focus on the min entropy r_{me} and consider directly the case with 1 Alice and 2 Bob with binary inputs and outputs, that is the scenario of the protocol we developed. We use the sequential characterization described in (6.4), which is particularly suited for numerical simulations. So Alice has projectors $\Lambda^{x_1}_{a_1}$, that quantify the probability of having output $a_1 \in \{\pm 1\}$ given the choice of the input $x_1 \in \{0,1\}$, and similarly Bobs have unified projectors $\Pi^{y_1 y_2}_{b_1 b_2}$, whose expectation values are the probability of Bob-1,Bob-2 measuring b_1, b_2 given the input choice y_1, y_2. Furthermore, without loss of generality randomness is extracted from the outcomes of input labeled with $x_1 = y_1 = y_2 = 0$. With that said, the problem we want to solve is computing

$$r_{me} = -\log_2(G),$$

where G is the guessing probability, defined as

$$\begin{aligned}
&G \equiv \max_{|\psi\rangle,\Lambda,\Pi,E}\left(\sum_{a_1,b_1,b_2}\langle\psi|\,\Lambda^0_{a_1}\Pi^{00}_{b_1b_2}E_{a_1b_1b_2}]\,|\psi\rangle\right)\\
&\text{such that } \langle\psi|\,\Lambda^{x_1}_{a_1}\Pi^{y_1y_2}_{b_1b_2}\,|\psi\rangle = p(a_1,b_1,b_2|x_1,y_1,y_2);\\
&\text{such that } E_{e_1e_2e_3}\succcurlyeq 0; \quad E_{e_1e_2e_3}E_{e_4e_5e_6} = \delta_{e_1e_4}\delta_{e_2e_5}\delta_{e_3e_6}E_{e_1e_2e_3}; \quad E^\dagger_{e_1e_2e_3} = E_{e_1e_2e_3};\\
&\text{such that } \Lambda^x_a\succcurlyeq 0; \quad \Lambda^x_{a_1}\Lambda^x_{a_2} = \delta_{a_1a_2}\Lambda^x_{a_1}; \quad (\Lambda^x_a)^\dagger = \Lambda^x_a;\\
&\text{such that } \Pi^{y_1y_2}_{b_1b_2}\succcurlyeq 0; \quad \Pi^{y_1y_2}_{b_1b_2}\Pi^{y_1y_2}_{b_3b_4} = \delta_{b_1b_3}\delta_{b_2b_4}\Pi^{y_1y_2}_{b_1b_2}; \quad (\Pi^{y_1y_2}_{b_1b_2})^\dagger = \Pi^{y_1y_2}_{b_1b_2},\\
&\text{such that } \Pi^{y_1y_2}_{b_1b_2}\Pi^{y_1y_3}_{b_3b_4} = 0, \quad \text{with } b_1\neq b_3;\\
&\text{such that } [\Lambda^x_a,\Pi^{y_1y_2}_{b_1b_2}] = [\Lambda^x_a,E_{e_1e_2e_3}] = [\Pi^{y_1y_2}_{b_1b_2},E_{e_1e_2e_3}] = 0;\\
&\text{such that } \sum_{e_1e_2e_3}E_{e_1e_2e_3} = \mathbb{I}, \quad \sum_{a_1}\Lambda^{x_1}_{a_1} = \mathbb{I}, \quad \forall x_1\in\{0,1\}, \quad \sum_{b_1b_2}\Pi^{y_1y_2}_{b_1b_2} = \mathbb{I}, \quad \forall y_1,y_2\in\{0,1\};\\
&\text{such that } \sum_{b_2}\Pi^{x_1x_2}_{b_1b_2} = \sum_{b_2}\Pi^{x_1x_3}_{b_1b_2}, \quad \forall x_1,x_2,x_3\in\{0,1\}, b_1\in\{\pm1\}, x_2\neq x_3.
\end{aligned} \tag{8.2}$$

Which is equation (5.15) generalized to this particular sequential case. Notice that we have also included the completeness relations

$$\begin{aligned}
&\sum_{e_1e_2e_3}E_{e_1e_2e_3} = \mathbb{I};\\
&\sum_{a_1}\Lambda^{x_1}_{a_1} = \mathbb{I}, \quad \forall x_1\in\{0,1\};\\
&\sum_{b_1b_2}\Pi^{y_1y_2}_{b_1b_2} = \mathbb{I}, \quad \forall y_1,y_2\in\{0,1\};
\end{aligned}$$

and the sequential linear constraints, see equation (6.4)

$$\sum_{b_2}\Pi^{x_1x_2}_{b_1b_2} = \sum_{b_2}\Pi^{x_1x_3}_{b_1b_2}, \quad \forall x_1,x_2,x_3\in\{0,1\}, b_1\in\{\pm1\}, x_2\neq x_3,$$

because for the sake of this example, we want to keep everything as simple as possible. Indeed, exploiting linear equations to build the reduced set and remove the redundancy, as explained in section 4.2, is a nice optimization but has the disadvantage of making the code more complex, since we have to deal with substitutions and extra steps. As we already discussed in chapter 5, the problem (8.2) is solvable order by order with the NPA hierarchy, where each order n is a SDP of the form

$$\begin{aligned}
&\text{maximize } \operatorname{Tr}[\beta^n\Gamma^n];\\
&\text{subject to } \operatorname{Tr}[F^T\Gamma^n] = g(p), \quad \forall F,g\in\mathcal{F}(\mathcal{O});\\
&\text{subject to } \Gamma^n\succcurlyeq 0,
\end{aligned} \tag{8.3}$$

where Γ^n is a n-th order certificate, defined in equation (4.13). The matrix β^n satisfy

$$\operatorname{Tr}[\beta^n\Gamma^n] = \sum_{a_1b_1b_2}p(a_1,b_1,b_2|x_1,y_1,y_2) \tag{8.4}$$

and the requirements

$$\operatorname{Tr}[F^T\Gamma^n] = g(p), \quad \forall F,g\in\mathcal{F}(\mathcal{O}), \tag{8.5}$$

are equivalent to the constraints of (8.2), rewritten as trace of product of matrices. There are many programs that can solve such SDP, for example SDPA.

