Intrinsic randomness in non-local theories:
quantification and amplification

Ph.D. Thesis

Ph.D. Candidate:
Chirag Dhara

Thesis Supervisor:
Dr. Antonio Acín

ICFO - Institut de Ciènces Fotòniques

Acknowledgements

The research presented in this thesis was conducted at the Institute of Photonic Sciences, Barcelona over four odd years under the supervision of Antonio Acín. I am indebted to him for all the stimulating discussions, advice and support over the years. His ability to combine work and life commitments yet always making the time to discuss our scientific and non-scientific problems has been an inspiration.

There is, of course, nothing like working beside friends! So it is my pleasure to thank all the members of my group, past and present, who have made these years memorable. Particular thanks go to Daniel, Rodrigo, Lars, Belén, Tobias, Anthony and Ariel, some of whom I've worked with and all of whom have been great friends.

My office mates, close friends and collaborators, Gonzalo and Giuseppe deserve a special mention. It has been a pleasure working with them and sharing thoughts on science, philosophy and the greatest invention since the wheel - pocket coffee!

I also owe a huge debt of gratitude to the members of the HR department at ICFO, Manuela, Anne, Cristina, Mery and Laia for their enormous assistance in these years and answering all my questions with almost otherworldly patience! They have undoubtedly been the single biggest factor in making my life (indeed that of almost everyone at ICFO) free of bureaucratic hurdles and frustrations and allowing me to focus on work.

I would also like to thank the three people who in the past have influenced my career above all others. The most important has been my mother, Lalitha. She has done what mothers do: love, support, sacrifice and encourage (and frustrate!) and I hardly need say more. The others are Ajay Patwardhan, my lecturer during BSc who introduced me to the world of research and N.D. Hari Dass, who advised me in a summer programme and has been a mentor and friend ever since.

Finally, and most importantly, I would like to reserve my greatest thanks for my partner, closest friend and confidant, Mona. Her love and unwavering

support through the most difficult stages of my work have been the most important resource for me in completing this thesis. It has been a long journey that I could not have made without her.

Abstract

Quantum mechanics was developed as a response to the inadequacy of classical physics in explaining certain physical phenomena. While it has proved immensely successful, it also presents several features that severely challenge our classicality based intuition. Randomness in quantum theory is one such and is the central theme of this dissertation.

Randomness is a notion we have an intuitive grasp on since it appears to abound in nature. It afflicts weather systems and financial markets and is explicitly used in sport and gambling. It is used in a wide range of scientific applications such as the simulation of genetic drift, population dynamics and molecular motion in fluids. Randomness (or the lack of it) is also central to philosophical concerns such as the existence of free will and anthropocentric notions of ethics and morality.

The conception of randomness has evolved dramatically along with physical theory. While all randomness in classical theory can be fully attributed to a lack of knowledge of the observer, quantum theory qualitatively departs by allowing the existence of objective or intrinsic randomness.

It is now known that intrinsic randomness is a generic feature of hypothetical theories larger than quantum theory called the non-signalling theories. They are usually studied with regards to a potential future completion of quantum mechanics or from the perspective of recognizing new physical principles describing nature. While several aspects have been studied to date, there has been little work in globally characterizing and quantifying randomness in quantum and non-signalling theories and the relationship between them. This dissertation is an attempt to fill this gap.

Beginning with the unavoidable assumption of a weak source of randomness in the universe, we characterize upper bounds on quantum and non-signalling randomness. We develop a simple symmetry argument that helps identify maximal randomness in quantum theory and demonstrate its use in several explicit examples. Furthermore, we show that maximal randomness is forbidden within general non-signalling theories and constitutes

a quantitative departure from quantum theory.

We next address (what was) an open question about randomness amplification. It is known that a single source of randomness cannot be amplified using classical resources alone. We show that using quantum resources on the other hand allows a full amplification of the weakest sources of randomness to maximal randomness even in the presence of supra-quantum adversaries. The significance of this result spans practical cryptographic scenarios as well as foundational concerns. It demonstrates that conditional on the smallest set of assumptions, the existence of the weakest randomness in the universe guarantees the existence of maximal randomness.

The next question we address is the quantification of intrinsic randomness in non-signalling correlations. While this is intractable in general, we identify cases where this can be quantified. We find that in these cases all observed randomness is intrinsic even relaxing the measurement independence assumption.

We finally turn to the study of the only known resource that allows generating certifiable intrinsic randomness in the laboratory *i.e.* entanglement. We address noisy quantum systems and calculate their entanglement dynamics under decoherence. We identify exact results for several realistic noise models and provide tight bounds in some other cases.

We conclude by putting our results into perspective, pointing out some drawbacks and future avenues of work in addressing these concerns.

Contents

Chapter 1

Introduction

Quantum theory started developing around the beginning of the twentieth century as a response to limitations of the prevailing classical theories. The dramatic failure in a physical explanation for the black body radiation spectrum (termed the ultraviolet catastrophe) was the proximate event that was solved by Planck's introduction of quanta of energy.

Quantum theory has since become the most successful theory in physics predicting observed behaviour with unprecedented accuracy in several domains of physics. It was successfully applied to describe scattering, matter-radiation interaction, nuclear decay and in condensed matter physics [ER85].

However, quantum theory also famously presents several counter intuitive and bizarre features such as the wave-particle duality, the uncertainty principle and non-locality. As a result, the study of the foundations of quantum theory has remained a subject of intense study right from its inception to date. The intrinsic randomness codified in the axiomatic structure of operational quantum theory is one such intriguing feature and is the subject of the present thesis.

1.1 Motivation for the study of randomness

Randomness is a notion that we understand and identify with at an intuitive level. It is usually associated with events with no intelligible pattern or predictability or whose underlying cause or structure is indiscernible. Randomness appears to abound in situations and events all around us. Coin flips are used for random initialization in sport, weather is a complex system with behaviour that appears unpredictable. Random throws of dice are used in gambling. In the physical sciences, randomness constitutes a valuable resource for applications such as cryptographic protocols [Gol01, Gol04, GRTZ02] or the numerical simulation of physical and biological systems [KTB11]. The mechanisms of evolution - like natural selection and genetic drift - work with the random variation generated by mutation [Sch44]. Randomness also occupies a central role in philosophical debates about the existence of free will[1] [Kan98] with the natural implications for anthropocentric concerns such as ethics and morality[2].

Important note on terminology. We use the terms randomness and unpredictability interchangeably unless specified otherwise. We distinguish between different flavours of randomness with the use of adjectives such as *classical* or *intrinsic* (discussed below).

Given the centrality of the notion of randomness in such varied disciplines of knowledge, the immediate question that begs itself is, *is the perceived randomness merely a reflection of the less-than-complete subjective state of knowledge of the observer or does genuine randomness indeed exist?* For example, coin flips appear random because of incomplete knowledge of parameters such as applied force, torque and interactions such as the friction due to air viscosity. However, given such knowledge, a coin flip is fully predictable. Additionally, identical initialization and interactions make the outcomes fully reproducible. Same is the case with the throw of dice. Weather systems also display what appears to be highly random behaviour. This is the outcome of the non-linear dynamics of such systems making them highly sensitive to initial conditions despite the dynamics being described by deterministic equations. Minor variations are amplified quickly resulting in their characteristic behaviour. In other words, the behaviour of such systems is deterministic and reproducible in principle if the initializing conditions can be made sufficiently accurate.

In fact, the mere possibility of the existence of objective randomness is forbidden within the confines of classical physics. Perfect knowledge of the positions and momenta of a system of classical particles at a given time, as well as of their interactions, allows one to predict their future (and also past) behaviour with complete certainty [Lap40]. Thus, any unpredictability observed in classical systems is but a manifestation of our imperfect description of the system. Henceforth, we term this as classical or deterministic randomness.

It was the advent of quantum physics that put into question this deter-

[1] Incompatibilism is a school of philosophy that considers determinism to be incompatible with free will. It is a view this author sympathizes with.

[2] http://www.bu.edu/law/central/jd/organizations/journals/bulr/documents/SCANLON.pdf

ministic viewpoint, as there exist experimental situations for which quantum theory gives predictions only in probabilistic terms, even if one has a perfect description of the preparation and interactions of the system. Nuclear decay, electronic transitions in atoms and vacuum fluctuations are examples of what is considered objectively random (quantum) behaviour.

In other words, quantum theory postulates the existence of (what we term henceforth) objective, intrinsic or genuine randomness qualitatively distinct from classical randomness. From a classical perspective,this is a highly counter intuitive phenomenon and a ”solution” was proposed in the early days of quantum physics: Quantum mechanics had to be incomplete [EPR35], and there should be a complete theory capable of providing deterministic predictions for all conceivable experiments. There would thus be no room for objective randomness, as any observed randomness would again be a consequence of our lack of control over hypothetical ”hidden variables” not treated by the quantum formalism.

This remained a burning philosophical question until path-breaking research by John Bell where he proved a no-go theorem [Bel64] implying that classical (local deterministic) hidden-variable theories are inconsistent with quantum mechanics. Therefore, none of these could ever render a deterministic completion to the quantum formalism. While determinism is an *ontological* assumption at the level of the hidden variables, it is known that the Bell inequalities can also be derived from the *operational* assumptions of signal locality (instantaneous communication impossible between separated observers) and predictability [CW12]. Thus, conditional on believing the validity of signal locality (also called no-signalling), Bell's theorem allows us to conclude that predictability must necessarily fail[3]. In other words, we are inexorably led to the conclusion that the known laws of physics indeed allow objective randomness to exist in nature.

The intrigue however deepens since there is a further subtle assumption used to derive the Bell theorem. Going under the name of *measurement independence or free choice*[4] it is the requirement that observers already possess a source of randomness before performing the Bell experiment. This leads to recursive logic in concluding the existence of randomness from the Bell theorem. In fact, by definition, super-deterministic models of nature postulating that all events in nature are fully pre-determined cannot be ruled out. In other words, at a philosophical level *the existence of objective*

[3] Here ”failure of predictability” is taken to mean in the strong sense that either any underlying model is indeterministic and if not, then the hidden variables are necessarily unknowable. Thus, failure of predictability -in this context- implies objective randomness.

[4] This is discussed at length in Chapter 2.

randomness, while allowed by the known laws of physics, must remain an un-testable assumption.

We proceed through the rest of this work under the implicit and critical assumption that a source of non-zero genuine randomness does indeed exist in our universe. The view we take here is that giving up this assumption entirely entails adverse consequences for both our current approach to science -such as the lack of free will [Gis10]- as well as our implicit understanding of the anthropocentric concepts previously alluded to. Thus, we believe ours to be a natural assumption.

1.2 Outline of thesis and major questions addressed

Here we trace the underlying theme and the work presented in the following chapters as well the links between them.

Having accepted the existence of sources of non-zero intrinsic randomness in the universe, the next logical question is if there are upper limits on how much randomness may be attained in theories of nature and how we may quantify such randomness. How does such a quantification depend on the specific mathematical framework being employed to describe the theory? It turns out that depending on whether we use the quantum framework or an expanded one called the no-signalling framework, there are different bounds on the maximum of objective randomness. We address this in detail in Chapters 3 and 4.

The next question we pose concerns the connection between the initial non-zero randomness to the maximum allowed randomness discussed above. More to the point, is it possible to *amplify* randomness fully? The answer to this question is a strict ”no” using exclusively classical resources [SV86]. We find that using quantum resources on the other hand makes this task possible in the broadest possible framework of no-signalling. This result and its far reaching ramifications are discussed in Chapter 5.

Staying with the theme of intrinsic randomness in the non-signalling framework, in Chapter 6 we identify broad criteria for certain scenarios that quantify the intrinsic randomness content in observed correlations. Our methods also constitute a significantly simpler proof of randomness amplification and produces a maximally random bit exponentially fast in the system size.

We conclude in Chapter 7 with a study of the only known physical resource known to generate intrinsic randomness, viz. entanglement. We study the dynamics of entanglement in an important class of multipartite entangled quantum states called the graph states evolving under noise (Chapter 7).

1.3 Contributions

This section summarizes the ideas and work that I, together with my co-workers, have developed over the course of my doctoral study.

1.3.1 Randomness in quantum theory certified by Bell inequalities.

We begin by studying randomness within quantum theory. We show that symmetries in Bell inequalities may be exploited to certify the existence (or lack thereof) of maximal randomness in quantum distributions maximally violating these inequalities. For this we require that such distributions be unique. We demonstrate uniqueness for several useful Bell inequalities and postulate it to be a general property of the quantum set of correlations - unlike the local or the non-signalling sets. We then use this for certificates of randomness in scenarios of higher complexity. We identify several scenarios where maximum randomness is attained within the quantum set [DPA12].

1.3.2 Maximal randomness precluded in maximally non-local theories.

It is well known that genuine randomness is completely precluded within classical theories or equivalently (as we call them here) the local set of correlations. Intrinsic randomness is a feature of general non-local, non-signalling theories of which quantum theory is a strict subset. We ask the question: what is the maximum allowed randomness in the largest possible set of non-local correlations respecting signal-locality? Intriguingly, we find that maximal randomness is forbidden for general no-signalling theories. We find upper bounds on the global randomness for the most general scenarios and find that in certain cases, the randomness diverges greatly from the maximum in the quantum set. This completes the theme of maximal (genuine) allowed randomness within the three sets of interest, local, quantum and no-signalling: No randomness in the local set, maximum in the quantum, but strictly less than maximum in the no-signalling [*In preparation*].

1.3.3 Full randomness amplification possible with quantum resources.

Randomness amplification is an informational task of using poor quality randomness and distilling higher quality randomness using available physical resources. It has been long known that this is impossible classically. However, it was recently shown to be possible in a limited case within quantum theory. We completed this project by showing that *full* amplification of randomness may be achieved by using quantum non-local resources. In other words, we showed that given a source generating the smallest amount of non-zero randomness we can use quantum correlations to amplify this randomness to that of a perfectly unbiased bit. We derive this result in a Device Independent manner which makes it conditional on only observing the desired violation of a Bell inequality. Thus, they are valid even in cryptographic settings allowing the existence of a supra-quantum adversaries. Full amplification has far reaching philosophical implications since it guarantees that fully random events are guaranteed to occur in our universe even if only slightly random events are assumed to exist [GMdlT$^+$12].

1.3.4 Observed randomness is fully genuine.

In the so-called device independent scenarios that are dealt with in this work, the source of the observed correlations is uncharacterised. Thus, the preparation of the observed correlations may in general be a mixture of extremal correlations, knowledge of which is hidden from the observers. Thus, one would generally expect the randomness observed in such correlations to contain some classical randomness associated with the lack of knowledge of the source. We show that, even so, one may choose certain correlations and scenarios where the observed randomness of appropriately defined functions completely excludes classical randomness. What makes this result even more significant is its validity even under almost complete relaxation of the freedom of choice assumption. Our criteria are general enough that for the first time, such results are derived for finite choices of parties, measurements and outcomes. The techniques can also be extended to provide a new and significantly simpler proof of full randomness amplification [*In preparation*].

1.3.5 Noisy entanglement evolution in graph states.

We finally change focus from characterization of randomness in non-local theories to the study of a physical resource strictly necessary for the existence of intrinsic randomness, namely, entanglement. We study entanglement dynamics in a very important class of quantum systems called the graph states. These include the GHZ states whose randomness properties we explore throughout this thesis. Graph states constitute resources for universal quantum computation and hence the study of their entanglement properties is of independent interest. We develop and expand on a computational tool which allows us to compute the entanglement dynamics of graph states under the action of an important class of noise channels, called the Pauli channels. We compute the decay of entanglement in systems of up to 14 qubits. The show that our method is scalable and generalize it to noise channels outside the afore said class with concrete examples [Dha09, ACC^{+}10].

Chapter 2

Preliminaries

In this chapter, we introduce the most important concepts and definitions that are necessary to make this thesis self-contained.

Section 2.1 introduces the concept of non-locality since this is a strictly necessary requirement for the existence of intrinsic randomness.

Section 2.2 is a brief introduction to the Device-Independent formalism which allows us to characterize randomness without reference to the internal working of the devices used and under the sole assumption of signal locality.

Non-locality imposes a natural classification in the space of correlations. We characterize the intrinsic randomness in each of these classes in Chapters 3 and 4. In preparation, Section 2.3 introduces the geometry of the correlation spaces.

Section 2.4 is a primer on the entropic measures of randomness we will use for our work and we end with a discussion about quantum entanglement in Section 2.5 since this is the topic of Chapter 7.

2.1 Non-locality

One of the clearest manifestations of the remarkable and highly counter-intuitive behaviour of quantum systems is the violation of the Bell inequalities [Bel64, Bel66, Bel87] by entangled states. This feature is termed non-locality. It captures the notion that statistics generated by measurements on entangled quantum systems do not allow simulation by strictly local resources. Such a simulation necessarily requires, in addition to local resources, some non-local resource such as communication [TB03].

As discussed in Chapter 1, the initial suggestions of the existence of local hidden variable models (LHVMs) simulating quantum statistics were an attempt to demonstrate the incompleteness of quantum theory [EPR35] which however failed in that particular case [Boh35]. However, the question remained open, at a philosophical level, whether the quantum description of nature was complete. It was after several decades that work by John Bell convincingly placed the question in the realm of experimentation and observation. His approach consisted of bounding the correlations that may be attained by LHVMs and demonstrating explicitly that there exist correlations in quantum theory that exceed these bounds. These bounds are usually called the Bell inequalities. It has been shown both theoretically as well as experimentally [ADR82] that these inequalities are violated by quantum probability distributions. The experimental confirmation is modulo certain well known loopholes that are difficult to treat technologically but are being actively addressed in different setups [TBZG98, WJS^{+}98, Row01, GMR^{+}12].

Non-locality is a necessary condition for the existence of intrinsic randomness. This is because the violation of Bell inequalities (namely, non-locality) implies the failure of the conjunction of signal locality and predictability as discussed in the previous chapter. Since signal locality has passed all experimental tests to date we may safely assume the failure of predictability, implying intrinsic randomness. This is an important observation that we will return to time and again.

A simple derivation of the Clauser-Horne-Shimony-Holt inequality

The easiest demonstration of Bell's idea which is directly amenable to experiment is called the Clauser-Horne-Shimony-Holt (CHSH) inequality after its authors [CHSH69]. It involves two parties, each choosing between two possible measurements. Each measurement may yield one of two possible outcomes (these are termed dichotomic measurements). We sketch a very simple derivation of this inequality and show that even so one can see how it is violated by quantum theory.

Consider two observers A and B (See Fig. 2.1) in different locations who perform several runs of the following experiment in order to gather statistics. Each observer is given a choice among two possible measurements. At every run of the experiment, the source distributes some physical state between them. Then, each observer must choose randomly among her options to make a measurement on the part of the system in her possession. We denote the measurement choice of A at a given round by x which could be either of x_0 or x_1 and the choice of B by y that may be either of y_0 or y_1. Each of these measurements is dichotomic i.e. having only two possible outcomes. We denote the outcomes of measurements x_i by a_i and those of y_i by b_i for $i = 0, 1$. While referring to the measurements with the variables x or y, the corresponding outcomes are referred to as a or b. With this notation, the statistics predicted by quantum theory are given by the Born rule, $P_{AB}(a, b|x, y, \rho) = \mathrm{tr}(\rho M_x^a M_y^b)$, where $M_x^a \geq 0$ and $M_y^b \geq 0$ constitute elements of a general quantum measurement that satisfy $\sum_a M_x^a = \mathbb{I}$ and $\sum_b M_y^b = \mathbb{I}$ for all the inputs.

Locality. The correlations $P_{AB}(a, b|x, y)$ are called local à la Bell or consistent with an LHVM iff,

$$P_{AB}(a, b|x, y) = \int_\lambda d\lambda \rho(\lambda) P_A(a|x, \lambda) P_B(b|y, \lambda), \tag{2.1}$$

where λ is distributed according to some well-defined density function $\rho(\lambda)$ and P_A and P_B are well defined local response functions. λ is understood to encode all the additional (unknown or unknowable) information required to assign locally probabilities to the outcomes of every possible measurement.

Determinism. In a deterministic model $P_A(a|x, \lambda) \in \{0, 1\}$ and $P_B(b|y, \lambda) \in \{0, 1\}$. While this is not required for the general stochastic model in Eqn. 2.1 it is known that determinism does not diminish generality [Fin82]. Using the standard notation of ± 1 to denote possible values of a_i and b_i, Eqn. 2.1 is then equivalent to

$$P_{AB}(a, b|x, y) = \int_\lambda d\lambda \rho(\lambda) \delta_A(a, f(x, \lambda)) \delta_B(b, g(y, \lambda)), \tag{2.2}$$

where $f(x, \lambda) \in \{-1, 1\}$ and $g(y, \lambda) \in \{-1, 1\}$ are deterministic functions with values specified fully by the corresponding input and the underlying hidden variable.

What the observation above tells us is that it is sufficient to consider deterministic local models to compute bounds on Bell inequalities. We use this observation in the following example.

Let us meditate on possible values of the expression

$$\mathcal{B} = a_0 b_0 + a_0 b_1 + a_1 b_0 - a_1 b_1, \tag{2.3}$$

for a deterministic local model at one run of the experiment where $\lambda \equiv \{a_0, a_1, b_0, b_1\} \in \{-1, +1\}^4$. It can be immediately verified that $\mathcal{B}_\lambda \in \{-2, 2\}$. Interpreting the above observation in the context of a full experiment we are only interested in the average of $\mathcal{B}$ over all the runs. Indeed, without the knowledge of the underlying hidden variables $\mathcal{B}$ cannot be computed at individual runs since A observes only one of a_0, a_1 and B only

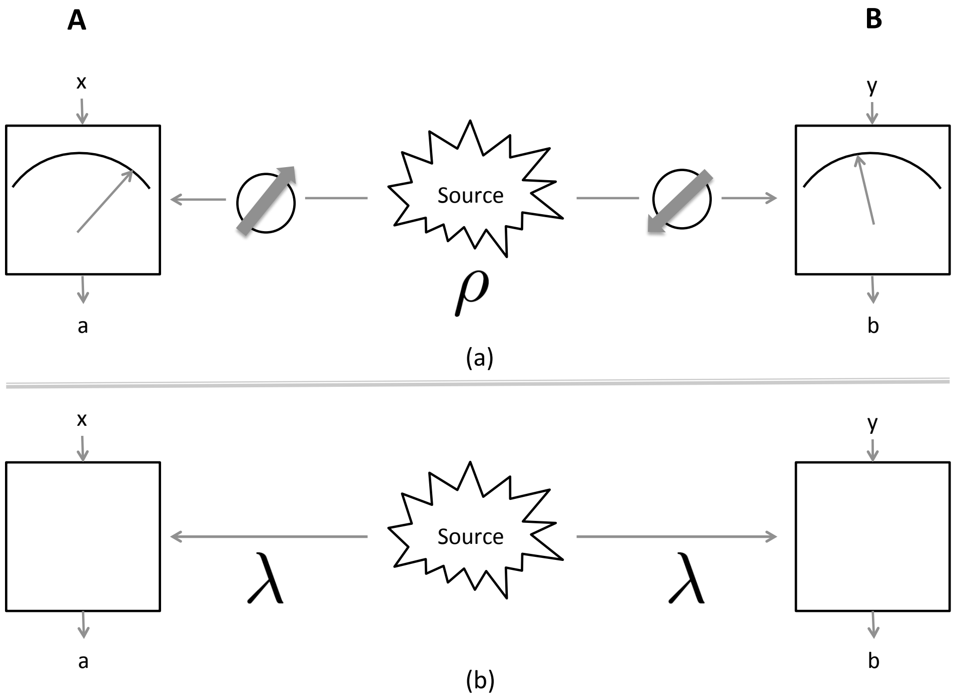

Figure 2.1: The simulation of correlations observed from measurements on quantum entangled particles with a hidden variable model. a) A source of quantum states of entangled particles (say, electrons) on which A and B measure observables (say, spin) x and y obtaining the corresponding outcomes a and b. b) The simulation of the correlations using only local hidden variables λ where the inputs are denoted by classical bits x and y yielding classical bits a and b as outcomes.

one of b_0, b_1. Then we ask: "Do the observed correlations allow description in terms of a deterministic LHVM?" To answer the question we begin by noticing that for any LHVM we have the bound,

$$\begin{aligned}\langle \mathcal{B} \rangle &= \langle a_0 b_0 \rangle + \langle a_0 b_1 \rangle + \langle a_1 b_0 \rangle - \langle a_1 b_1 \rangle \\ &\leq 2. \end{aligned} \tag{2.4}$$

This expression is called the CHSH inequality. Now the remarkable bit: It turns out that *there exist quantum correlations that violate this inequality.* For instance, for the maximally entangled state $|\psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ and the choice of measurements $x_0 = \sigma_x$, $x_1 = \sigma_z$ and $y_0 = (\sigma_x + \sigma_z)/\sqrt{2}$, $y_1 = (\sigma_x - \sigma_z)/\sqrt{2}$ the value of the CHSH operator is $\langle \mathcal{B}_\psi \rangle = 2\sqrt{2}$. This proves the existence of quantum correlations that cannot be simulated using local resources alone.

The consequences of this remarkably simple observation are profound indeed as discussed earlier in this text. However, there are certain critical assumptions that are used in the derivation of the Bell inequalities. Any violation of those assumptions could imply a failure in the conclusions[1]. We discuss these assumptions in the next section where the more general device-independent formalism is developed. Under this formalism, the assumptions become much more transparent and easier to place into context.

2.2 The Device Independent Formalism

The Device Independent (DI) formalism has been an outgrowth of the formalism of the Bell inequalities. As was observed in the previous section, the bipartite CHSH inequality studied the correlations in the statistics generated by certain entangled states denoted by $P(a,b|x,y)$. The CHSH is an example of $N = 2$ parties each performing $M = 2$ measurements of $d = 2$ outcomes. We denote Bell scenarios with the shorthand (N, M, d) in which case the CHSH represents the simplest possible one: the $(2, 2, 2)$. The correlations in a general (N, M, d) scenario are encapsulated in the object $P(a_1, a_2, \ldots, a_N|x_1, x_2, \ldots, x_N)$, where $x_i \in \{1, \ldots, M\}$ is the measurement choice of the ith party yielding outcome $a_i \in \{0, \ldots, d-1\}$ for $i = 1, \ldots, N$.

Underlying the DI approach is the simple observation that the object $P(a_1, \ldots, a_N|x_1, \ldots, x_N)$ requires no knowledge about the precise physical processes generating the experimental results. It is neutral with respect to the underlying states, the dimension of the respective Hilbert spaces and the description of the physical measurement devices. *In fact, the only quantity of importance in this approach is $P(a_1, \ldots, a_N|x_1, \ldots, x_N)$.* Thus, in this approach, physical devices are replaced with black boxes, the internal working of which are of no relevance since only the measured statistics are used for the analyses.

The DI study of correlations and the violation of Bell inequalities are immensely useful from both a foundational and applications point of view. Foundationally, this approach allows the characterization of physical quantities as independent of the mathematical framework as possible. In fact, the only framework is that of no-signalling. Non-signalling is the assumption that information propagation speed is finite and we discuss it in greater detail further in this section. The DI approach allows us to characterize the properties of no-signalling probability distributions independent of whether they

[1] The afore mentioned *loopholes* in Bell experiments are some of the technological/ practical inadequacies that may result in such failure.

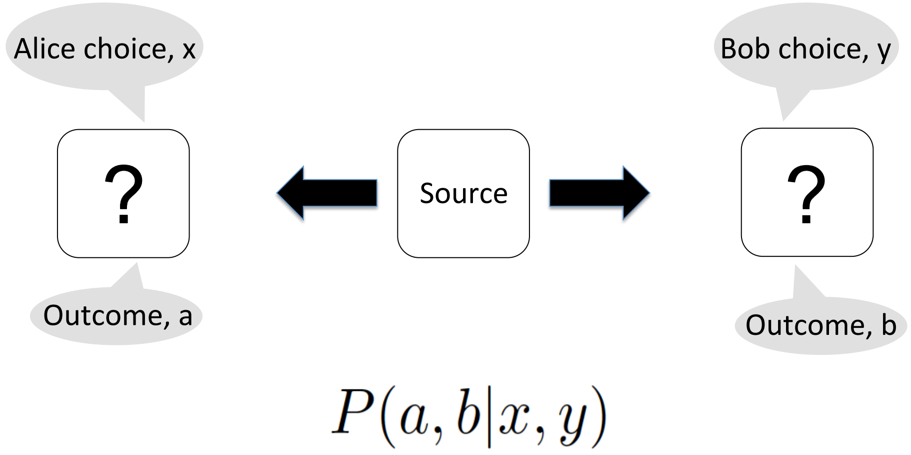

Figure 2.2: Schematic for a bipartite Device Independent Bell experiment. The statistics gathered at the end of the experiment is independent of the internal working of the source and the measurement devices and thus makes the conclusions dependent on only a few assumptions such as non-signalling.

were generated by a quantum system. It is highly desirable, from a foundations perspective, to study the properties of such correlations since it allows us to identify properties that are generic among all non-local theories vs those that are specific to quantum theory alone [MAG06]. Such approaches have even motivated the formulation of plausible physical principles such as Information Causality [PPK$^+$09], Macroscopic Locality [NW10] and Local Orthogonality [FSA$^+$12] that may serve to distinguish quantum theory from the larger set of no-signalling theories. Several important results connected with DI randomness expansion [Col07, PAM$^+$10b, VV12a] and (full) randomness amplification [CR12b, GMdlT$^+$12] have also been obtained. Much work has also focussed on a DI quantum information processing and cryptography [ABG$^+$07, Eke91, BHK05, DMPA11, VV12b].

We illustrate these ideas and the explain the underlying assumptions using the schematic in Fig. 2.2. While the illustrated system is bipartite, this is only for convenience of representation. All the statements and assumptions that follow are made for the most general case (N, M, d). While the object $P(a_1 \dots a_N|x_1 \dots x_N)$ is obtained in a DI fashion, all applications require calculating its Bell violation. However, even the mere requirement that this quantity is *definable* entails the implicit assumption that the source can generate identical states which are independent of one another at every run of the experiment. This assumption is called the *i.i.d.* assumption. From the foundational physical point of view, this is a reasonable assumption. However, in cryptographic scenarios we would like to avoid even this assumption [BCH$^+$02] which is also the case in Chapter 5.

For the moment, making the *i.i.d.* assumption, we can define probabilities as $P(a_1,\ldots,a_N|x_1,\ldots,x_N) = \lim_{N(x_1,\ldots,x_N)\to\infty} \frac{N(a_1,\ldots,a_N|x_1,\ldots,x_N)}{N(x_1,\ldots,x_N)}$. From now, we also use the shorthand notation $\mathbf{a} = \{a_1, a_2, \ldots, a_N\}$ and $\mathbf{x} = \{x_1, x_2, \ldots, x_N\}$ where necessary or convenient.

Assumptions. Here we discuss the most important assumptions made in deriving the Bell inequalities and by extension in the DI formalism.

1. *Non-signalling.* Non-signalling is the most important assumption made in the Device Independent formalism. It basically forbids the existence of super-luminal communication. Non-signalling is an *operational* assumption referring to observable and experimentally verifiable phenomena. In fact, this assumption is generic to a large part of modern physics and is one of its most well tested hypotheses. Thus, we may be confident of its validity.

   At the level of correlations, non-signalling is translated into the requirement that the outcome of a measurement performed at one location is independent of the choice of (simultaneous) measurements at other locations. It is a natural requirement in that it allows associating well-defined statistics to marginal distributions as expressed below:

$$\begin{aligned}&P(a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_N|x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_N)\\&=\sum_{a_i} P(a_1,\ldots,a_i,\ldots,a_N|x_1,\ldots,x_i,\ldots,x_N)\\&=\sum_{a_i} P(a_1,\ldots,a_i,\ldots,a_N|x_1,\ldots,x_i',\ldots x_N),\end{aligned}$$

   where x_i and x_i' are distinct measurement inputs of party i. This ensures that the marginal of all the parties but i is well defined.

2. *Measurement independence.* Informally, the assumption of measurement independence or free choice is the notion that the every party has access to a source of private randomness to choose her measurement at every run of the experiment. This source may be a pseudo-random generator or a source of quantum randomness like nuclear decay or even her own "free will". Hence, this is also referred to as the free will

assumption. Mathematically, it implies that the input choice x_i is independent of any underlying hidden variables λ: $p(x_i|\lambda) = p(x_i)$. This assumption may however be rejected as being too strong and recently much work has been directed towards weakening this assumption as far as possible [BG10, Hal10, Hal11, GMdlT^{+}12, KHS^{+}12].

2.3 The geometry of correlations

The Bell inequalities impose a natural structure in the space of correlations. In particular, we may distinguish between the local and the non-local correlations and the latter are further classified as quantum or non-signalling. We introduce these ideas with the bipartite $(2, 2, 2)$ scenario since it is the easiest in which to discuss the geometry of correlation spaces (See Fig. 2.3 for a 2d projection of an 8d space).

Given some statistical correlations in the form a non-signalling distribution $P(a, b|x, y)$, we test the CHSH expression $\langle\mathcal{B}\rangle$ given in Eqn. 2.4. Then the following is known:

- **Local**. If $P(a, b|x, y) = \int_\lambda d\lambda\rho(\lambda)P(a|x, \lambda)P(b|y, \lambda)$ then $P(a, b|x, y) \in \mathcal{L}$ where $\mathcal{L}$ denotes the local set. The distribution can be simulated by a LHVM with hidden parameters λ. In this case, $\langle\mathcal{B}\rangle_P \leq 2$.

  The local set is known to be a convex set having a finite set of vertices and facets making it a polytope. The vertices of this polytope represent the deterministic strategies [Fin82].

- **Quantum**. If one can express the given correlations as $P(a, b|x, y) = \mathrm{tr}(\rho M_x^a \otimes M_y^b)$ for a quantum state $\rho \geq 0$ of unit trace $\mathrm{tr}(\rho) = 1$ and a set of general measurement operators $\{M_x^a\}_A$ and $\{M_y^b\}_B$ satisfying $\sum_a M_x^a = \mathbb{I}$ and $\sum_b M_y^b = \mathbb{I}$ for all x and y then the distribution belongs to the quantum set $\mathcal{Q}$. In this case, $\langle\mathcal{B}\rangle_P \leq 2\sqrt{2}$.