8.2 Converting to SDP

The difficult part of the above discussion is converting the maximization problem (8.2) of non-commutative operators to the semi definite programming form of equation (8.3), in particular:

- First of all we need to generate the whole set S_n, as defined in (5.14);
- Then, as shown in section 4.7, we need to map the indices of Γ^n to the expectation values of operators in S_n, and especially for large n the task is not easy;
- Finally we need to rewrite each constraint of equation (8.2) to their matrix form (8.5) and similarly we need to find a matrix β^n that satisfy (8.4);
- On top of this, the resulting SDP (8.3) has to be in a valid format that can be inputted to SDPA.

Luckily, such algorithm has already been created [14], and implemented in a python library called ncpol2sdpa, that does this conversion for any generic polynomial optimization problem of operators, and inputs the result directly to SDPA. Therefore, the only missing link is writing a program that generates:

1. The list of operators $\Lambda^x_a, \Pi^{y_1 y_2}_{b_1 b_1}, E_{e_1 e_2 e_3}$;
2. Their whole set of constraints, listed in equation (8.2), including the one with the observed probability distribution;
3. The target operator to maximize

$$\sum_{a_1, b_1, b_2} \Lambda^0_{a_1} \Pi^{00}_{b_1 b_2} E_{a_1 b_1 b_2},$$

in a valid format that can be inputted to ncpol2sdpa.

8.3 Concrete implementation

Let's see how each step can be concretely implemented in python. As we said the first step is defining the list of projectors

$$\Lambda^x_a; \quad \Pi^{y_1 y_2}_{b_1 b_1}; \quad E_{e_1 e_2 e_3},$$

which is done by using HermitianOperator objects of the sympy library:

```python
from sympy.physics.quantum import HermitianOperator

# map with Alice's projectors
alice_projectors = {}

# Loop over all possible inputs 0,1
for x in [0,1]:
    # Loop over all possible outputs -1,+1
    for a in [-1,1]:
        # For each input and output create an Hermitian operator named L_x_a
        projector_name = "L_{}_{}"
        # and store it in the map.
        alice_projectors[(x,a)]= HermitianOperator(projector_name.format(x,a))

print(alice_projectors)
```

Listing 8.1: Creating Alice's projectors

Which prints the following

```
{(0, -1): L_0_-1, (0, 1): L_0_1, (1, -1): L_1_-1, (1, 1): L_1_1},
```

so basically we have built a dictionary that maps

$$(x, a) \to \Lambda_a^x.$$

In a similar way we define Bob and Eve projectors

```
# map with Bob1 and Bob2 projectors
bob_projectors = {}
... # derivation skipped since is the same as Alice

# map with Eve's projectors
eve_projectors = {}
...
```

in such a way that bob_projectors is a dictionary that maps

$$(y_1, y_2, b_1, b_2) \to \Pi_{b_1 b_2}^{y_1 y_2},$$

and analogously eve_projectors maps

$$(e_1, e_2, e_3) \to E_{e_1,e_2,e_3}.$$

At this point we pass to step 2), which consists in generating the constraints of equation (8.2), for instance

$$\Lambda_{a_1}^x \Lambda_{a_2}^x = 0, \quad \text{for } a_1 \neq a_2,$$

can be implemented as:

```
# map with Alice's projectors
alice_projectors = []
# ... previous code snippet to fill alice_projectors ...

# list with constraints
constraints = {}
# loop over all Alice's inputs
for x1 in [0,1]:
    # Loop over all the possible different outputs a1 = 1,-1 and a2 != a1
    for a1 in [-1,1]:
        for a2 in {-1,1}.difference({a1}):
            # fetch the two projectors
            op1 = alice_projectors[(x1,a1)]
            op2 = alice_projectors[(x1,a2)]
            # insert the constraint op1*op2 = 0 in the list
            constraints.append(op1*op2)

print(constraints)
```

Listing 8.2: Orthogonality Alice constraints

The code prints

```
[ L_0_-1 * L_0_1 ,  L_0_1 * L_0_-1 ,  L_1_-1 * L_1_1 ,  L_1_1 * L_1_-1 ]
```

which is the expected result. Notice that we are implicitly assuming that each constraint is zero

```
constraints[j] = 0,  with  0 <= j < len(constraints)
```

because it is the format required by ncpol2sdpa. So for example, the completeness relations

$$\mathbb{I} = \sum_{a_1} \Lambda_{a_1}^{x_1}$$

have to be implemented in the form

$$\mathbb{I} - \sum_{a_1} \Lambda_{a_1}^{x_1} = 0,$$

as shown in the following code snippet:

```python
# list with constraints
constraints = []
# loop over all alice projectors
for x in [0,1]:
    # constraint 1 - sum_of_projectors = 0
    expr = 1
    for a in [-1,1]:
        expr -= alice_projectors[(x,a)]
    constraints.append(expr)
```

Listing 8.3: Alice projectors completeness relation

All other constraints are derived in a similar way, and we are not showing their implementation. The last and final step is generating the objective we want to maximize

$$\sum_{a_1,b_1,b_2} \Lambda^0_{a_1} \Pi^{00}_{b_1 b_2} E_{a_1 b_1 b_2},$$

which is done quite easily

```python
# objective to maximize
objective = 0
# loop over all possible outcomes of alice bob-1 and bob-2
for a1 in [-1,1]:
    for b1 in [-1,1]:
        for b2 in [-1,1]:
            # fetch the projectors
            alice_p = alice_projectors[(0,a1)]
            eve_p = eve_projectors[(a1,b1,b2)]
            bob_p = bob_projectors[(0,0,b1,b2)]
            # sum their product to the objective
            objective += alice_p*bob_p*eve_p
```

Listing 8.4: Objective to maximize

Finally we pass everything to the ncpol2sdpa library:

```python
import ncpol2sdpa as ncp

# list with all projectors of all users
op = [*eve_projectors.values(), *alice_projectors.values(), *bob_projectors.values()]
# level of the NPA hierarchy
npa_level = 2
# SDP solver
solver = 'sdpa'

# perform relaxation and solve
sdp = ncp.SdpRelaxation(op, normalized=True)
sdp.get_relaxation(level = npa_level,
                   momentequalities = constraints,
                   objective = -objective)
sdp.solve(solver)

# print the result (guessing probability G)
print(-sdp.dual)
```

Listing 8.5: Calling ncpol2sdpa

8.4 Intermediate orders

As explained in section 4.7 the running time needed to find a valid certificate at second order of the NPA hierarchy could be too big. In such cases it is convenient considering an intermediate level, such as

$$1 + AB + AE + BE.$$

Concretely, it is done by storing the extra monomials we want to add in a list

```python
extra_monomials = []
# loop over all alice and bob projectors
for a_op in alice_projectors.values():
    for b_op in bob_projectors.values():
        # store the monomials of the form A*B
        extra_monomials.append(a_op*b_op)

# ... To the same for AE + BE ...
```

Listing 8.6: generating AB monomials

that is passed to ncpol2sdpa as parameter when performing the relaxation

```python

# ... previous code snippets ...

# notice that this time the level is 1
sdp.get_relaxation(level = 1,
                   momentequalities = constraints,
                   objective = -objective
                   # extra monomials are added here
                   extramonomials = extra_monomials)

# ... solve and print the result ...
```

Listing 8.7: Calling ncpol2sdpa

8.5 A possible optimization

The example we made is simple and outputs the correct result, however it is not optimized. Indeed, apart from using reduced sets, ncpol2sdpa allows us to further reduce the number of variables with substitutions: a substitution is a special constraint of the form

$$\text{product of operators} = \text{product of operators},$$

so for example

$$\begin{aligned} &\Lambda^x_{a_1}\Lambda^x_{a_2} = \Lambda^x_{a_1}\delta_{a_1 a_2}, \quad \forall a_1, a_2 \in \{-1,1\}, x \in \{0,1\}, \\ &\Lambda^x_a \Pi^y_b = \Pi^y_b \Lambda^x_a \end{aligned}$$

are substitutions, while

$$\mathbb{I} = \Lambda^x_1 + \Lambda^x_{-1}, \quad \forall x \in \{0,+1\}$$

are not. So the idea is that, at the n-th order of the NPA hierarchy, we apply each substitution to all elements of S_n. This will result in a new set $\widetilde{S}_n$ with fewer elements (for example all variables that contains $\Lambda^x_{a_1}\Lambda^x_{a_2} = 0$ are removed), hence the corresponding matrix Γ^n is smaller in size and in particular finding a valid certificate is easier. Concretely, substitutions are inserted in a map, as we show in the following code snippet where we implement Alice's idempotence relations:

```python
# map with substitutions
subs = {}
# insert the idempotence relations for Alice
for x1 in [0,1]:
    for a1 in [-1,1]:
        alice_op = alice_projectors[(x1,a1)]
        subs[alice_op*alice_op] = alice_op
print(subs)
```

Listing 8.8: Alice idempotence as substitution

which prints

```
{ L_0_-1 **2:  L_0_-1 ,  L_0_1 **2:  L_0_1 ,
  L_1_-1 **2:  L_1_-1 ,  L_1_1 **2:  L_1_1 }.
```

All other substitutions are implemented in a similar way and we are not showing the implementation. Finally, the map is passed as parameter when performing the relaxation

```python
import ncpol2sdpa as ncp

# ... as previous code snippet

# perform relaxation and solve
sdp = ncp.SdpRelaxation(op, normalized=True)
sdp.get_relaxation(level = npa_level,
                   substitutions = subs,
                   # constriants no longer contains substitutions
                   momentequalities = constraints,
                   objective = -objective)
sdp.solve(solver)

# print the result (guessing probability G)
print(-sdp.dual)
```

Listing 8.9: Passing substitutions to ncpol2sdpa

and the library performs the optimization internally.