  $\mathcal{Q}$ is known to be a convex set with infinite vertices [Tsi87], thus not a polytope.

- **No-signalling**. If the given correlations cannot be represented in either of the two forms above, then they belong to the non-signalling set $\mathcal{NS}$. In this case, $\langle\mathcal{B}\rangle_P \leq 4$,

  The non-signalling set is also convex and known to be a polytope. These vertices are generally called the extremal boxes [BP05] while in the special case of the $(2, 2, 2)$ are called the PR boxes [PR94].

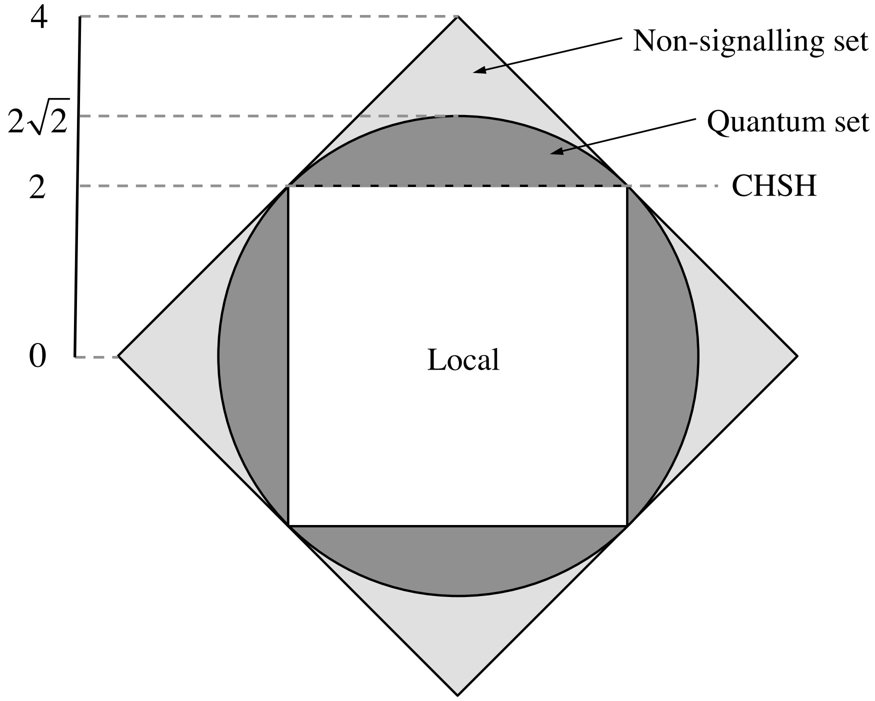

Figure 2.3: The geometry of the local, quantum and non-signalling sets for a bipartite $(2, 2, 2)$ scenario. The tight Bell inequalities are all symmetries of the CHSH inequality while the extremal boxes are all symmetries of the PR box. The numbers to the left indicate the value of the CHSH inequality. The relationship between the sets is $\mathcal{L} \subset \mathcal{Q} \subset \mathcal{NS}$.

We have the strict inclusion $\mathcal{L} \subset \mathcal{Q} \subset \mathcal{NS}$. The reason the $(2, 2, 2)$ has simple geometry is because all extremal boxes are known to be equivalent to each other under re-labellings of the inputs, outcomes and parties [BLM$^+$05]. Moreover, every Bell inequality is equivalent under symmetries to the CHSH inequality [Fin82].

However, in progressing to higher scenarios (N, M, d) the geometry quickly becomes far more complex. For instance, the $(3, 2, 2)$ already presents 46 distinct classes of extremal boxes and the same number of classes of Bell inequalities [PBS11, Ś03, Fri12]. The classification of local, quantum and non-signalling sets and their ordering relation follows the same logic as before and we represent the case $(3, 2, 2)$ with only the qualitative figure 2.4. The exact classification of inequalities and extremal boxes for higher scenarios is unknown in general and only a few partial results exist [BLM$^+$05, BP05].

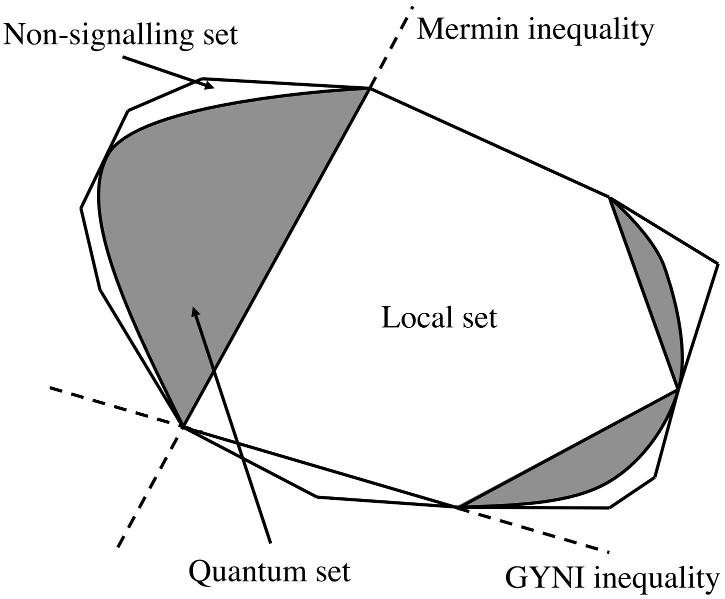

Figure 2.4: The qualitatively richer geometry of the local, quantum and non-signalling correlations for the $(3, 2, 2)$ scenario. Two non-equivalent Bell inequalities are indicated. The Mermin inequality is algebraically violated by the quantum set while the Guess Your Neighbour Input (GYNI) inequality is violated only by non-signalling points, but not quantum. In total, there are 46 non-equivalent classes of extremal boxes and as many classes of Bell inequalities. The ordering relation between the various classes follows as before the relationship $\mathcal{L} \subset \mathcal{Q} \subset \mathcal{NS}$.

2.4 Randomness

In this section, we summarize some of the most relevant entropic definitions of randomness and their properties. We also justify the use of *min-entropy* as our preferred definition in this work.

Entropy as a measure of unpredictability. Entropy in thermodynamics and statistical mechanics is often understood to be a measure of the disorder in the given system. This lends itself to the intuition that entropy constitutes a good measure of randomness. Three of the most common definitions of entropy used in information theory [CT91] are as follows:

Definition 1. ***(entropy measures)****. Let X be a discrete random variable with possible values $\{x_1, x_2, \ldots, x_n\}$ with respective weights $\{p_1, p_2, \ldots, p_n\}$. Then*

1. *The Rényi entropy [R6́1] of X is*

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \left(\sum_i p_i^\alpha \right) \tag{2.5}$$

2. *the Shannon entropy [Sha48] of X is,*

$$H_{Sh}(X) = -\sum_i p(x_i) \log p(x_i). \tag{2.6}$$

3. *the min-entropy of X is,*

$$H_\infty(X) = -\log \left(\max_i p(x_i) \right). \tag{2.7}$$

where all logs are to the base 2.

The Shannon and min-entropy are special cases of the Rényi entropy. For $\alpha \to 1$, it can be shown that the Rényi entropy converges to the Shannon entropy while for $\alpha \to \infty$, it converges to the min-entropy. The reason the later two quantities are explicitly defined above is because of their significance in information theory.

Henceforth, we will mainly be concerned by the latter two quantities. These quantities satisfy the following properties that we would expect from a randomness measure.

Lemma 2. ***(properties of entropy)**. Each of the entropy measures* $H \in \{H_{Sh}, H_\infty\}$ *satisfies for the random variables* X *and* Y*:*

- $H(X) \geq 0$*, with equality iff* X *is supported on a single element.*
- $H(X) \leq \log |Supp(X)|$*, with equality iff* X *is uniform on* $Supp(X)$[2].
- *if* X*,* Y *are independent, then* $H(X,Y) = H(X) + H(Y)$*,*
- *for every deterministic function* f*, we have* $H(f(X)) \leq H(X)$*, and*
- *for every* X*, we have* $H_\infty(X) \leq H_\alpha(X)$ *implying in particular* $H_\infty(X) \leq H_{Sh}(X)$*.*

[2] $Supp(X)$ denotes the support of X.

The Shannon entropy is related to the resources required to store information or alternatively, to the compressibility of information. This is called the Shannon's source coding theorem [CT91, NC00]. While the formal statement of the theorem is a departure of the main theme of this section, we do state its consequences and limitations as a measure of randomness. For example, suppose a random variable (equivalently information source) can take any one of four symbols 1,2,3 or 4. Naï vely, storage of this information requires 2 bits for each use of the source without compression. And indeed this is the case, if each symbol is equi-probable, that is, occurs with a probability of 1/4. However, for any non-uniform distribution of the symbols, the noiseless source coding theorem states that the information can be compressed on average to less than 2 bits. For example, for a source producing the symbols with the probabilities 1/8, 1/16, 1/16 and 3/4, the average storage required is only $H_{Sh} = 1.19$ bits. Thus, the source coding theorem assures us that the average storage space required is much less than the naïve value of 2 bits.

However, now consider the possible value of the random variable over just one run of the experiment and the issue of *randomness* in X rather than its *compressibility*. Intuitively, since X takes the value 4 with a relatively large probability and all the others with much smaller probability, we expect the randomness to be less than one bit. By this logic, the Shannon entropy $H_{Sh} = 1.19$ bits is clearly too large but the min-entropy $H_\infty = -\log 3/4 = 0.42$ bits gives a more intuitively satisfying measure of the randomness. This observation, while very hand-waving and qualitative, is useful as an aid to intuition in distinguishing the relation between compressibility and Shannon entropy and between randomness and min-entropy.

Min-entropy is related to the *guessing probability* of the outcome of a given random variable. This can be seen from Eqn. 2.7. It is the favoured measure of randomness used in the theory of randomness extractors [Rao07]. The operational interpretation as well as several useful properties of min-entropy were proved in [Ren05, KRS09].

Now we have established min-entropy as our measure of randomness of choice, we now distinguish between *classical or deterministic randomness* associated with our lack of knowledge and *genuine, intrinsic or objective randomness* as found exclusively within non-local theories. In particular, our task is to quantify the intrinsic randomness inherent in a given probability distribution depending on whether we view that distribution as belonging to quantum set alone or to the larger no-signalling set. We expand on these points in the following.

2.4.1 Randomness within the quantum formalism

Within the quantum theory, complete knowledge of the preparation of a system allows us to describe it with a pure state ψ. In such systems, all observed randomness is intrinsic since there is no randomness stemming from a lack of knowledge. Systems lacking an unambiguous pure state description are instead represented as a mixture of pure states, $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$. These are called mixed states and generally include both intrinsic randomness as well as classical randomness associated with out lack of knowledge of the exact preparation of the system. We are interested in characterizing the first type of randomness.

Definition 3. ***(Randomness in pure states)****. Let $|\psi\rangle \in H_A \otimes H_B$ be a bipartite pure state. Then the randomness of an outcome pair (a, b) resulting from the measurement of the observables $\hat{A}$ and $\hat{B}$ may be characterized by the guessing probability,*

$$G(\hat{A}, \hat{B}, \psi) = \max_{a,b} P(a, b|\hat{A}, \hat{B}, \psi) \tag{2.8}$$

with the min-entropy defined as $H_\infty(\hat{A}, \hat{B}, \psi) = -\log_2 G(\hat{A}, \hat{B}, \psi)$ from Eqn. (2.7).

As an example of the application of this formula[3], consider a state $|\psi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, $\hat{A} = \sigma_z$, $\hat{B} = \sigma_x$. Then, $P(a, b|\hat{A}, \hat{B}, \psi) = 1/4$ for all $a, b = 0, 1$ giving $G(\hat{A}, \hat{B}, \psi) = 1/4$. Thus for this uniform distribution we have $H_\infty = 2$ bits. This is the maximum possible randomness given this scenario since there are two observer each making a measurement of two outcomes (1 bit each).

By extension, the maximum randomness in a (N, M, d) scenario would be N dits of randomness or equivalently $N \log_2 d$ bits of randomness.

Definition 4. ***(Randomness in mixed states)****. Let $\rho \in \mathcal{O}(H_A \otimes H_B)$ be a bipartite mixed state. The intrinsic randomness in ρ associated with measurements $\hat{A}$ and $\hat{B}$ is given by the optimised guessing probability,*

$$G(\hat{A}, \hat{B}, \rho) = \max_{\{p_i, \psi_i\}} \sum_i p_i G(\hat{A}, \hat{B}, \psi_i), \tag{2.9}$$

where $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$. As before, $H_\infty(\hat{A}, \hat{B}, \rho) = -\log_2 G(\hat{A}, \hat{B}, \rho)$.

The optimization is performed in order to remove the classical randomness associated with the lack of knowledge of the exact preparation of a

[3] example chosen from [AMP12]

mixed state. In cryptographic adversarial terms, it quantifies the minimum randomness perceived by a quantum adversary correlated with ρ with knowledge of its preparation.

Randomness in a general quantum probability distribution. We are finally prepared to define randomness in a DI manner. We do this in the following for a general distribution $P_Q(a,b|x,y)$ known to come from a quantum system but where the precise states of measurements or the internal workings of the devices are unknown.

Definition 5. ***(DI randomness in quantum probability distributions).*** *The intrinsic randomness content of the distribution $P_Q(a,b|x,y)$ is,*

$$G(x,y,P_Q) = \max_{\{\rho,M\}\to P_Q} G(\hat{A}_x,\hat{B}_y,\rho) \tag{2.10}$$

where the optimization is performed over all quantum realizations $\{\rho,M\}$ compatible with $P_Q(a,b|x,y)$.

2.4.2 Randomness in non-signalling distributions

Let $P(a,b|x,y)$ be a non-signalling distribution. We can define two notions of randomness with regards to this distribution: the observed and the intrinsic.

Definition 6. ***(Observed randomness).*** *The observed randomness G_{obs} for a non-signalling $P(a,b|x,y)$ is given by,*

$$G_{\text{obs}}(x,y,P) = \max_{a,b} P(a,b|x,y). \tag{2.11}$$

Randomness quantified by this definition takes no account of the preparation of the the given distribution, which in general could be mixture of non-signalling extremal distributions (including local and quantum ones). Thus, it is expected to include a contribution from classical randomness associated with this lack knowledge in addition to intrinsic randomness.

In general, it is only if $P \equiv P^{\text{ex}}$ *i.e.* an extremal point of the no-signalling set that this definition is equivalent to the *intrinsic* randomness of P, since in this case there is no "missing" knowledge. If P is non-extremal, however, we define the intrinsic randomness of P as below.

Definition 7. ***(Intrinsic randomness).*** *The intrinsic randomness G_{int}*

for a non-signalling $P(a,b|x,y)$ is given by,

$$\begin{aligned} & G_{\text{int}}(x,y,P) = \max_{\{p_j, P_j^{\text{ex}}\}} \sum_j p_j G_{\text{obs}}(x,y,P_j^{\text{ex}}) \\ & \textit{subject to:} \\ & P(a,b|x,y) = \sum_j p_j P_j^{\text{ex}}(a,b|x,y). \end{aligned} \tag{2.12}$$

Note the analogy with the definition of randomness of mixed quantum states.

2.4.3 A comment on algorithmic definitions of randomness

Before concluding the discussion on randomness it is appropriate to comment on the quantification of randomness coming from the field of algorithmic information theory [Cha87]. A part of this subject deals with the randomness in bit strings and has developed some formidable notions of randomness. For infinite strings, the *Martin Löf* randomness [ML66] is a robust definition which satisfies all the intuitive properties we may expect from a random string, such as incompressibility and the lack of a shorter description of the string. While it is unknown exactly how quantum objective randomness relates to Martin Löf randomness, we believe that it is reasonable to assume that the two quantities are equivalent [Cal04].

The situation is more tricky for finite strings since there is no unifying notion of randomness in this case. Shannon entropy (already alluded to) of a randomness source and the Kolmogorov complexity [LV08] of a bit string are the two most important concepts. Using notions from Kolmogorov complexity, a finite bit string is defined to be random if it lacks a shorter description than itself in some universal description language. While within computation theory, this is a reasonable definition, we take the view that nothing can be concluded about a finite bit string without reference to the physical system generating the randomness: is it classical or quantum? On appending the next bits from the source, do we get a correspondingly larger incompressible string? For these reasons, we interpret Kolmogorov complexity (of finite bit strings) also as relating to the notion of the *compressibility* alone (as with Shannon entropy), but not to randomness in the sense of predictability.

2.5 Quantum Entanglement

We have noted before that non-locality is necessary for intrinsic randomness. However, non-locality in physical systems occurs only for those that are entangled. In other words, entanglement is the only physical resource known that allows non-locality and thus intrinsic randomness to exist. Thus, we turn to the study of entanglement as a resource. We study the dynamics of entanglement in a very important class of quantum systems termed graph states [BR01, RB01, DAB03, RBB03, HEB04, ACC$^+$10]. Since graph states are networks of Ising type interaction and code words for universal measurement based quantum computation, the study of entanglement in these systems is also significant independent of the theme of randomness.

Entanglement refers to the existence of global states of composite systems that cannot be written as a product of the states of the individual subsystems. Another way of stating the above is that complete knowledge of the global state of a composite system does not imply a complete knowledge of the subsystems which it consists of. This has no counterpart in classical theory.

Among the first papers to recognize entanglement was the EPR paper alluded to before [EPR35] as well as Schrödinger [Sch35]. While the authors of the former (EPR) regarded the existence of entanglement as a paradox indicating the inadequacy of quantum mechanics, the latter (Schrödinger) believed it to be an essential component of quantum mechanics. Despite the lack of consensus during the early days, entanglement has now been firmly established as an essential part of the formalism of quantum theory. The modern consensus considers entanglement to be a key resource in several informational tasks such as quantum dense coding, teleportation and swapping as well as in quantum cryptography and the speed up of quantum algorithms. Please see [NC00, HHHH09] and the references within for an exhaustive discussion.

We now define entanglement formally as well as some measures of entanglement which will be used in Chapter 7. Note that, entanglement usually has a negative definition: We define states that are *separable* and entangled states are understood to be precisely those that are *non-separable.*

2.5.1 Bipartite entanglement

Pure states. A pure state $|\psi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ with subsystems of local dimension d_A and d_B is separable iff it can be written as a product of vectors corresponding to the respective subsystems, i.e.

$$|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle. \tag{2.13}$$

In general, any bipartite pure state can be expressed using the Schmidt decomposition [NC00] as,

$$|\psi_{AB}\rangle = \sum_{i=1}^{r(\psi)} q_i |i_A\rangle \otimes |i_B\rangle, \tag{2.14}$$

where q_i are non-negative real numbers called the Schmidt coefficients satisfying $\sum_i q_i^2 = 1$ while $\{|i_A\rangle\}$ and $\{|i_B\rangle\}$ are orthonormal bases of $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. In general, $r(\psi) \leq \min[d_A, d_B]$ is called the Schmidt rank of ψ and is equal to either of the ranks of the reduced operators $\rho_A = \mathrm{tr}_B[|\psi_{AB}\rangle\langle\psi_{AB}|]$ or $\rho_B = \mathrm{tr}_A[|\psi_{AB}\rangle\langle\psi_{AB}|]$.

Given Eqn. (2.14), a quantitative restatement of the condition for entanglement: $|\psi_{AB}\rangle$ is separable iff $q_i = \delta_i^1$. Thus, any state requiring more than one Schmidt coefficient in Eqn. (2.14) is entangled.

And example of a pure entangled state is the singlet state, $|\phi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$.

Mixed states. In general, we deal with mixed states in the laboratory rather than pure states because of imperfections in the preparation procedures and decoherence. Hence we move next to the definition of entanglement in bipartite mixed states.

A bipartite mixed state ρ_{AB} defined on $\mathcal{H}_A \otimes \mathcal{H}_B$ is separable [Wer89] iff it cannot be represented by states of the form,

$$\rho_{AB} = \sum_{i=1}^{k} p_i \; \rho_A^i \otimes \rho_B^i, \tag{2.15}$$

where ρ_A^i and ρ_B^i are defined on $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively. These local density operators can be chosen to be pure for $\dim(\mathcal{H}_{AB}) < \infty$. Then, $k \leq \dim(\mathcal{H}_{AB})^2$ [HHHH09].

Separability criteria are generally hard to check for mixed states. For example, for the Werner states [Wer89] defined as $\rho_{AB}^W = v|\phi^-\rangle\langle\phi^-| + \frac{1-v}{4}\mathbb{I}$, we may apply the PPT criterion [Per96] to find that ρ_{AB}^W is separable for visibility $v \leq 1/3$. Thus, it is entangled for all $1/3 < v \leq 1$.

2.5.2 Multipartite entanglement

Multipartite entanglement is qualitatively far richer than bipartite entanglement. For a n-partite system of $n > 2$ one may distinguish genuine

n-partite entanglement from more restricted flavours of entanglement. Before proceeding to the formal definitions, we illustrate these differences with some examples.

It turns out that tripartite states are sufficient to demonstrate the complexity of the notions of entanglement in multipartite systems. Let us consider first the state,

$$\rho = p\ |+\rangle\langle+|^{\otimes 3} + (1-p)\ |-\rangle\langle-|^{\otimes 3}, \tag{2.16}$$

where $\{|+\rangle, |-\rangle\}$ is the eigenbasis of σ_x. This is clearly a separable state from a simple generalization of the notions of separability developed for the bipartite case. An example of a tripartite entangled state is the so-called GHZ state [GHZ89],

$$|\psi_{GHZ}\rangle = (|000\rangle + |111\rangle)/2. \tag{2.17}$$

These are the intuitive generalizations of the notion of separable and entangled states developed for the bipartite scenario. However, in a departure from the latter, there exist separable and entangled tripartite states which are not captured by the bipartite definitions. Let us meditate over the state,

$$|\psi_{ABC}\rangle = \frac{1}{2}\left(|00\rangle_{AB} + |11\rangle_{AB}\right) \otimes |0\rangle_C. \tag{2.18}$$

This state is clearly not ”fully” tripartite entangled since C is manifestly uncorrelated from the other two subsystems.

A more complicated example of a family of tripartite mixed states is of the form,

$$\rho_{ABC} = \sum_i p_i \rho_A^i \rho_{BC}^i + \sum_i r_i \rho_B^i \rho_{AC}^i + \sum_i q_i \rho_C^i \rho_{AB}^i. \tag{2.19}$$

Here, ρ_{ABC} is a mixture of states that are each entangled in two parties while uncorrelated from the third. We term this as a 2-entangled state.

These examples demonstrate the different flavours of entanglement that exist in multipartite systems. In this context, the examples of Eqns. 2.16 and 2.17 may be termed as fully separable and fully entangled states respectively while the examples of Eqns. 2.18 and 2.19 may be called 2-entangled pure and mixed states respective. We can use these examples to formulate our criteria for multipartite full and partial separability (and thus entanglement).

Definition 8. *(**Full multipartite separability**). An n-partite pure state is called fully separable iff $|\psi_{A_1 A_2 \ldots A_n}\rangle = |\psi_{A_1}\rangle \otimes |\psi_{A_2}\rangle \otimes \cdots \otimes |\psi_{A_n}\rangle$ while an n-partite mixed state is called fully separable iff $\rho_{A_1 A_2 \ldots A_n} = \sum_{i=1}^{k} p_i \rho^i_{A_1} \otimes \rho^i_{A_2} \otimes \cdots \otimes \rho^i_{A_n}$ [Wer89]. Any state not fully separable is called entangled.*

This definition does not guarantee that a given state is genuinely n-partite entangled. We formulate next the criterion to decide if a pure state is indeed genuinely n-partite entangled.

Definition 9. *(**Genuine multipartite entanglement**). An n-partite pure state is genuinely n-partite entangled iff every bipartition yields mixed reduced density operators.*

The intuitive reasoning is that this condition ensures that the state is not a product across any bipartite cut. Graph states are an important class of genuinely n-partite entangled multipartite states, entanglement dynamics of which are studied in detail in Chapter 7.

2.5.3 Quantifiers of entanglement

Since entanglement is a critical quantum resource it is important to quantify it for applications in the fields of quantum communication [BBP$^+$96, BDSW96] and algorithms (see Chapter 7). We do not give an exhaustive account of the various measures of entanglement but focus on the intuitive properties required of a ”good” measure of entanglement and the particular definition that will be used later in the text.

In fact, there is only one property that is considered critical for an entanglement quantifier which is monotonicity under Local Operations and Classical Communication (LOCC). This requires that the entanglement must be non-increasing under any LOCC operation $\mathcal{L}$,

$$E(\mathcal{L}(\rho)) \leq E(\rho) \tag{2.20}$$

Other properties such as asymptotic continuity and convexity are often useful and are satisfied by many measures of entanglement. However, they are not necessary. While there are several measures of entanglement relevant to different situations, we focus on *negativity* since it will be used later. We consider a (multipartite) state ρ in which we choose a certain bipartition.

Definition 10. *(**Negativity**). Negativity [ZHSL98, VW02] is defined as the absolute value of the sum of the negative eigenvalues of the given density* *matrix partially transposed with respect to the considered bipartition.*

$$\begin{aligned}\mathcal{N} &= \frac{\|\rho^T\|_1 - 1}{2} \\ &= |\sum_{\lambda<0} \lambda|, \qquad (2.21)\end{aligned}$$

where λ *denote the eigenvalues of* ρ^T*.*

Negativity is a convex entanglement monotone. Further properties are discussed where relevant (Chapter 7).

Chapter 3

Maximal quantum randomness

As we have discussed as some length in the preceding chapters, quantum theory incorporates intrinsic randomness in its framework having no classical counterpart. Non-locality is a necessary condition for the existence of intrinsic randomness. However, the relationship between the two physical quantities beyond this is troubled and we are still far from understanding the exact relation between them. For instance, probability distributions with maximal non-locality does not necessarily contain maximal randomness [AMP12]. Conversely, distributions with arbitrarily small non-locality may contain almost maximal randomness [AMP12]. This informs us that naïvely expecting maximally non-local quantum distributions to demonstrate maximal randomness is incorrect. One of the primary reasons why the relationship is hard to characterize is the lack of a general characterization of the boundary of quantum correlations. The best techniques known so far use a hierarchy of semi-definite programs that bound the quantum set asymptotically [NPA07]. However, the complexity and computational resources required at higher levels make the problem infeasible. Hence, it is not yet known if the problem of identifying the quantum boundary is even decidable [WCPG11]. Along these lines, identifying those quantum set-ups, namely Bell tests, which offer the highest possible randomness would be a highly desirable result, relevant to our theme of maximal randomness as well as for applications such as cryptography and others requiring high randomness sources.

Consider as an example the standard Clauser-Horne-Shimony-Holt (CHSH) inequality [CHSH69], $I_{CHSH} = \langle A_1B_1\rangle + \langle A_1B_2\rangle + \langle A_2B_1\rangle - \langle A_2B_2\rangle$. For the Tsirelson correlations maximally violating the CHSH, any measurement output by any of the parties provides a perfect random bit. That is, the corresponding probability distribution contains *locally* the maximum possible of one bit of randomness for every party and every measurement setting. However, there are strictly less than 2 random bits ($H_{\infty}^{AB} = 1.23$ bits) *globally*, as any pair of local measurements gives correlated results. Now, consider the following modification of the CHSH inequality, $I_{\eta} = \langle A_1 B_1 \rangle + \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle - \langle A_2 B_2 \rangle + \eta \langle A_1 \rangle$. At the point of maximal quantum violation, only the measurement A_2 defines a perfect random bit [AMP12]. Why this setting and not the others? Why all of them in the case of CHSH? More in general, does maximal (global) randomness occur for quantum correlations at all and if so what measurement settings much be chosen? What relationship do they bear with maximal non-locality?

3.1 Results

Our main result is to recover all previously known results on maximal randomness in quantum distributions and identify several new cases where maximal randomness exists. Furthermore, not all the possible measurement settings in a Bell type experiment can be used to certify maximal randomness so we provide a simple criterion to infer when and which settings in a Bell test are optimal for randomness extraction.

Given a Bell inequality, our method (i) assumes that the quantum probability distribution attaining its maximal violation is unique and (ii) exploits the symmetries of the inequality. We show how this method reproduces all known results relating Bell tests and maximal randomness. Moreover, based on our construction, we provide Bell tests certifying the maximal global randomness in a robust manner, that is, Bell tests for which there exist measurements by the N parties providing N random bits. We also provide a geometric interpretation of our findings. Finally, we discuss the existence of uniqueness and show that it is known to exist in several important cases either analytically or from numerical computation.

Bell's Inequalities	Randomness		Uniqueness
	Local	Global	
CHSH $(2, 2, 2)$	1-bit	$<$ 2-bits	anl
CGLMP $(2, M, d)$	1-dit	$<$ 2-dits	num
Chain $(2, M, 2)$	1-bit	2-bits (M_{odd})	num
Mermin $(N_{odd}, 2, 2)$	1-bit	N-bits	anl
Mermin $(N_{even}, 2, 2)$	1-bit	$(N-1)$-bits	anl
gMermin $(N_{even}, 2, 2)$	1-bit	N-bits	anl
gMermin $(3, 2, d)$	1-dit	3-dits	–

Table 3.1: All results derived in this chapter for local and global randomness and whether uniqueness is known analytically (anl) or numerically (num) in each case. The last two Bell inequalities refer to the ones in Eqns. 3.6 and 3.7 respectively.

3.2 Background

Before proceeding, we recall some important points. Since we often work with $(N, M, d = 2)$ cases we can use the following useful parametrization,

$$\begin{aligned} P(\mathbf{a}|\mathbf{x}) = \frac{1}{2^N}(1 &+ \sum_{i=1}^{N} a_i \langle A_i \rangle + \sum_{i<j} a_i a_j \langle A_i A_j \rangle \\ &+ \sum_{i<j<k} a_i a_j a_k \langle A_i A_j A_k \rangle + \cdots + a_1 a_2 \ldots a_N \langle A_1 A_2 \ldots A_N \rangle). \end{aligned} \tag{3.1}$$

Here, measurement outputs are labelled by ± 1 and $\langle A_i \ldots A_j \rangle$ are the standard correlators $\langle A_i \ldots A_j \rangle = Pr(a_1 \ldots a_N = +1 | x_1 = i, \ldots, x_N = j) - Pr(a_1 \ldots a_N = -1 | x_1 = i, \ldots, x_N = j)$.

Also note that in a general (N, M, d) scenario maximum local randomness is $\log_2 d$ bits while maximum global, $N \log_2 d$.

3.3 Methods

The main result of this chapter is a simple method to infer when some settings in a Bell test provide maximal randomness. We assume in what follows that the quantum distribution attaining the maximal quantum violation of the Bell inequality is unique[1] (discussed later). Under this assumption, we

[1] In the usual sense of the word: There is no other quantum distribution attaining the same Bell violation.

show how symmetries in the Bell inequality under permutation of measurement results, possibly together with permutations of measurement settings, lead to maximal randomness. Our method, then, can be summarized as follows: *uniqueness plus symmetries implies maximal randomness.*

3.3.1 Maximal local randomness

To illustrate our method, we return to the examples given above. Consider again the CHSH inequality and denote by $\mathcal{P}^*$ the distribution attaining its maximal quantum violation, namely $I_{CHSH}(\mathcal{P}^*) = 2\sqrt{2}$. Note that in this case, this distribution is known to be unique [Tsi87]. The symmetry transformation $\mathcal{T}_s$: $a_{1,2} \mapsto -a_{1,2}$ and $b_{1,2} \mapsto -b_{1,2}$ flips the signs of all the one-body correlators, $\langle A_i \rangle$ and $\langle B_j \rangle$, while keeps unchanged all two-body correlators, $\langle A_i B_j \rangle$. Applying $\mathcal{T}_s$ to P^* we obtain a new distribution $\mathcal{T}_s(\mathcal{P}^*) = \mathcal{P}^{**}$ with

$$\langle A_i \rangle^{**} = -\langle A_i \rangle^*, \qquad \langle B_j \rangle^{**} = -\langle B_j \rangle^*, \tag{3.2}$$

and that also maximally violates CHSH. Because of the uniqueness of the distribution, $\mathcal{P}^* = \mathcal{P}^{**}$ and all one-body correlators (3.2) must be zero, which certifies 1 bit of *local* randomness (for both parties). Moving to I_η, the transformation $a_2 \mapsto -a_2$, $B_1 \leftrightarrow B_2$, flips the value of $\langle A_2 \rangle$ without changing the value of I_η. Under the assumption of uniqueness, this proves that the setting A_2 is fully random. A little thought shows that it is impossible to construct similar transformations for the other local measurements. Our argument, then, easily reproduces the known results for these two inequalities.

As mentioned, our method applies to any Bell inequality with symmetries. The previous argument for the CHSH inequality can be easily generalized to all the chained inequalities of Refs. [BC90, BKP06]. Under the assumption of uniqueness, these inequalities always certify 1-dit of local randomness. The chained Bell inequalities can be compactly represented as [BKP06]:

$$C_d^M = \sum_{i=1}^{M} \langle [A_i - B_i]_d \rangle + \langle [B_i - A_{i+1}]_d \rangle \geqslant d - 1 \tag{3.3}$$

where A_i, $B_j \in \{0, \ldots, d-1\}$ are measurement choices for Alice and Bob and $A_{M+1} = A_1 + 1$. The square brackets denote sum modulo d.

Let P attain the quantum maximum of C_d^M. The transformation $\mathcal{T}$: $a_i \mapsto a_i + 1$ and $b_i \mapsto b_i + 1$ for every i changes the value of the marginal distributions of Alice and Bob but leaves the terms in C_d^M unchanged. Applying

$\mathcal{T}$ to P and assuming it to be unique, it follows that all local distributions of Alice and Bob must be uniform. In other words, the chained inequality certifies $\log_2 d$ bits of local randomness for every measurement by each party.

3.3.2 Maximal global randomness

A natural open question is whether there exist Bell tests in the (N, M, d) scenario that allow certifying the maximal possible randomness, namely $N \log_2 d$ bits. Some progress on this question was obtained in [AMP12], where it was shown how to get arbitrarily close to two random bits in the $(2, 2, 2)$ scenario. However the corresponding correlations are non-robust. Here, we show how our method can be easily applied to design Bell tests allowing maximal randomness certification in a robust manner.