8.6 Adapting to Von Neumann entropy

The procedure we just showed can be easily adapted to simulating the number of bits generated per round with the Von Neumann entropy r_{vn}. Indeed, as described in section 5.3, we have to solve the following optimization problem

$$
\begin{aligned}
& c_m + \sum_{i=1}^{m-1} \inf \left(\frac{w_i}{t_i \ln 2} \sum_{a,b_1,b_2} \langle\psi| \Lambda_a^0 \Pi_{b_1b_2}^{00} (Z_{ab_1b_2} + Z_{ab_1b_2}^\dagger + (1-t_i) Z_{ab_1b_2}^\dagger Z_{ab_1b_2}) + t_i Z_{ab_1b_2} Z_{ab_1b_2}^\dagger |\psi\rangle \right) \\
& \text{such that } \langle\psi| \Lambda_a^x \Pi_{b_1b_2}^{y_1y_2} |\psi\rangle = p(a,b_1,b_2|x,y_1,y_2); \\
& \text{such that } \Lambda_a^x \succcurlyeq 0; \quad \Lambda_{a_1}^x \Lambda_{a_2}^x = \delta_{a_1a_2} \Lambda_{a_1}^x; \quad (\Lambda_a^x)^\dagger = \Lambda_a^x; \\
& \text{such that } \Pi_{b_1b_2}^{y_1y_2} \succcurlyeq 0; \quad \Pi_{b_1b_2}^{y_1y_2} \Pi_{b_3b_4}^{y_1y_2} = \delta_{b_1b_3}\delta_{b_2b_4} \Pi_{b_1b_2}^{y_1y_2}; \quad (\Pi_{b_1b_2}^{y_1y_2})^\dagger = \Pi_{b_1b_2}^{y_1y_2}, \\
& \text{such that } \Pi_{b_1b_2}^{y_1y_2} \Pi_{b_3b_4}^{y_1y_3} = 0, \quad \text{with } b_1 \neq b_3; \\
& \text{such that } [\Lambda_a^x, \Pi_{b_1b_2}^{y_1y_2}] = [\Lambda_a^x, Z_{cb}] = [\Pi_{b_1b_2}^{y_1y_2}, Z_{ac}] = [\Lambda_a^x, Z_{cb}^\dagger] = [\Pi_{b_1b_2}^{y_1y_2}, Z_{ac}^\dagger] = 0 \\
& \text{such that } \sum_{a_1} \Lambda_{a_1}^{x_1} = \mathbb{I}, \quad \forall x_1 \in \{0,1\}, \quad \sum_{b_1b_2} \Pi_{b_1b_2}^{y_1y_2} = \mathbb{I}, \quad \forall y_1, y_2 \in \{0,1\}; \\
& \text{such that } \sum_{b_2} \Pi_{b_1b_2}^{x_1x_2} = \sum_{b_2} \Pi_{b_1b_2}^{x_1x_3}, \quad \forall x_1, x_2, x_3 \in \{0,1\}, b_1 \in \{\pm 1\}, x_2 \neq x_3. \\
& \text{where } c_m = \sum_{i=1}^{m-1} \frac{w_i}{t_i \ln 2}; \\
& \text{where } m \in \mathbb{N} \text{ and } t_i, w_i \text{ are the nodes and weight of a } m\text{-point Gauss-Radau rule with } t_m = 1,
\end{aligned} \tag{8.6}
$$

which is basically equation (5.17), with the swap approximation (5.19) and without the bounded condition (5.18), adapted to our particular case with 1 Alice and 2 Bobs with two binary inputs and outputs. As before we are considering also the completeness and sequential linear equations, in order to keep the code as simple as possible. We have the following similarities with the min-entropy case, see equation (8.2):

- The number of Alice and Bobs projectors and their constraints are the same in both cases;
- In both cases the only constraint between Alice/Bobs and Eve operators is that they commute.

Therefore, we can reuse the code previously developed and change only the objective function and Eve's operators, which are now non-hermitian. Concretely, the latter can be implemented as follows

```python
from sympy.physics.quantum import HermitianOperator, Operator
from sympy.physics.quantum.dagger import Dagger
alice_projectors = {}
bob_projectors = {}

# ... alice and bob projectors maps are filled as before ...

# maps to store Eve's operators and their dagger
eve_operators = {}
eve_operators_dagger = {}
for i in [-1,1]:
    for j in [-1,1]:
        for k in [-1,1]:
            projector_name = "Z_{}_{}_{}"
            new_op = Operator(projector_name.format(i,j,k))
            eve_operators[(i,j,k)]= new_op
            eve_operators_dagger[(i,j,k)] = Dagger(new_op)
```

Listing 8.10: Operators for the Von Neumann case

So eve_projectors maps

$$(a_1, b_1, b_2) \to Z_{a_1 b_1 b_2},$$

and similarly eve_projectors_dagger maps

$$(a_1, b_1, b_2) \to Z^{\dagger}_{a_1 b_1 b_2}.$$

To generate the objective function instead, we need first of all a library to compute Gauss-Radau coefficients t_i and w_i for any given $m \in \mathbb{N}$. There are many possibilities, we decided to use *Chaospy* that does the task in a few lines of code:

```python
import chaospy
def generate_quadrature(m):
    t, w = chaospy.quad_gauss_radau(m, chaospy.Uniform(0,1), 1)
    return t[0], w

m = 4
t, w = generate_quadrature(m)
```