We start with the bipartite case. Maximal global randomness is impossible in the CHSH case, as at the point of maximal violation all settings are correlated. Maximal global randomness, however, can be certified as soon as another measurement is included. More in general, consider the chained inequalities for an odd number of two-outcome measurements. We move to the notation $a_i, b_j = \pm 1$ and reexpress (3.3) as follows:

$$C_2^M = \left| \sum_{i=1}^{M} \langle A_i B_i \rangle + \sum_{i=1}^{M-1} \langle A_{i+1} B_i \rangle - \langle A_1 B_M \rangle \right| \tag{3.4}$$

where A_i, $B_j = \pm 1$. Let $M = 2k+1$. As above, we consider a transformation leaving C_2^M unchanged but under which $\langle A_1 B_{k+1} \rangle \mapsto -\langle A_1 B_{k+1} \rangle$. Such a transformation is: $\mathcal{T}$: $a_1 \mapsto -a_1$, $B_{1+i} \leftrightarrow B_{M-i}$, $A_{2+i} \leftrightarrow A_{M-i}$ $\forall i$ $0 \leq i \leq k-1$. Assuming that the distribution maximally violating (3.4) is unique leads to $\langle A_1 B_{k+1} \rangle = 0$. The previous results show that $\langle A_1 \rangle = 0 = \langle B_{k+1} \rangle$. These together certify 2 bits of global randomness for (A_1, B_{k+1}). Similar arguments certify maximal randomness in all inputs of the form (A_l, B_{k+l}) $\forall$ $1 \leq l \leq k$. Analogous to the case for CHSH, maximal randomness cannot be certified for those measurement combinations appearing in the chained inequality, as they display non-zero correlations. The previous results rely on the assumption of uniqueness, which is unknown for the case of the chained inequality with $M > 2$. We then follow [PAM$^+$10b] and apply the techniques in [NPA07, NPA08] to get an upper bound on the randomness of (A_1, B_2) for the chained inequality with 3 measurement settings. The obtained results corroborate the presence of maximal global randomness, up to numerical accuracy.

3.3.3 Maximal randomness from multipartite Bell inequalities

We show that randomness may be deduced in multipartite systems using appropriate Bell inequalities. We treat the cases of even and odd N separately since they require different Bell inequalities for the demonstration of full global randomness.

Case when N is odd

We consider the Mermin inequalities [Mer90] and prove that they allow certifying up to N bits of global randomness for arbitrary odd N. The Mermin inequalities belong to a class of Bell-Klyshko inequalities of N parties that are defined recursively as,

$$M_N = \frac{1}{2}M_{N-1}(A_N + A'_N) + \frac{1}{2}M'_{N-1}(A_N - A'_N) \tag{3.5}$$

where M_2 is the CHSH inequality and M'_{N-1} is obtained from M_{N-1} by exchanging all A_j and A'_j. The Mermin inequalities are precisely those for odd N.

Let M_N denote a Mermin inequality of $N = 2J + 1$ sites. Party i, with $i = 1, \ldots, N$ has a choice between two dichotomic measurements, A_i and A'_i. It is easily checked that for odd N, M_N contains only full correlators with an odd number of primes. We show, using symmetry arguments, that at the point of maximal quantum violation every correlator $\langle A_i \ldots A_j \rangle$ (involving an arbitrary number of measurements) that does not appear in M_N is identically zero. This automatically implies that any combination of N settings not appearing in the inequality define N random bits.

To see this, first take a specific N-body correlator not appearing in M_N, $\langle X_1 X_2 \ldots X_N \rangle$ where $X_i = A_i$ or A'_i but such that the total number of primed A is an even number. Denote the outcome of X_i by x_i. Choose any of the parties, say the first one, and denote by $Corr(X_1)$ the set of all correlators of arbitrary size containing X_1 plus possibly other settings X_i with $i > 1$. We would like to show that every element belonging to $Corr(X_1)$ is equal to zero for the unique distribution maximally violating the inequality. Let us consider the transformation $\mathcal{S}_1 : \{x_1 \mapsto -x_1$, and x_j untouched $\forall j > 1\}$. This maps $Corr(X_1) \mapsto -Corr(X_1)$. The terms in M_N remains unchanged if we complement $\mathcal{S}_1$ with $\mathcal{S}'_1 : \{x'_j \mapsto -x'_j \forall j > 1\}$, where we use $(A'_i)' = A_i$. In fact, note that for the original even primed term we started with, $\mathcal{S}'_1 \circ \mathcal{S}_1 \langle X_1 X_2 \ldots X_N \rangle = -\langle X_1 X_2 \ldots X_N \rangle$. The Mermin inequality consists only of odd-parity full-correlators. Any such a term can be obtained from $\langle X_1 X_2 \ldots X_N \rangle$ by swapping inputs at an odd number of places. However, the transformation $\mathcal{S}_1' \circ \mathcal{S}_1$ is such that at every site, either the outcome of A_i or A_i' flips sign but not both. Hence, $\mathcal{S}_1' \circ \mathcal{S}_1$ applied on any correlator obtained by an odd number of local swaps on $\langle X_1 X_2 \ldots X_N \rangle$ gains an additional factor of -1 *for each swapped site* relative to $\mathcal{S}_1' \circ \mathcal{S}_1 \langle X_1 X_2 \ldots X_N \rangle$. Thus, M_N remains unchanged. It remains to study the effect of $\mathcal{S}_1'$ on $Corr(X_1)$. Since $X_j' \notin Corr(X_1)$, this set is unmodified under $\mathcal{S}_1'$, so $\mathcal{S}_1' \circ \mathcal{S}_1$ maps $Corr(X_1) \mapsto -Corr(X_1)$. We then conclude from uniqueness that all the correlators in $Corr(X_1)$ must be zero. The same argument can be run for any party, and then for any full-correlator with an even number of primes, proving the result.

Case when N is even

It is worth mentioning first that similar arguments to above may be applied to the Bell-Klyshko inequalities for even N allowing the certification of $(N-1)$ bits of randomness.

However, there exist a different class of inequalities [HCLB11] for which full global randomness exists for even N. These inequalities are defined as,

$$\sum_x (-1)^{f(x)} \delta_0^{\bigoplus_{j=1}^N x_j} \langle A_1^{x_1} A_2^{x_2} \ldots A_N^{x_N} \rangle \leq C \tag{3.6}$$

where $f(x) = \bigoplus_{j=1}^{N-2} x_j (\bigoplus_{k=j+1}^{n-1} x_k)$ and $x_i \in \{0,1\}$ is the choice of measurement input. It can be shown that the local bound C is strictly lower than 2^{N-1}, the quantum limit.

Symmetry arguments can be applied [Hob] to show that maximal violation of this inequality provides full N bits of randomness for any input string not appearing in the inequality. The uniqueness for this inequality is discussed in the next section.

Case when $d > 2$

The examples considered so far have involved dichotomic observables. We finally study a simple generalization of the tripartite Mermin inequality [GLB+12] for arbitrary outcomes to demonstrate that our method works more generally. The inequality introduced by Grandjean *et al.* was,

$$\begin{aligned} \mathcal{M}_d = & \langle [A_2 - B_1 + C_1]_d \rangle + \langle [A_1 + B_2 - C_1]_d \rangle \\ & + \langle [-A_1 + B_1 + C_2]_d \rangle + \langle [-A_2 - B_2 - C_2 - 1]_d \rangle \\ & \geq d - 1. \end{aligned} \tag{3.7}$$

It can be easily verified that the same symmetry transformations used for the regular Mermin inequality of $d = 2$ can be used to certify 3 dits of randomness for any distribution maximally violating $\mathcal{M}_d$. However, uniqueness is unknown for this inequality and must remain at the level of an assumption.

3.4 Geometric interpretation

The previous argument crucially relies on the assumption that there is a unique quantum distribution attaining the maximal violation of a given Bell inequality. For some cases, such as Mermin $(N, 2, 2)$, this uniqueness has been proven [FFW11, WW01, Che04] and, then, it is no longer an assumption. Moreover, these arguments also apply to the class of inequalities of Eqn. 3.6-being full correlator inequalities- and so have a unique quantum saturation. For the chained inequality, we have numerical evidence using the techniques from [NPA07] that the distribution saturating it is unique in the $(2, 3, 2)$ and $(2, 4, 2)$ cases. The check involved using semi-definite programming to maximize different correlators and marginals under the constraint of maximal Bell violation (at the so-called $1 + AB$ level of the hierarchy). We found that no matter which correlator is maximized, the resultant distribution is the same up to numerical error.

From a geometrical point of view, it is natural to expect that the maximal violation of a generic Bell inequality is attained by a unique point. The set of quantum correlations defines a convex set in the space of probability distributions $P(a_1, \ldots, a_N|x_1, \ldots, x_N)$. A Bell inequality is a hyperplane in this space. The maximal quantum violation corresponds to the point in which the hyperplane, *i.e.* the Bell inequality, becomes tangent to the set of quantum correlations. Since the set is convex, this point is expected to be unique, in general. Of course, there may be situations for which this is not true. So far the only exceptions we have found from numerics are for *lifted* Bell inequalities. A tight Bell inequality of a smaller space can be lifted in a sense made precise in [Pir05] to a tight Bell inequality in a higher space, either with more parties, measurements or outcomes. For example, $(CHSH - 2)_{AB} \otimes C_1 \leq 0$ is a tight Bell inequality of $(3, 2, 2)$ in which party C only applies one measurement. It is easy to see that there are several quantum realizations attaining the maximal violation of this inequality. However, it may be argued that these Bell inequalities should be properly be considered as belonging to a lower dimensional space.

One should, however, be careful when following this geometrical intu-

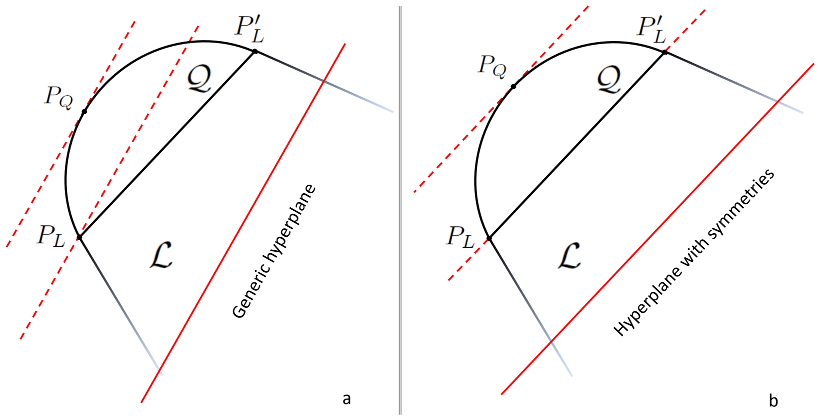

Figure 3.1: a) A generic hyperplane generally does not have symmetries and has a unique maximum in both the local and the quantum sets. b) A hyperplane with symmetries (such as the CHSH) precludes uniqueness in the local set but still allows for a unique maximum in the quantum set.

ition. Note that the previous argument does not make use of any quantum property. In fact, the set of classical correlations is also convex and, thus, a generic hyperplane is expected to become tangent at a unique extremal point, see Fig. 3.1a. However, randomness cannot be certified by classical correlations. The reason is that our method applies only to Bell inequalities that are symmetric under permutation of some of the measurement results, possibly assisted by permutations of measurements. It is easy to see that, within the local set, any symmetry under permutations of the results can be immediately used to construct another extremal and deterministic point saturating the inequality, thus breaking uniqueness (see Fig. 3.1b).

How do these considerations extend to general non-signalling correlations? While this is treated in the next chapter, we just pointed out here that the chained inequality allows certifying at most one bit of global randomness [JM05], as opposed to the two bits in the quantum case. This implies that there is more than one non-signalling point saturating the inequality. Understanding why randomness certification, based on uniqueness and symmetries, behaves so differently in the quantum set is an interesting question that deserves further investigation. From a speculative point of view, the fact that the quantum set is not a polytope, as opposed to the set of classical and non-signalling correlations, may play a key role in these considerations.

3.5 Discussion

Our argument is based on the simultaneous existence of uniqueness and symmetries. While in the classical case the needed symmetries immediately break the uniqueness of the maximal violation, this is no longer the case for quantum correlations, as implied by our results. Furthermore, we are yet to find an example where results from our symmetry arguments are in contradiction with numerical results where such computation was possible. For instance, for the I3322 [Fro81, Ś03, CG04] inequality, there are no symmetry arguments possible in order to certify maximal local randomness and, in fact, the known maximal quantum violation of the inequality gives non-uniform marginals [Ver].

While our simple recipe does not constitute a formal proof of randomness unless uniqueness is proven it still turns out to be very useful to find the right Bell inequalities and measurements allowing maximal randomness certification. Indeed, the results derived following our method can later be confirmed using the techniques from [PAM$^+$10b, NPA07]. In this sense, we are not aware of any Bell test leading to maximal randomness, local or global, that cannot be explained using our method. Our findings indicate that settings not appearing in the Bell inequality may have more global randomness than those appearing in the inequality. Moreover, using our method, we easily demonstrated the existence of Bell tests allowing maximal global randomness. Finally, our work opens new perspectives on the relation between randomness and non-locality that deserve further investigation.

This recipe demonstrates explicitly that maximal randomness is indeed attained in the set of quantum correlations and identifies several scenarios where this occurs. This is particular remarkable placed in the context of the results for the larger non-signalling set derived in the following chapter.

Chapter 4

Maximal non-signalling randomness

Quantum mechanics is the most successful theory of physics with remarkably precise predictions and successful application in the widest range of topics in physics. Why then would we study general non-signalling theories? One of the motivations for the study is that it allows us to understand properties thought to be generically ”quantum” in the context of these extended theories. For instance it turns out that monogamy of correlations, the impossibility of perfect cloning and privacy of correlations are a consequence of non-locality alone [MAG06] rather than a property of specifically quantum correlations. On the flip side, it has also been demonstrated that, given PR boxes [PR94], communication complexity [Dam05] and dynamics [SPG05] are rendered trivial. Such results further understanding of what may be considered properties genuinely associated with quantum theory alone. In addition, the quantification and comparison of properties between the two sets provide insight into why stronger-than-quantum correlations are forbidden in nature. Information causality [PPK$^+$09], macroscopic locality [NW10] and local orthogonality [FSA$^+$12] are examples of principles that have been proposed along these lines.

Intrinsic randomness is also known to be a generic feature of non-local correlations [PR94, MAG06, PAM$^+$10a, CW12]. However, there has been little work so far on quantifying it in the non-signalling set. Results are known for particular scenarios but systematic studies are limited on account of the rapidly increasingly complexity of the set of non-local correlations with increasing system size N, measurements M and outcomes d. The theme of this chapter is similar to the one in the previous chapter, but in the context of the non-signalling set. Namely, we investigate whether maximal global randomness is attained by extremal non-signalling correlations. If not, we are interested in finding quantitative upper bounds on the randomness. Before we present our results, we review results already known or that may be easily deduced from the literature thus far. Some of these were found in a different context and are often single examples rather than a systematic treatment.

The simplest scenario we can define is the $(2, 2, 2)$ where it is known that the PR boxes are the singular extremal boxes and encode only 1 bit of global randomness for any measurement. The extremal boxes characterizing the non-local sets for $(2, M, 2)$ [JM05] and $(2, 2, d)$ [BLM$^+$05] have also been found and the corresponding maximum allowed global randomness is 1 bit and 1 dit respectively. These two instances are somewhat "trivial" in the sense that they admit only a single class of extremal boxes in either case. The only scenario where a full non-trivial (in the sense of the existence of multiple non-equivalent classes) characterization of extremal boxes has been successful, to our knowledge, is the $(3, 2, 2)$ scenario [PBS11]. The maximal randomness among all these classes and measurement combinations is found to be $\log_2 6 = 2.58$ bits. Due to the non-trivial nature of this scenario, it provides important insight into the existence (or lack thereof) of maximal global randomness.

Certain other examples have also been studied in the context of monogamy and security of correlations. One of the first such was by Barrett and coworkers [BKP06] where they showed that all non-signalling correlations maximally violating the chained inequality [BC90] in the limit of infinite measurements encodes 1 dit of randomness for arbitrary local dimension in the scenario $(2, M, d)$. This result naturally generalized to $N - 1$ dits of randomness in chained inequalities for infinite measurements in the (N, M, d) scenario [AGCA12].

The common thread running through these results is that in no case do we find the maximal possible global randomness defined for the given scenario. Is this merely an accident resulting from our inability in characterizing more general examples? Or is it indicative of an underlying property of general non-signalling theories which distinguishes them from quantum theory? Our main result is to show that the answer to the latter question is "Yes". It turns out that randomness in general non-signalling theories is indeed forbidden from attaining the maximal possible randomness one could expect from the description of the scenario.

4.1 Results

The randomness of extremal non-signalling distributions for $(N, M, 2)$ is given by a guessing probability satisfying the bound $G(\mathbf{x}, P^{ex}) \geq \frac{1}{2^N-1}$. Since the algebraic maximum randomness for this scenario corresponds to a guessing probability of $G = 1/2^N$, this constitutes the first demonstration that non-signalling theories do not allow maximal randomness *even in principle*. For $(N, M, d > 2)$ the bound on the guessing probability is, $G(\mathbf{x}, P^{ex}) = \frac{1}{d^N-(d-1)^N}$. This compares with the algebraic maximum of $G = 1/d^N$.

We discuss the tightness of these bounds after the formal proofs.

4.2 Methods

4.2.1 Bound for $\mathbf{d = 2}$

Theorem 11. *(**Conditions for non-extremality for** $d = 2$). Let $P(\mathbf{a}|\mathbf{x})$ be a non-signalling distribution of a $(N, M, d = 2)$ scenario. If $\exists$ $\mathbf{x_0}$ such that $P(\mathbf{a}|\mathbf{x_0}) > 0$ $\forall$ $\mathbf{a}$ then P is non-extremal.*

Before proving the theorem, we recall that a general non-signalling distribution P satisfies the following properties:

1. Non-signalling: The marginals of P are well-defined.
2. Normalization: $\sum_{\mathbf{a}} P(\mathbf{a}|\mathbf{x}) = 1$ for all $\mathbf{x}$.
3. Non-negativity: $P(\mathbf{a}|\mathbf{x}) \geq 0$.

Furthermore, for $d = 2$, we can use the same correlator decomposition for P as in the previous chapter. We restate it here for convenience,

$$\begin{aligned} P(\mathbf{a}|\mathbf{x}) = \frac{1}{2^N}(1 + \sum_{i=1}^{N} a_i \langle A_i \rangle + \sum_{i<j} a_i a_j \langle A_i A_j \rangle \\ + \sum_{i<j<k} a_i a_j a_k \langle A_i A_j A_k \rangle + \cdots + a_1 a_2 \ldots a_N \langle A_1 A_2 \ldots A_N \rangle), \end{aligned} \tag{4.1}$$

where all k-body correlators for $1 \leq k \leq N$ satisfy $|\langle . \rangle_k| \leq 1$. Any P defined by this parametrization satisfies the first two conditions and only the third needs to be additionally imposed.

Proof. (*Theorem 11*) We demonstrate the non-extremality of P by explicit decomposition. In particular, we show that under the given assumptions we can construct two non-signalling distributions P_+ and P_- such that,

$$P(\mathbf{a}|\mathbf{x}) = \frac{1}{2}P_+(\mathbf{a}|\mathbf{x}) + \frac{1}{2}P_-(\mathbf{a}|\mathbf{x}). \tag{4.2}$$

Let $\epsilon = \min\{P_{min}(\mathbf{x_0}), 1-|\langle A_1 \dots A_N\rangle_{\mathbf{x_0}}|\}$ where $P_{min}(\mathbf{x_0}) = \min_{\mathbf{a}} P(\mathbf{a}|\mathbf{x_0})$. The full correlator in question cannot take its extremal values of ±1 since otherwise the definition $\langle A_1 \dots A_N\rangle_{\mathbf{x_0}} = P(a_1 \cdot a_2 \dots a_N = 1|\mathbf{x_0}) - P(a_1 \cdot a_2 \dots a_N = -1|\mathbf{x_0})$ implies many outcomes of zero probability in $P(\mathbf{a}|\mathbf{x_0})$ violating the assumptions in the theorem. These observations provide us the strict positivity $\epsilon > 0$. We can now define each new distribution (using its equivalent correlator parametrization) as identical to P *except only* the full-correlator corresponding to $\mathbf{x_0}$. This we set to be respectively $\langle A_1A_2 \dots A_N\rangle^{\pm}_{\mathbf{x_0}} = \langle A_1A_2 \dots A_N\rangle_{\mathbf{x_0}} \pm \epsilon/2$. It only remains to show that $P_\pm$ are well-defined non-signalling distributions.

Since the only correlator in which they differ from P is $\langle A_1 \dots A_N\rangle^{\pm}_{\mathbf{x_0}}$, we must show that $P_\pm(\mathbf{a}|\mathbf{x_0}) > 0$ for all $\mathbf{a}$.

This follows from the relations,

$$\begin{aligned}
P_\pm(\mathbf{a}|\mathbf{x_0}) &= P(\mathbf{a}|\mathbf{x_0}) \pm \frac{1}{2^N} a_1 \dots a_N \ \epsilon/2 \\
&\geq P_{min}(\mathbf{x_0}) - \frac{1}{2^N}|a_1 \dots a_N \ \epsilon/2| \\
&= P_{min}(\mathbf{x_0}) - \frac{1}{2^{N+1}} \ \epsilon \\
&\geq P_{min}(\mathbf{x_0}) - \frac{1}{2^{N+1}} P_{min}(\mathbf{x_0}) \\
&> 0,
\end{aligned}$$

from the definition of ϵ. □

From Theorem 11, it immediately follows that if a non-signalling distribution $P(\mathbf{a}|\mathbf{x})$ is uniformly distributed in any input $\mathbf{x} = \mathbf{x_0}$, then it is non-extremal. Thus, maximal randomness is forbidden in any extremal distribution. The convex decomposition of Eqn. 4.2 can be effected unless $P(\mathbf{a}|\mathbf{x_0}) = 0$ for some outcome string $\mathbf{a}$. This may be used to upper bound the randomness for an extremal distribution. Indeed, if the probability is zero for *for just one* outcome corresponding to $\mathbf{x_0}$ then we can lower bound the guessing probability assuming a uniform distribution on the other outcomes. This is given by,

$$G(\mathbf{x_0}, P^{ex}) = \max_{\mathbf{a}} P^{ex}(\mathbf{a}|\mathbf{x_0}) \geq \frac{1}{2^N - 1}. \tag{4.3}$$

This is effectively an upper bound on the randomness defined by the min-entropy. It provides an in-principle upper bound to attainable randomness within the set of non-signalling correlations and suggests that counting the least number of zeros in an extremal distribution is a means of bounding randomness. We use this observation in the following generalization to (N, M, d).

4.2.2 Bound for $\mathbf{d > 2}$

For scenarios $(N, M, d > 2)$, the parametrization of Eqn. 4.1 does not provide adequate free parameters. While it is possible to write an equivalent parametrization, we opt for a different approach. We parameterize the given distribution with the least number of required terms by normalization and non-signalling.

The full distribution is $P(a_1, a_2, \ldots, a_N|x_1, x_2, \ldots, x_N)$ where $1 \leq x_i \leq M$ and $0 \leq a_i \leq d-1$ for every party. P is subject to the non-signalling, normalization and non-negativity as before. These constraints reduce the number of free parameters and P can be parameterized in terms of a truncated set. We fix some notation which will be useful in what follows.

1. The input string we extract randomness from is $\mathbf{x_0}$.
2. $\bar{\mathbf{a}} = \{\bar{a}_1, \ldots, \bar{a}_N\}$ where the outcomes take the truncated values $0 \leq \bar{a}_i \leq d-2$.
3. $S^{N-1}(\bar{\mathbf{a}}) = \{P(\bar{a}_1, \ldots, \bar{a}_{i-1}, \bar{a}_{i+1}, \ldots, \bar{a}_N|\mathbf{x_0})$ for all $1 \leq i \leq N\}$ for a chosen outcome $\bar{\mathbf{a}}$. This is the set of all the $N-1$ party marginals for a given outcome string $\bar{\mathbf{a}}$.
4. $\mathbb{T} = \{P(\bar{\mathbf{a}}|\mathbf{x_0})\}$ for all the outcomes $\bar{\mathbf{a}}$.

The following example aids understanding the new notation. Let $N = 3$, $M = 2$, $d = 3$ and $\mathbf{x_0} = (0, 0, 0)$ where the measurement labels are taken to be 0 or 1 for each party. Then, the outcomes satisfy $0 \leq a_1, a_2, a_3 \leq 2$ and the truncated set, $0 \leq \bar{a}_1, \bar{a}_2, \bar{a}_3 \leq 1$. In this case,

$$S^2(\bar{\mathbf{a}} = (0,1,1)) = \{P_{BC}(1,1|0,0), P_{AC}(0,1|0,0), P_{AB}(0,1|0,0)\}$$
$$\mathbb{T} = \{P_{ABC}(0,0,0|\mathbf{x_0}), P_{ABC}(0,0,1|\mathbf{x_0}), \ldots, P_{ABC}(1,1,1|\mathbf{x_0})\},$$

so that $|\mathbb{T}| = 8$.

Now, $P(\mathbf{a}|\mathbf{x})$ may be parameterized in terms of a truncated set of its terms as,

$$\begin{aligned} P(\mathbf{a}|\mathbf{x}) = h[&P(\bar{a}_1|x_1), \ldots, P(\bar{a}_N|x_N); \\ &P(\bar{a}_i, \bar{a}_j|x_i, x_j) \text{ for all } i < j; \\ &P(\bar{a}_i, \bar{a}_j, \bar{a}_k|x_i, x_j, x_k) \text{ for all } i < j < k; \\ &\vdots \\ &P(\bar{a}_1, \ldots, \bar{a}_N|x_1, \ldots, x_N)], \end{aligned} \tag{4.4}$$

with the terms being related by the usual ordering between marginals, of the form, $P(\bar{a}_1|x_1) \geq P(\bar{a}_1, \bar{a}_2|x_1, x_2) \geq \cdots \geq P(\bar{a}_1, \ldots, \bar{a}_N|x_1, \ldots, x_N)$ etc. A set of parameters (on the right hand side) in Eqn. 4.4 lying in the interval $[0, 1]$ and satisfying this ordering relation specifies a valid non-signalling distribution $P(\mathbf{a}|\mathbf{x})$. We can now state the next theorem on the non-extremality of P for arbitrary outcomes.

Theorem 12. ***(Conditions for non-extremality for*** $\mathbf{d > 2}$***).*** *Let* $P(\mathbf{a}|\mathbf{x})$ *be a non-signalling distribution of the* (N, M, d) *scenario. If* $\exists\ \mathbf{a'}$ *such that an element of* $\mathbb{T}$ *say* $P(\bar{\mathbf{a}}'|\mathbf{x_0}) > 0$ *and* $P(\bar{\mathbf{a}}'|\mathbf{x_0}) \neq \alpha\ \forall\ \alpha \in S^{N-1}(\bar{\mathbf{a}}')$*, then* P *is non-extremal.*

Proof. The proof is constructive.

We can define two new distributions $P_{1,2}(\mathbf{a}|\mathbf{x})$ parameterizing them as in Eqn. 4.4. We set all parameters of $P_{1,2}(\mathbf{a}|\mathbf{x})$ equal to the corresponding ones for $P(\mathbf{a}|\mathbf{x})$, except for just the N-partite parameter in question. This term is used for the decomposition:

$$\begin{aligned} P_1(\bar{\mathbf{a}}'|\mathbf{x_0}) &= P(\bar{\mathbf{a}}'|\mathbf{x_0}) + \epsilon/2 \\ P_2(\bar{\mathbf{a}}'|\mathbf{x_0}) &= P(\bar{\mathbf{a}}'|\mathbf{x_0}) - \epsilon/2 \end{aligned} \tag{4.5}$$

where $\epsilon = \min\{P(\bar{\mathbf{a}}'|\mathbf{x_0}),\ D(P(\bar{\mathbf{a}}'|\mathbf{x_0}), S^{N-1}(\bar{\mathbf{a}}'))\}$. For a given distribution p and a set S the quantity D is defined by $D(p, S) = \min\{|p - \alpha| \text{ such that } \alpha \in S\}$ and is a notion of distance of p from S. Clearly $\epsilon > 0$.

The new distributions satisfy non-negativity by construction. Normalization is accounted for by the terms *not* appearing in the parametrization and they satisfy non-signalling since the marginals of $P_{1,2}$ are identical to those of P. The ordering relationship between the marginals are also preserved for this choice of ϵ. Thus, we can effect the decomposition,

$$P(\mathbf{a}|\mathbf{x}) = \frac{1}{2}P_1(\mathbf{a}|\mathbf{x}) + \frac{1}{2}P_2(\mathbf{a}|\mathbf{x}). \tag{4.6}$$

□

Theorem 13. *(**Maximum randomness for $\mathbf{d > 2}$**). Given an extremal non-signalling distribution $P^{ex}(\mathbf{a}|\mathbf{x})$ in (N, M, d), the maximal randomness for any input $\mathbf{x} = \mathbf{x_0}$ is given by, $G(\mathbf{x_0}, P^{ex}) = \frac{1}{d^N - (d-1)^N}$.*

We prove the result by showing that there must exist at least $(d-1)^N$ outcomes with probability zero in $P(\mathbf{a}|\mathbf{x_0})$. Then the guessing probability is least for a uniform distribution over the other terms *i.e.* $G(\mathbf{x_0}, P^{ex}) \geq \frac{1}{d^N - (d-1)^N}$, proving the result. Towards this end, we observe that $\mathbb{T}$ defined earlier has $|\mathbb{T}| = (d-1)^N$ terms. Thus, if each element of $\mathbb{T}$ is identically 0, then the result follows trivially. However, more generally, there will exist non-zero elements in $\mathbb{T}$. In this case, Theorem 12 allows us to conclude that any such non-zero element must necessarily equal at least one of its $(N-1)$-party marginals to render decomposition in Eqn. 4.6 impossible. We now provide the formal proof of Theorem 13 showing that $(d-1)^N$ zeros can be guaranteed in $P^{ex}(\mathbf{a}|\mathbf{x_0})$ by mapping every non-zero element of $\mathbb{T}$ to a unique term guaranteed to be 0.

Proof. (*Theorem 13*). We frame the proof as a protocol scanning $\mathbb{T}$ for non-zero elements and then updating it, replacing the said non-zero term with another term guaranteed to be zero under the assumed constraints. At the end of the protocol $\mathbb{T}$ can be mapped to a tensor of all zeros $\mathbf{0}$ proving our result.

It is however useful to first illustrate the point explicitly by revisiting the previous example of $(N = 3, M = 2, d = 3)$. Suppose the non-zero element of $\mathbb{T}$ is $P_{ABC}(\bar{\mathbf{a}} = (0, 1, 1)|0, 0, 0)$ and is equal to a marginal $P_{AC}(0, 1|0, 0) \in S^2(\bar{\mathbf{a}} = (0, 1, 1))$.

In this case, we necessarily get two other elements identically equal to zero. One of these belongs to the set $\mathbb{T}$ and the other does not. Namely, $0 = P_{ABC}(0, 0, 1|0, 0, 0) \in \mathbb{T}$ while $0 = P_{ABC}(0, 2, 1|0, 0, 0) \notin \mathbb{T}$. The latter term is used to replace the original non-zero term in $\mathbb{T}$. We can thus systematically associate a term guaranteed to be zero for every non-zero term to create a mapping between $\mathbb{T}$ and $\mathbf{0}$.

The general protocol is the following.

1. Scan $\mathbb{T}$ until a non-zero term is found. We denote this term by $p(\bar{a}_1, \ldots, \bar{a}_N)$.

2. Compare $p(\bar{a}_1, \ldots, \bar{a}_N)$ with the set of its marginals $S^{N-1}(\bar{a}_1, \ldots, \bar{a}_N)$. It must necessarily equal at least one of the marginals (consistent with

the rest of the terms in $\mathbb{T}$) taking into account Theorem 12 and that $P(\mathbf{a}|\mathbf{x})$ is extremal. Suppose,

$$p(\bar{a}_1, \dots, \bar{a}_N) = p(\bar{a}_1, \dots, \bar{a}_{i-1}, \bar{a}_{i+1}, \dots, \bar{a}_N), \tag{4.7}$$

for some i. This implies,

i. certain zero terms in $\mathbb{T}$. Namely, $p(\bar{a}_1, \dots, \bar{a}_{i-1}, a_i, \bar{a}_{i+1}, \dots, \bar{a}_N) = 0$ for all $0 \leq a_i \neq \bar{a}_i \leq d-2$.

ii. a zero term *not* in $\mathbb{T}$. Namely, $p(\bar{a}_1, \dots, \bar{a}_{i-1}, d{-}1, \bar{a}_{i+1}, \dots, \bar{a}_N) = 0$.

3. Update $\mathbb{T}$ by setting all terms given in Step 2.i explicitly equal to zero. Also replace the original non-zero term $p(\bar{a}_1, \dots, \bar{a}_N)$ by the term in Step 2.ii, $p(\bar{a}_1, \dots, \bar{a}_{i-1}, d-1, \bar{a}_{i+1}, \dots, \bar{a}_N)$ This replacement is unique of all other terms already in $\mathbb{T}$. This is because all elements that map to $p(\bar{a}_1, \dots, \bar{a}_{i-1}, d{-}1, \bar{a}_{i+1}, \dots, \bar{a}_N)$ are exactly those in Step 2.i, which are set to zero at this iteration. Hence, further iterations of this protocol skips those elements.

4. Go to Step 1 and repeat until all non-zero elements in $\mathbb{T}$ are replaced by zero elements.

□

4.3 Discussion

We have demonstrated that randomness in the general set of non-signalling correlations is generically upper bounded, in contrast to the quantum set. These results add a layer of subtlety to the notion of intrinsic randomness being a general property of non-signalling correlations. While this is true, we have shown that there are important quantitative differences between the two sets.

Our results are not tight. The bounds for $(2, 2, 2)$ and $(3, 2, 2)$ scenarios directly observed from the given classes of extremal boxes appear to satisfy $G(\mathbf{x}, P^{ex}) \geq \frac{1}{2^N - 2}$. A result of this form is derived in [BLM$^+$05] using methods that ours bears a close relation to. For $N = 2$, we have derived $(d-1)^2$ zeros while they demonstrate the marginally larger $d(d-1)$ zeros. The tightness in their bound relative to ours arises because their subject of study was a general characterization of extremal boxes in the specific scenario of $(2, 2, d)$, while our scenario is completely general.