Listing 8.11: Generating Gauss-Radau coefficients with Chaospy

the output of generate_quadrature are two lists t,w of length m such that

$$t[i] = t_{i+1}, \quad w[i] = w_{i+1}, \quad \forall i \in 0, ..., m-1.$$

At this point we have everything needed to generate the objective function

```python
import numpy as np

# ... Define operator maps as in previous code snippets ...

m = 4
t, w = generate_quadrature(m)

# k integer in the range [0, m-1]
def generate_objective(k):
    res = 0
    ck = w[k]/(t[k]*np.log(2))
    # loop over all possible outputs
    for a1 in [-1, 1]:
        for b1 in [-1, 1]:
            for b2 in [-1,1]:
                # extract the operators from the maps
                a_p = alice_projectors[(0,a1)]
                b_p = bob_projectors[(0, 0, b1, b2)]
                e_p = eve_operators[(a1, b1, b2)]
                e_p_d = eve_operators_dagger[(a1, b1, b2)]
                # update the result
                res += a_p*b_p*(e_p + e_p_d + (1-t[k])*e_p_d*e_p)
                res += t[k]*e_p*e_p_d
    return ck*res
```

Listing 8.12: Generating the objective function

and finally we perform m consecutive SDP and sum their results

```
import ncpol2sdpa as ncp

# ... previous code snippets ...

constraints = []
# ... find all constraints as in the min-entrop case ...

# list with all operators of all users (no need to insert eve_projectors_dagger)
op = [*eve_operators.values(), *alice_projectors.values(), *bob_projectors.values()]

sdp = ncp.SdpRelaxation(op, normalized = True)
sdp.get_relaxation(level = 2, momentequalities = constraints)
res = 0.0
for k in range(m):
    ck = w[k]/(t[k]*np.log(2))
    # update the objective and solve
    sdp.set_objective(generate_objective(k))
    sdp.solve('sdpa')
    res += (ck + sdp.dual)
# print the result (lower bound on r_vn)
print(res)
```

Listing 8.13: Finding lower bounds on r_{vn}

the printed result is a lower bound on r_{vn}. As in the min-entropy case, the running time and memory needed to solve the problem at second order in NPA hierarchy could be too big. In such cases a possibility is considering an intermediate level, for example

$$1 + AB + AZ + AZ^{\dagger} + BZ + BZ^{\dagger} + ZZ^{\dagger},$$

or we could even consider higher order terms that appear in (8.6), such as

$$1 + AB + AZ + AZ^{\dagger} + BZ + BZ^{\dagger} + ZZ^{\dagger} + ABZ + ABZ^{\dagger} + ABZZ^{\dagger} \tag{8.7}$$

8.7 Examples

In chapter 7 we quantified the device-independent rate of extraction of random bits, equations (7.34) and (7.38), for a large family of sequential protocols. In this section we will verify some particular cases by using the numerical tools developed and also simulate what happens if we add noise.

Protocol 1 We begin with the sequential extension of the following protocol with 1-Alice and 1-Bob and operators

$$\begin{aligned} A_0 &= \sigma_z; \quad B_0 = \sigma_x; \\ A_1 &= -\sin(\epsilon)\sigma_z + \cos(\epsilon)\sigma_x; \\ B_1 &= \cos(\epsilon)\sigma_z - \sin(\epsilon)\sigma_x; \\ &\text{where } \epsilon \in \left(0, \frac{\pi}{6}\right], \end{aligned}$$

which are measured on the shared state

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

It has been proposed in [15], where they proved self-testing results for state and operators, and showed that it's possible to generate 2 bits of randomness per round from the outcomes of A_0 and B_0 for any value of ϵ. By following table (7.2), we propose the following sequential extension

Person	Measurement 1	Measurement 2
Alice 1	σ_z	$\cos(\epsilon)\sigma_x - \sin(\epsilon)\sigma_z$
Bob 1	$w(\sigma_x, \theta)$	$-\sin(\epsilon)\sigma_x + \cos(\epsilon)\sigma_z$
Bob 2	σ_z	σ_x

Table 8.1: Operators for protocol 1.

where we have fixed the parameter $\delta = \frac{\pi}{2}$. From equation (7.34) we derive the expected number of bits generated per round with the min entropy

$$r_{me} = 3 - \log_2 \left[1 + \sin(2\theta)\right], \quad \theta \in \left(0, \frac{\pi}{4}\right), \tag{8.8}$$

which asymptotically converges to 3 as $\theta \to 0$. Notice in particular that r_{me} doesn't depend on ϵ but only on the strength of the weak measurement θ. We fixed $\epsilon = \frac{\pi}{12}$ and ran many simulations at intermediate NPA level $1 + AB$ for different values of θ and different values of the noise p, see equation (8.1), and plotted the results in figure 8.1. We can conclude that the noiseless case fits very well equation (8.8) and that r_{me} decays really fast as p increases.