However, our bound is tight in the limit $d \to \infty$ for a fixed N. The bound given by Theorem 13 in this limit upper bounds the randomness in extremal boxes to $(N-1)$-dits whereas the results of [AGCA12] indicate that this is also a lower bound in the specific instance they consider. It is an open question if quantum distributions attain the maximal of N-dits in this scenario. If so, the difference in the maximum randomness between the quantum and non-signalling sets becomes large in this limit.

These results also connect to the symmetry argument of the previous chapter. The non-existence of maximal global randomness implies that either symmetries of Bell inequalities allowing full global randomness or uniqueness of the non-signalling Bell violation does not hold for every non-signalling scenario. This is indeed known to be true in those scenarios where the extremal boxes have been characterized. In the $(2,2,2)$ scenario, for instance, the PR box is the unique extremal box maximally violating the CHSH inequality but as we have already seen, there are no symmetries that allow global randomness for this inequality. Another example is the class of Mermin inequalities. Here, the symmetries do indeed exist allowing certification of global randomness, however uniqueness does not. There are multiple classes of inequalities that maximally violate the Mermin inequality each allowing several symmetries. In this sense our present results provide further evidence of the intimate relationship between symmetries and randomness.

Chapter 5

Full randomness amplification

Understanding whether nature is perfectly pre-determined or there are intrinsically random processes is a fundamental question that has attracted the interest of multiple thinkers, ranging from philosophers and mathematicians to physicists or neuroscientists. Today this question is also important from a practical perspective, as random bits constitute a valuable resource for applications such as cryptographic protocols, gambling, or the numerical simulation of physical and biological systems.

In the Introduction we discussed the notion of intrinsic versus classical randomness at length. Since they form the very crux of the work in this chapter, we quickly recall the arguments made and expand on the most significant parts.

Randomness observed in classical systems is not intrinsic to the theory but merely a manifestation of our imperfect description of the system. In other words, it is what we have termed classical or deterministic randomness.

It was the development of quantum physics that first challenged this deterministic view of nature. It was suggested that quantum mechanics could be *completed* with hidden variables [EPR35] which would again render nature fully deterministic. However Bell's no-go theorem [Bel64] conclusively excluded the possibility of such a project. More precisely, Bell found that all hidden-variable theories compatible with a local causal structure predict that any correlations among space-like separated events satisfy a series of inequalities, which have now come to be known as Bell inequalities. Bell inequalities, in turn, are violated by some correlations among quantum particles. This is, of course, non-locality.

However non-locality alone does not necessarily imply the existence of fully unpredictable processes in nature. The reasons behind this are subtle. First of all, unpredictable processes could be certified only if the non-signalling principle holds. In fact, Bohm's theory is both deterministic and able to reproduce all quantum predictions [Boh52], but it is incompatible with non-signalling. Thus, we assume throughout the validity of the non-signalling principle. Yet, even within the non-signalling framework, it is still not possible to infer the existence of fully random processes from the mere observation of non-local correlations. This is due to the fact that Bell tests require measurement settings chosen at random, but the actual randomness in such choices can never be certified. The extremal example is given when the settings are determined in advance. Then, any Bell violation can easily be explained in terms of deterministic models. As a matter of fact, super-deterministic models, which postulate that all phenomena in the universe, including our own mental processes, are fully pre-programmed, are by definition impossible to rule out.

These considerations imply that the strongest result on the existence of randomness one can hope for using quantum non-locality is stated by the following possibility: Given a source that produces an arbitrarily small but non-zero amount of randomness, can one still certify the existence of completely random processes? Our main result is to provide an affirmative answer to this question. Our results, then, imply that the existence of correlations as those predicted by quantum physics forces us into a dichotomic choice: Either we postulate super-deterministic models in which all events in nature are fully pre-determined, or we accept the existence of fully unpredictable events.

Besides the philosophical and physics-foundational implications, our results provide a protocol for full randomness amplification using quantum non-locality. Randomness amplification is an information-theoretic task whose goal is to use an input source of imperfectly random bits to produce perfect random bits. Santha and Vazirani proved that randomness amplification is impossible using classical resources [SV86]. This is in a sense intuitive, in view of the absence of any intrinsic randomness in classical physics. In the quantum regime, randomness amplification has been recently studied by Colbeck and Renner [CR12b]. They proved how input bits with very high initial randomness can be mapped into arbitrarily pure random bits, and conjectured that randomness amplification should be possible for any initial randomness [CR12b]. Our results also solve this conjecture, as we show that quantum non-locality can be exploited to attain *full randomness amplification.*

Finally, before presenting our results, it is worth commenting on previous works on randomness in connection with quantum non-locality. In [PAM$^+$10b] it was shown how to bound the intrinsic randomness generated in a Bell test. These bounds can be used for device-independent randomness expansion, following a proposal by Colbeck [Col07], and to achieve a quadratic expansion of the amount of random bits (see [AMP12, PM11, FGS11, VV12a] for further works on device-independent randomness expansion). Note however that, in randomness expansion, one assumes instead, from the very beginning, the existence of an input seed of free random bits, and the main goal is to expand this into a larger sequence. The figure of merit there is the ratio between the length of the final and initial strings of free random bits. Finally, other recent works have analysed how a lack of randomness in the measurement choices affects a Bell test [KPB06, BG10, Hal10] and the randomness generated in it [KHS$^+$12].

5.1 Results

5.1.1 Definition of the scenario

From an information perspective, our goal is to construct a protocol for full randomness amplification based on quantum non-locality. In randomness amplification, one aims at producing random bits that are arbitrarily uncorrelated from all the events that may have been a potential cause of them, i.e. arbitrarily free, from many uses of an input source $\mathcal{S}$ of imperfectly random bits. In general, $\mathcal{S}$ produces a sequence of bits $x_1, x_2, \ldots x_j, \ldots$, with $x_j = 0$ or 1 for all j, see Fig. 5.1. Each bit j contains some randomness, in the sense that the probability $P\left(x_j|rest\right)$ that it takes a given value x_j, conditioned on the rest of the bits produced by the source and any pre-existing variable e, is such that

$$\epsilon \leq P\left(x_j|rest\right) \leq 1-\epsilon \tag{5.1}$$

for all j and e, where $0 < \epsilon \leq 1/2$. Here, $rest = \{x_i\}_{i\neq j} \cup e$ where the variable e can correspond to any event that could be a possible cause of bit x_j (different from the rest of the bits generated by the source). Therefore, e represents events contained in the space-time region lying outside the future light-cone of x_j. Free random bits correspond to $\epsilon = \frac{1}{2}$; while deterministic ones, i.e. those predictable with certainty by an observer with access to e, to $\epsilon = 0$. More precisely, when $\epsilon = 0$ the bound (5.1) is trivial and no randomness can be certified. We refer to $\mathcal{S}$ as an ϵ-source, and to any bit satisfying (5.1) as an ϵ-free bit. The aim is then to generate, from arbitrarily

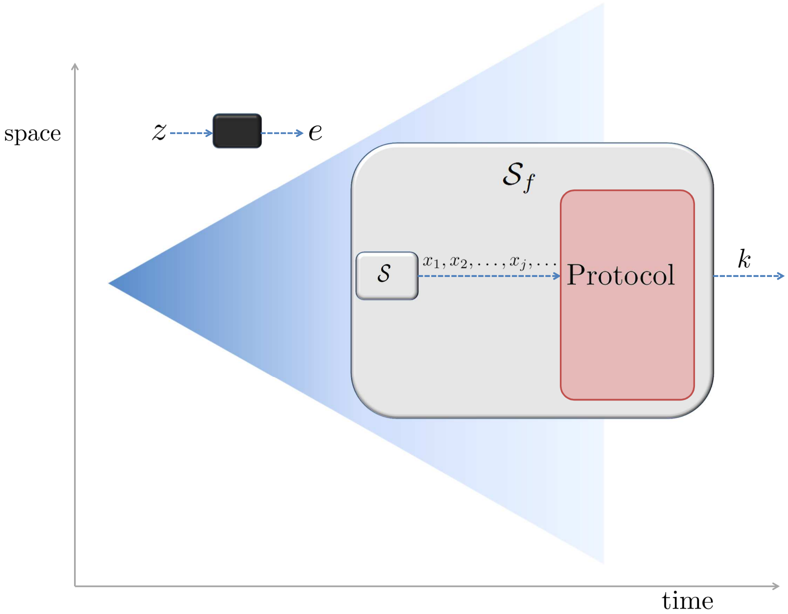

Figure 5.1: **Local causal structure and randomness amplification**. A source $\mathcal{S}$ produces a sequence $x_1, x_2, \ldots x_j, \ldots$ of imperfect random bits. The goal of randomness amplification is to produce a new source $\mathcal{S}_f$ of perfect random bits, that is, to process the initial bits to get a final bit k fully uncorrelated (free) from any potential cause of it. All space-time events outside the future light-cone of k may have been in its past light-cone before and therefore constitute a potential cause of it. Any such event can be modelled by a measurement z, with an outcome e, on some physical system. This system may be under the control of an adversary Eve, interested in predicting the value of k.

many uses of $\mathcal{S}$, a final source $\mathcal{S}_f$ of ϵ_f arbitrarily close to 1/2. If this is possible, no cause e can be assigned to the bits produced by $\mathcal{S}_f$, which are then fully unpredictable. Note that efficiency issues, such as the rate of uses of $\mathcal{S}$ required per final bit generated by $\mathcal{S}_f$ do not play any role in randomness amplification. The relevant figure of merit is just the quality, measured by ϵ_f, of the final bits. Thus, without loss of generality, we restrict our analysis to the problem of generating a single final free random bit k.

The randomness amplification protocols we consider exploit quantum non-locality. This idea was introduced in [CR12b], where a protocol was presented in which the source $\mathcal{S}$ is used to choose the measurement settings by two distant observers, Alice and Bob, in a Bell test [BC90] involving two entangled quantum particles. The measurement outcome obtained by one of the observers, say Alice, in one of the experimental runs (also chosen with $\mathcal{S}$) defines the output random bit. Colbeck and Renner proved how input bits with very high randomness, of $0.442 < \epsilon \leq 0.5$, can be mapped into arbitrarily free random bits of $\epsilon_f \to 1/2$. In our case, the input ϵ-source $\mathcal{S}$ is used to choose the measurement settings in a multipartite Bell test involving a number of observers that depends both on the input ϵ and the target ϵ_f. After verifying that the expected Bell violation is obtained, the measurement outcomes are combined to define the final bit k. For pedagogical reasons, we adopt a cryptographic perspective and assume the worst-case scenario where all the devices we use may have been prepared by an adversary Eve equipped with arbitrary non-signalling resources, possibly even supra-quantum ones. In the preparation, Eve may have also had access to $\mathcal{S}$ and correlated the bits it produces with some physical system at her disposal, represented by a black box in Fig. 5.1. Without loss of generality, we can assume that Eve can reveal the value of e at any stage of the protocol by measuring this system. Full randomness amplification is then equivalent to proving that Eve's correlations with k can be made arbitrarily small.

Let us make two clarifications about the causal structure of the protocol. Firstly, note that the bits used by each party are drawn from local sources. However, since all those sources are not independent but hold correlations between each other, they form collectively a Santha-Vazirani source with the properties previously described. Secondly, note that throughout the protocol, all measurements define space-like separated events.

5.1.2 Partial randomness from GHZ paradoxes

One of the Bell tests for which quantum correlations achieve the maximal non-signalling violation are known as Greenberger-Horne-Zeilinger (GHZ) paradoxes [GHZ89] and are necessary for full randomness amplification. This is due to the fact that unless the maximal non-signalling violation is attained, for sufficiently small ϵ, Eve may fake the observed correlations with classical deterministic resources. Nevertheless, GHZ paradoxes are not sufficient. In fact, given (i) maximal violation of the 3-party GHZ paradox [GHZ89] and (ii) any function of the measurement outcomes, it is always possible to find non-signalling correlations that assign a deterministic value to the said function. For example, if the function selects outcome c, one can have a PR-box shared between A and B with C being deterministic since such correlations are known to satisfy point (i). This observation can be checked for all unbiased functions mapping $\{0,1\}^3$ to $\{0,1\}$ (there are $\binom{8}{4}$ of those) through a linear program analogous to the one used in the proof of the

Lemma below.

For five parties though, the latter happens not to hold any longer. Consider now any correlations attaining the maximal violation of the five-party Mermin inequality [Mer90]. In each run of this Bell test, measurements (inputs) $\mathbf{x} = (x_1, \ldots, x_5)$ on five distant black boxes generate 5 outcomes (outputs) $\mathbf{a} = (a_1, \ldots, a_5)$, distributed according to a non-signalling conditional probability distribution $P(\mathbf{a}|\mathbf{x})$. Both inputs and outputs are bits, as they can take two possible values, $x_i, a_i \in \{0, 1\}$ with $i = 1, \ldots, 5$. The inequality can be written as

$$\sum_{\mathbf{a},\mathbf{x}} I(\mathbf{a}, \mathbf{x}) P(\mathbf{a}|\mathbf{x}) \geq 6 \ , \tag{5.2}$$

with coefficients

$$I(\mathbf{a}, \mathbf{x}) = (a_1 \oplus a_2 \oplus a_3 \oplus a_4 \oplus a_5)\, \delta_{\mathbf{x} \in \mathcal{X}_0} + (a_1 \oplus a_2 \oplus a_3 \oplus a_4 \oplus a_5 \oplus 1)\, \delta_{\mathbf{x} \in \mathcal{X}_1} \ , \tag{5.3}$$

where

$$\delta_{\mathbf{x} \in \mathcal{X}_s} = \left\{ \begin{array}{ll} 1 & \text{if } \mathbf{x} \in \mathcal{X}_s \\ 0 & \text{if } \mathbf{x} \notin \mathcal{X}_s \end{array} \right. ,$$

and

$$\mathcal{X}_0 = \Big\{ \mathbf{x} \ \Big| \ \sum_{i=1}^{5} x_i = 1 \Big\} \cup \Big\{ \mathbf{x} \ \Big| \ \sum_{i=1}^{5} x_i = 5 \Big\}, \tag{5.4}$$

$$\mathcal{X}_1 = \Big\{ \mathbf{x} \ \Big| \ \sum_{i=1}^{5} x_i = 3 \Big\}. \tag{5.5}$$

That is, only half of all possible combinations of inputs, namely those in $\mathcal{X} = \mathcal{X}_0 \cup \mathcal{X}_1$, appear in the Bell inequality.

The maximal, non-signalling and algebraic, violation of the inequality corresponds to the situation in which the left-hand side of (5.2) is zero. The key property of inequality (5.2) is that its maximal violation can be attained by quantum correlations and furthermore, one can construct a function of the outcomes that is not completely determined. Take the bit corresponding to the majority-vote function of the outcomes of any subset of three out of the five observers, say the first three. This function is equal to zero if at least two of the three bits are equal to zero, and equal to one otherwise. We show that Eve's predictability on this bit is at most 3/4. We state this result in the following Lemma:

Lemma. *Let a five-party non-signalling conditional probability distribution* $P(\mathbf{a}|\mathbf{x})$ *in which inputs* $\mathbf{x} = (x_1, \ldots, x_5)$ *and outputs* $\mathbf{a} = (a_1, \ldots, a_5)$ *are bits. Consider the bit* $\mathrm{maj}(\mathbf{a}) \in \{0, 1\}$ *defined by the majority-vote function of any subset consisting of three of the five measurement outcomes, say the first three,* a_1*,* a_2 *and* a_3*. Then, all non-signalling correlations attaining the maximal violation of the 5-party Mermin inequality are such that the probability that* $\mathrm{maj}(\mathbf{a})$ *takes a given value, say 0, is bounded by*

$$1/4 \leq P\left(\mathrm{maj}(\mathbf{a}) = 0\right) \leq 3/4. \tag{5.6}$$

Proof. This result was obtained by solving a linear program. Therefore, the proof is numeric, but exact. Formally, let $P(\mathbf{a}|\mathbf{x})$ be a 5-partite non-signalling probability distribution. For $\mathbf{x} = \mathbf{x}_0 \in \mathcal{X}$, we performed the maximization,

$$\begin{aligned} P_{max} = \max_{P} \;& P(\mathrm{maj}(\mathbf{a}) = 0|\mathbf{x}_0) \\ & \text{subject to} \\ & I(\mathbf{a}, \mathbf{x}) \cdot P(\mathbf{a}|\mathbf{x}) = 0 \end{aligned} \tag{5.7}$$

which yields the value $P_{max} = 3/4$. Since the same result holds for $P(\mathrm{maj}(\mathbf{a}) = 1|\mathbf{x}_0)$, we get the bound $1/4 \leq P(\mathrm{maj}(\mathbf{a}) = 0) \leq 3/4$.

□

As a further remark, note that a lower bound to P_{max} can easily be obtained by noticing that one can construct conditional probability distributions $P(\mathbf{a}|\mathbf{x})$ that maximally violate 5-partite Mermin inequality (5.2) for which at most one of the output bits (say a_1) is deterministically fixed to either 0 or 1. If the other two output bits (a_2, a_3) were to be completely random, the majority-vote of the three of them $\mathrm{maj}(a_1, a_2, a_3)$ could be guessed with a probability of 3/4. Our numerical results say that this turns out to be an optimal strategy.