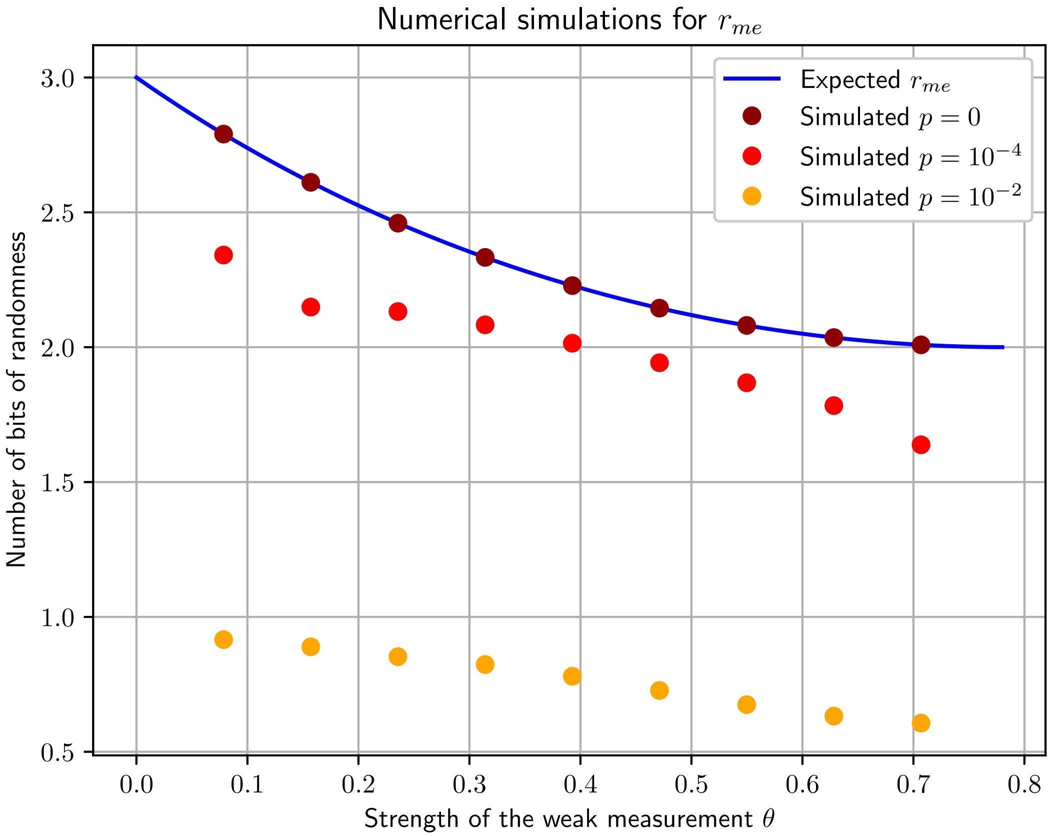

Figure 8.1: Numerical simulations for protocol 1

Notice that the sequential extension is always convenient, since from equation (8.8) follows that

$$r_{me} > 2, \quad \forall \theta \in \left(0, \frac{\pi}{4}\right).$$

The expected rate of extraction with the Von Neumann entropy can be found from equation (7.38)

$$r_{vn} = 3 - \frac{1+\sin(2\theta)}{2}\log_2\big[1+\sin(2\theta)\big] - \frac{1-\sin(2\theta)}{2}\log_2\big[1-\sin(2\theta)\big],$$

due to the higher running time, we ran only three noiseless simulations, with the intermediate NPA level (8.7) and Gauss-Radau with four points $m = 4$. Results are plotted in figure 8.2. The lowers bounds found are not very tight, but they're regardless above the min-entropy curve (for $\theta = 0.1$ and $\theta = 0.4$):

$$\Delta r(\theta = 0.1) \equiv r_{vn}(0.1) - r_{me}(0.1) \approx 0.13;$$
$$\Delta r(\theta = 0.4) \equiv r_{vn}(0.4) - r_{me}(0.4) \approx 0.19.$$

For $\theta = 0.7$ instead it is slightly below the min-entropy curve:

$$\Delta r(\theta = 0.7) \equiv r_{vn}(0.7) - r_{me}(0.7) \approx -0.03.$$

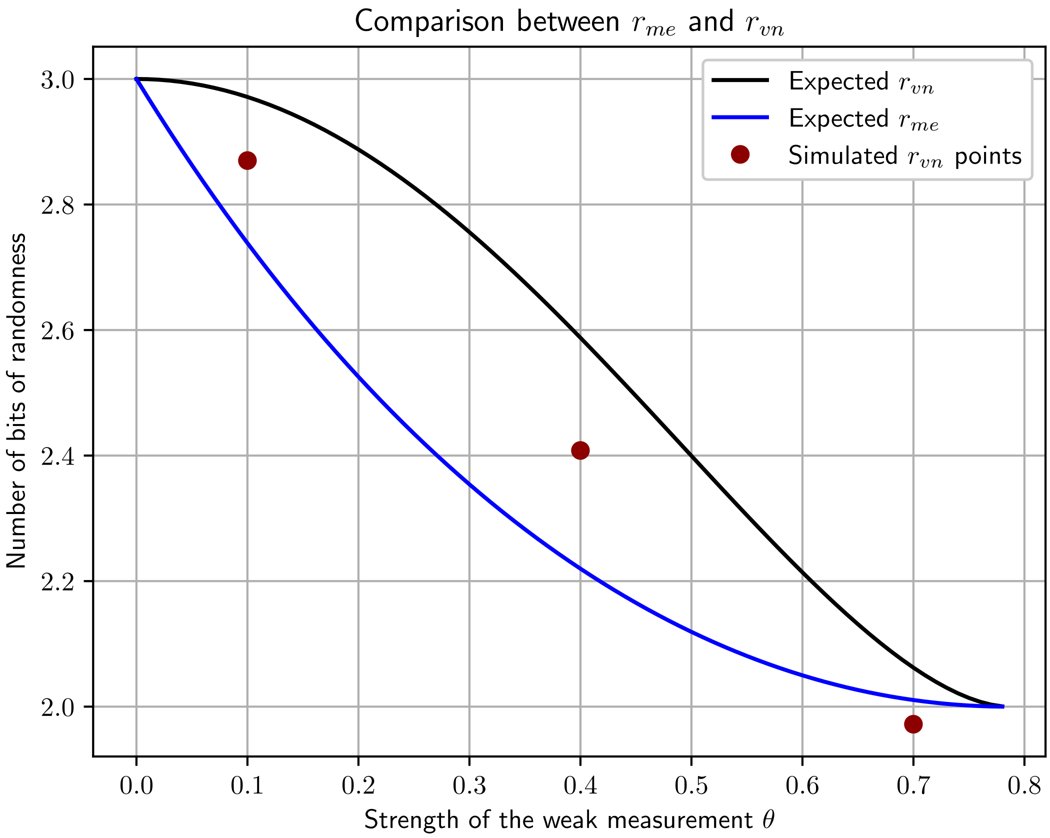

Figure 8.2: Comparison between r_{me} and r_{vn} for protocol 1

Protocol 2 The second protocol we consider is an extension of the CHSH. Recall that the non-sequential version consists in 1-Alice and 1-Bob measuring the following operators

$$A_0 = \frac{\sigma_z + \sigma_x}{\sqrt{2}};$$
$$A_1 = \frac{\sigma_x - \sigma_z}{\sqrt{2}};$$
$$B_0 = \sigma_z; \quad B_1 = \sigma_x,$$

measured on the shared state

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

which is equation (2.7) with the two users swapped. In section 3.8 we proved self-testing results for both state and measurements, therefore we can apply what we found in chapter 7 and the sequential extension follows from table (7.2):

Person	Measurement 1	Measurement 2
Alice 1	$\frac{1}{\sqrt{2}}(\sigma_z + \sigma_x)$	$\frac{1}{\sqrt{2}}(\sigma_x - \sigma_z)$
Bob 1	$w(\sigma_x, \theta)$	σ_z
Bob 2	σ_z	σ_x

Table 8.2: Operators for protocol 2.

Where again we fixed $\delta = \frac{\pi}{2}$. This protocol has been already studied in [10], in the case of local randomness extraction (i.e. random bits are generated only from outcomes of Bob-1 and Bob-2) and it has been found $r_{me} = 2$. From equations (7.34) we derive that the extraction rate is

$$r_{me} = 3 - \log_2 \left[1 + \frac{1}{\sqrt{2}} \left(\sin(2\theta) + \cos(2\theta) \right) \right], \quad \text{with } \theta \in \left(0, \frac{\pi}{4}\right). \tag{8.9}$$

Notice that $r_{me} \geq 2$ and therefore extracting randomness globally is convenient. As in the first protocol we simulated its expression at NPA level $1 + AB$ for different values of θ and noise level p. Results are plotted in figure 8.3 and we can conclude that the noiseless case fits very well equation (8.9).

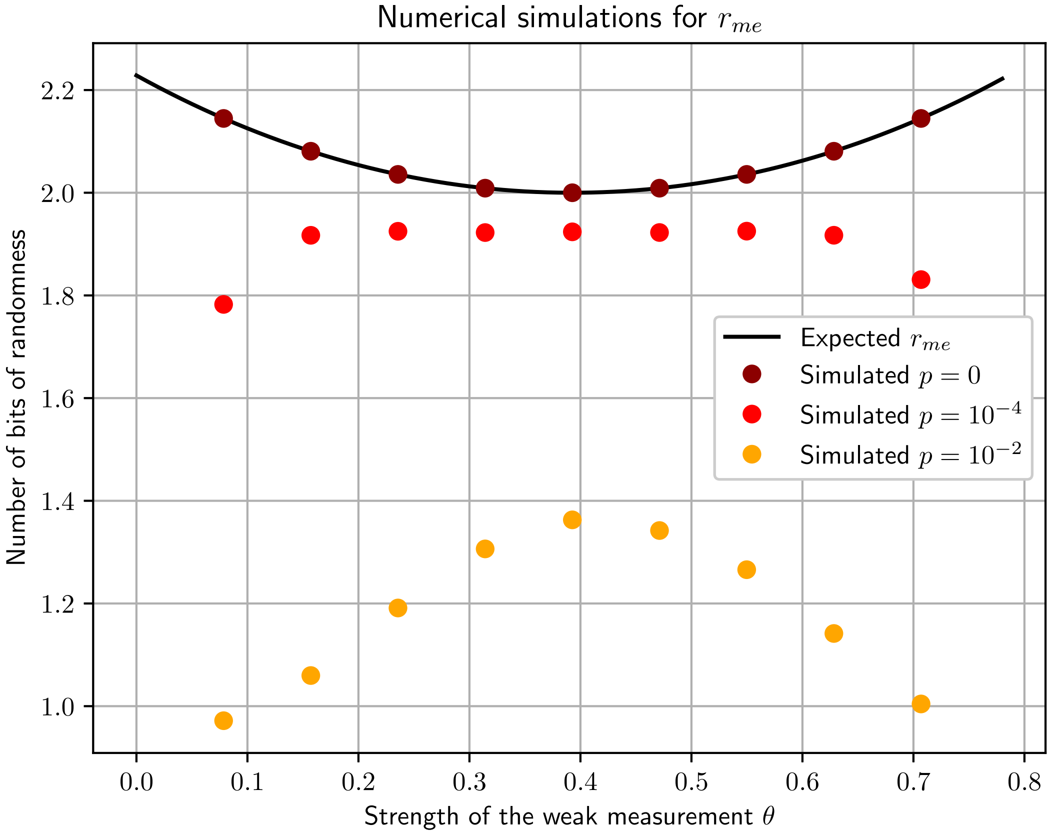

Figure 8.3: Numerical simulations for protocol 2

Bibliography

[1] J. S. Bell. "On the Einstein-Podolsky-Rosen paradox". In: *Physics Physique Fizika* 1 (1964), pp. 195–200. DOI: 10.1103/PhysicsPhysiqueFizika.1.195.

[2] Rutvij Bhavsar, Sammy Ragy, and Roger Colbeck. "Improved device-independent randomness expansion rates using two sided randomness". In: *New Journal of Physics* 25.9 (Sept. 2023), p. 093035. ISSN: 1367-2630. DOI: 10.1088/1367-2630/acf393. URL: http://dx.doi.org/10.1088/1367-2630/acf393.

[3] Joseph Bowles, Flavio Baccari, and Alexia Salavrakos. "Bounding sets of sequential quantum correlations and device-independent randomness certification". In: *Quantum* 4 (Oct. 2020), p. 344. ISSN: 2521-327X. DOI: 10.22331/q-2020-10-19-344. URL: http://dx.doi.org/10.22331/q-2020-10-19-344.

[4] Peter Brown, Hamza Fawzi, and Omar Fawzi. *Device-independent lower bounds on the conditional von Neumann entropy.* 2023. arXiv: 2106.13692 [quant-ph]. URL: https://arxiv.org/abs/2106.13692.

[5] Nicolas Brunner et al. "Bell nonlocality". In: *Reviews of Modern Physics* 86.2 (Apr. 2014), pp. 419–478. ISSN: 1539-0756. DOI: 10.1103/revmodphys.86.419. URL: http://dx.doi.org/10.1103/RevModPhys.86.419.

[6] Paul Busch et al. "Dilation Theory". In: *Quantum Measurement.* Cham: Springer International Publishing, 2016, pp. 137–162. ISBN: 978-3-319-43389-9. DOI: 10.1007/978-3-319-43389-9_7. URL: https://doi.org/10.1007/978-3-319-43389-9_7.

[7] A. Einstein, B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" In: *Phys. Rev.* 47 (10 May 1935), pp. 777–780. DOI: 10.1103/PhysRev.47.777. URL: https://link.aps.org/doi/10.1103/PhysRev.47.777.

[8] M. Navascués et al. "A Physical Approach to Tsirelson's Problem". In: *Foundations of Physics* 42.8 (Mar. 2012), pp. 985–995. ISSN: 1572-9516. DOI: 10.1007/s10701-012-9641-0. URL: http://dx.doi.org/10.1007/s10701-012-9641-0.

[9] Miguel Navascués, Stefano Pironio, and Antonio Acín. "A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations". In: *New Journal of Physics* 10.7 (July 2008), p. 073013. ISSN: 1367-2630. DOI: 10.1088/1367-2630/10/7/073013. URL: http://dx.doi.org/10.1088/1367-2630/10/7/073013.

[10] Matteo Padovan et al. *Geometry of sequential quantum correlations and robust randomness certification.* 2023. arXiv: 2309.12286 [quant-ph]. URL: https://arxiv.org/abs/2309.12286.

[11] Itamar Pitowsky. "The range of quantum probability". In: *Journal of Mathematical Physics* 27.6 (June 1986), pp. 1556–1565. ISSN: 0022-2488. DOI: 10.1063/1.527066. eprint: https://pubs.aip.org/aip/jmp/article-pdf/27/6/1556/19154432/1556_1_online.pdf. URL: https://doi.org/10.1063/1.527066.

[12] Florian A. Potra and Stephen J. Wright. "Interior-point methods". In: *Journal of Computational and Applied Mathematics* 124.1 (2000). Numerical Analysis 2000. Vol. IV: Optimization and Nonlinear Equations, pp. 281–302. ISSN: 0377-0427. DOI: https://doi.org/10.1016/S0377-0427(00)00433-7. URL: https://www.sciencedirect.com/science/article/pii/S0377042700004337.

[13] Marco Tomamichel. *A Framework for Non-Asymptotic Quantum Information Theory.* 2013. arXiv: 1203.2142 [quant-ph]. URL: https://arxiv.org/abs/1203.2142.

[14] Peter Wittek. “Algorithm 950: Ncpol2sdpa—Sparse Semidefinite Programming Relaxations for Polynomial Optimization Problems of Noncommuting Variables”. In: *ACM Transactions on Mathematical Software* 41.3 (June 2015), pp. 1–12. ISSN: 1557-7295. DOI: 10.1145/2699464. URL: http://dx.doi.org/10.1145/2699464.

[15] Lewis Wooltorton, Peter Brown, and Roger Colbeck. “Tight Analytic Bound on the Trade-Off between Device-Independent Randomness and Nonlocality”. In: *Physical Review Letters* 129.15 (Oct. 2022). ISSN: 1079-7114. DOI: 10.1103/physrevlett.129.150403. URL: http://dx.doi.org/10.1103/PhysRevLett.129.150403.