The previous lemma strongly suggests that, given an ϵ-source with any $0 < \epsilon \leq 1/2$ and quantum five-party non-local resources, it should be possible to design a protocol to obtain an ϵ_i-source of $\epsilon_i = 1/4$. We do not explore this possibility here, but rather use the partial unpredictability in the five-party Mermin Bell test as building block of our protocol for full randomness amplification. To complete it, we must equip it with two essential components: (i) an *estimation procedure* that verifies that the untrusted devices do yield the required Bell violation; and (ii) a *distillation procedure* that, from sufficiently many ϵ_i-bits generated in the 5-party Bell experiment, distills a single final ϵ_f-source of $\epsilon_f \to 1/2$. Towards these ends, we consider a more complex Bell test involving N groups of five observers (quintuplets) each.

5.1.3 A protocol for full randomness amplification

Our protocol for randomness amplification uses as resources the ϵ-source $\mathcal{S}$ and $5N$ quantum systems. Each of the quantum systems is abstractly modelled by a black box with binary input x and output a. The protocol processes classically the bits generated by $\mathcal{S}$ and by the quantum boxes. The result of the protocol is a classical symbol k, associated to an abort/no-abort decision. If the protocol is not aborted, k encodes the final output bit, with possible values 0 or 1. However when the protocol is aborted, the symbol $\varnothing$ is assigned to k (instead of a numerical value), representing the fact that the bit is empty. The formal steps of the protocol are, see also Fig. 5.2:

1. $\mathcal{S}$ is used to generate N quintuple-bits $\mathbf{x}_1, \dots \mathbf{x}_N$, which constitute the inputs for the $5N$ boxes. The boxes then provide N output quintuple-bits $\mathbf{a}_1, \dots \mathbf{a}_N$.

2. The quintuplets such that $\mathbf{x} \notin \mathcal{X}$ are discarded. The protocol is aborted if the number of remaining quintuplets is less than $N/3$.

3. The quintuplets left after step 2 are organized in N_b blocks each one having N_d quintuplets. The number N_b of blocks is chosen to be a power of 2. For the sake of simplicity, we relabel the index running over the remaining quintuplets, namely $\mathbf{x}_1, \dots \mathbf{x}_{N_b N_d}$ and outputs $\mathbf{a}_1, \dots \mathbf{a}_{N_b N_d}$. The input and output of the j-th block are defined as $y_j = (\mathbf{x}_{(j-1)N_d+1}, \dots \mathbf{x}_{(j-1)N_d+N_d})$ and $b_j = (\mathbf{a}_{(j-1)N_d+1}, \dots \mathbf{a}_{(j-1)N_d+N_d})$ respectively, with $j \in \{1, \dots, N_b\}$. The random variable $l \in \{1, \dots N_b\}$ is generated by using $\log_2 N_b$ further bits from $\mathcal{S}$. The value of l specifies which block (b_l, y_l) is chosen to generate k, i.e. the distilling block. We define $(\tilde{b}, \tilde{y}) = (b_l, y_l)$. The other $N_b - 1$ blocks are used to check the Bell violation.

4. The function
$$r[b, y] = \begin{cases} 1 & \text{if } \; I(\mathbf{a}_1, \mathbf{x}_1) = \dots = I(\mathbf{a}_{N_d}, \mathbf{x}_{N_d}) = 0 \\ 0 & \text{otherwise} \end{cases} \tag{5.8}$$
tells whether block (b, y) features the right correlations ($r = 1$) or the wrong ones ($r = 0$), in the sense of being compatible with the maximal violation of inequality (5.2). This function is computed for all blocks but the distilling one. The protocols is aborted unless all of them give the right correlations,
$$g = \prod_{j=1, j\neq l}^{N_b} r[b_j, y_j] = \begin{cases} 1 & \text{not abort} \\ 0 & \text{abort} \end{cases} . \tag{5.9}$$

Note that the abort/no-abort decision is independent of whether the distilling block l is right or wrong.

5. If the protocol is not aborted then k is assigned a bit generated from $b_l = (\mathbf{a}_1, \dots \mathbf{a}_{N_d})$ as
$$k = f(\mathrm{maj}(\mathbf{a}_1), \dots \mathrm{maj}(\mathbf{a}_{N_d})) \ . \tag{5.10}$$
Here $f : \{0,1\}^{N_d} \to \{0,1\}$ is a function whose existence is proven in Appendix A (along with other technical details), while $\mathrm{maj}(\mathbf{a}_i) \in \{0,1\}$ is the majority-vote among the three first bits of the quintuple string $\mathbf{a}_i$. If the protocol is aborted it sets $k = \varnothing$.

At the end of the protocol, k is potentially correlated with the settings of the distilling block $\tilde{y} = y_l$, the bit g in (5.9), and the bits
$$t = [l, (b_1, y_1), \dots (b_{l-1}, y_{l-1}), (b_{l+1}, y_{l+1}), \dots (b_{N_b}, y_{N_b})].$$
Additionally, an eavesdropper Eve might have a physical system correlated with k, which she may measure at any instance of the protocol. This system is not necessarily classical or quantum, the only assumption about it is that measuring it does not produce instantaneous signalling anywhere else. We label all possible measurements Eve can perform with the classical variable z, and with e the corresponding outcome. In summary, after performing the protocol all the relevant information is $k, \tilde{y}, t, g, e, z$, with statistics described by an unknown conditional probability distribution $P(k, \tilde{y}, t, g, e|z)$.

To assess the security of our protocol for full randomness amplification, we have to show that with a high probability the distribution describing the protocol when not aborted is indistinguishable from the distribution $P_{\mathrm{ideal}}(k, \tilde{y}, t, g, e|zg = 1) = \frac{1}{2} P(\tilde{y}, t, e|zg = 1)$ describing an ideal free random bit. For later purposes, it is convenient to cover the case when the protocol is aborted with an equivalent notation: if the protocol is aborted, we define $P(k, \tilde{y}, t, e|zg = 0) = \delta_k^{\varnothing} P(\tilde{y}, t, e|zg = 0)$ and $P_{\mathrm{ideal}}(k, \tilde{y}, t, e|zg = 0) = \delta_k^{\varnothing} P(\tilde{y}, t, e|zg = 0)$, where $\delta_k^{k'}$ is a Kronecker's delta. In this case, it is immediate that $P = P_{\mathrm{ideal}}$, as the locally generated symbol $\varnothing$ is always uncorrelated to the environment. To quantify the indistinguishability between P and P_{ideal}, we consider the scenario in which an observer, having access to all the information $k, \tilde{y}, t, g, e, z$, has to correctly distinguish between these two distributions. We denote by $P(\text{guess})$ the optimal probability of correctly guessing between the two distributions. This

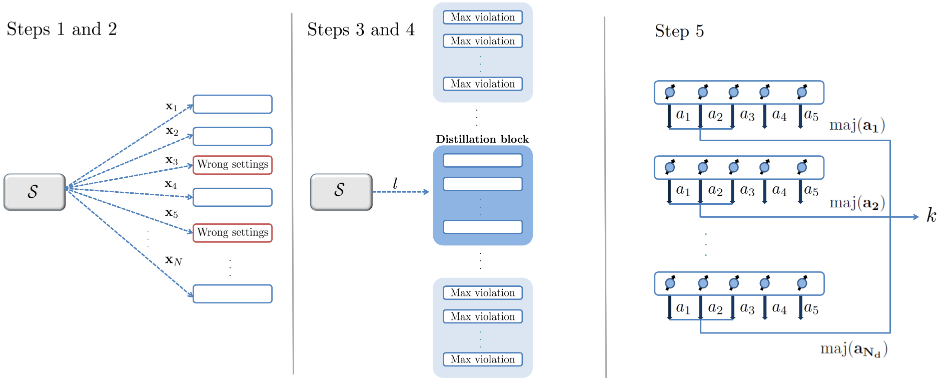

Figure 5.2: **Protocol for full randomness amplification based on quantum nonlocality**. In the first two steps, all N quintuplets measure their devices, where the choice of measurement is done using the ϵ-source $\mathcal{S}$; the quintuplets whose settings happen not to take place in the five-party Mermin inequality are discarded (in red). Note that although $\mathcal{S}$ is shown as a single source, it represents the collection of sources at the location of each observer correlated by (5.1). In steps 3 and 4, the remaining quintuplets are grouped into blocks. One of the blocks is chosen as the distillation block, using again $\mathcal{S}$, while the others are used to check the Bell violation. In the fifth step, the random bit k is extracted from the distillation block.

probability reads

$$P(\text{guess}) = \frac{1}{2} + \frac{1}{4} \sum_{k,\tilde{y},t,g} \max_{z} \sum_{e} \Big| P(k,\tilde{y},t,g,e|z) - P_{\text{ideal}}(k,\tilde{y},t,g,e|z) \Big|, \tag{5.11}$$

where the second term can be understood as (one fourth of) the variational distance between P and P_{ideal} generalized to the case when the distributions are conditioned on an input z [Mas09]. If the protocol is such that this guessing probability can be made arbitrarily close to 1/2, it generates a distribution P that is basically undistinguishable from the ideal one. This is known as "universally-composable security", and accounts for the strongest notion of cryptographic security (see [Can01] and [Mas09]). It implies that the protocol produces a random bit that is secure (free) in any context. In particular, it remains secure even if the adversary Eve has access to $\tilde{y}$, t and g. At this point, we can state the main result of our work.

Main result: *Given an ϵ-source with any $0 < \epsilon \leq 1/2$, a perfect free random bit k can be obtained using quantum non-local correlations.*

This result follows from the following theorem.

Theorem. *Consider the previous protocol for randomness amplification and the conditional probability distribution $P(k, \tilde{y}, t, g, e|z)$ describing the statistics of the bits $k, \tilde{y}, t, g$ generated during its execution and any possible system with input z and output e correlated to them. The probability $P(\text{guess})$ of correctly guessing between this distribution and the ideal distribution $P_{\text{ideal}}(k, \tilde{y}, t, g, e|z)$ is such that*

$$P(\text{guess}) \ \leq \ \frac{1}{2} + \frac{3\sqrt{N_d}}{2}\left[\alpha^{N_d} + 2\,N_b^{\log_2(1-\epsilon)}\left(32\beta\epsilon^{-5}\right)^{N_d}\right] \ . \qquad (5.12)$$

where α and β are real numbers such that $0 < \alpha < 1 < \beta$.

Proof. The proof of this Theoremis rather technical and and has been moved to Appendix A. □

Now, the right-hand side of (5.12) can be made arbitrary close to 1/2, for instance by setting $N_b = \left(32\,\beta\,\epsilon^{-5}\right)^{2N_d/|\log_2(1-\epsilon)|}$ and increasing N_d subject to fulfilling the condition $N_d N_b \geq N/3$. [Note that $\log_2(1-\epsilon) < 0$.] In the limit $P(\text{guess}) \to 1/2$, the bit k generated by the protocol is indistinguishable from an ideal free random bit.

To complete the argument, we must show that quantum resources can indeed successfully implement our protocol. It can be immediately verified that the qubit measurements X or Y on the quantum state $|\Psi\rangle = \frac{1}{\sqrt{2}}(|00000\rangle + |11111\rangle)$, with $|0\rangle$ and $|1\rangle$ the eigenstates of the Z qubit basis, yield correlations that maximally violate the five-partite Mermin inequality in question. This completes our main result.

Finally, we would like to conclude by explaining the main intuitions behind the proof of the previous theorem. As mentioned, the protocol builds on the 5-party Mermin inequality because it is the simplest GHZ paradox allowing some randomness certification. The estimation part, given by step 4, is rather standard and inspired by estimation techniques introduced in [BHK05], which were also used in [CR12b] in the context of randomness amplification. The most subtle part is the distillation of the final bit in step 5. Naively, and leaving aside estimation issues, one could argue that it is nothing but a classical processing by means of the function f of the imperfect random bits obtained via the N_d quintuplets. But this seems in contradiction with the result by Santha and Vazirani proving that it is impossible to extract by classical means a perfect free random bit from imperfect ones [SV86]. This intuition is however wrong. The reason is because in our protocol the randomness of the imperfect bits is certified by a Bell violation, which is impossible classically. Indeed, the Bell certification allows applying techniques similar to those obtained in Ref. [Mas09] in the context of privacy amplification against non-signalling eavesdroppers. There, it was shown how to amplify the privacy, that is the unpredictability, of one of the measurement outcomes of bipartite correlations violating a Bell inequality. The key point is that the amplification, or distillation, is attained in a *deterministic* manner. That is, contrary to standard approaches, the privacy amplification process described in [Mas09] does not consume any randomness. Clearly, these deterministic techniques are extremely convenient for our randomness amplification scenario. In fact, the distillation part in our protocol can be seen as the translation of the privacy amplification techniques of Ref. [Mas09] to our more complex scenario, involving now 5-party non-local correlations and a function of three of the measurement outcomes.

5.2 Discussion

In summary, we have presented a protocol that, using quantum non-local resources, attains *full randomness amplification*. This task is impossible classically and was not known to be possible in the quantum regime. As our goal was to prove full randomness amplification, our analysis focuses on the noise-free case. In fact, the noisy case only makes sense if one does not aim at perfect random bits and bounds the amount of randomness in the final bit. Then, it should be possible to adapt our protocol in order to get a bound on the noise it tolerates. Other open questions that naturally follow from our results consist of studying randomness amplification against quantum eavesdroppers, or the search of protocols in the bipartite scenario.

From a more fundamental perspective, our results imply that there exist experiments whose outcomes are fully unpredictable. The only two assumptions for this conclusion are the existence of events with an arbitrarily small but non-zero amount of randomness and the validity of the no-signalling principle. Dropping the former implies accepting a super-deterministic view where no randomness exist, so that we experience a fully pre-determined reality. This alternative challenges both our usual scientific [Gis10] and philosophical presumptions. Dropping the latter, in turn, implies abandoning a local causal structure for events in space-time. However, this is one of the most fundamental notions of special relativity, and without which even the very meaning of randomness or predictability would be unclear, as these concepts implicitly rely on the cause-effect principle.

Chapter 6

The intrinsic content of observed randomness

Randomness comes in two flavours: Classical randomness associated with subjective lack of knowledge and the objective or intrinsic randomness. In quantum theory, randomness from pure states is fully intrinsic while those from mixed states necessarily encode classical randomness associated with our lack of knowledge of the preparation. The analogues in the non-signalling set for pure quantum states are the extremal boxes while for the mixed quantum states, the non-extremal ones. Exploiting this analogy further, we expect that a given general non-signalling distribution contains *both* classical and intrinsic randomness unless the distribution happens to be extremal in the non-signalling set.

Fig. 6.1 is a useful qualitative geometric picture which serves to clarify the general idea and to explain the scenario we work with. Given some non-local distribution P_{obs}, its intrinsic randomness content is quantitatively dependent on whether we use the quantum or non-signalling framework. For example, the Tsirelson correlations [Tsi87] in the $(2, 2, 2)$ scenario considered strictly within the quantum set yields 1.23 bits of randomness [AMP12]. However, its randomness in the larger non-signalling set is a much smaller 0.34 bits[1]. Another example is the GHZ correlations [GHZ89] which contain (considering the tripartite states in particular) 3 bits of randomness within the quantum set. However, in the non-signalling set it is just 1 bit[2] as can be seen from the extremal boxes characterized in [PBS11]. In fact, it is generally the case that the intrinsic randomness of a point considered to be

[1] This can be computed easily from Fig. 2.3

[2] This can be computed from Definition 7 in Chapter 2.

embedded in the non-signalling set is lower than its intrinsic randomness within the quantum set. The reason is simply that there are more general decompositions possible within the non-signalling set which increases our ignorance about its underlying preparation. It is in this context that we now pose the question that is the theme of this chapter: is it possible to certify that observed randomness is fully intrinsic for some observed quantum correlations P_{obs}?

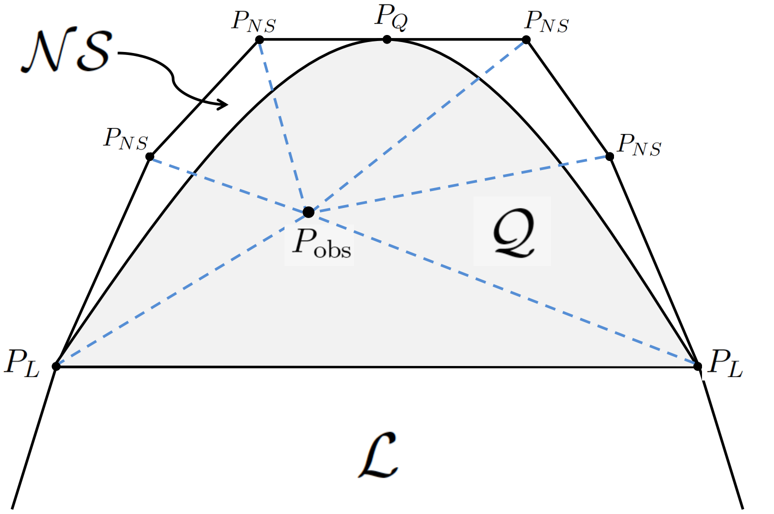

Figure 6.1: Qualitative picture of the local, quantum and no-signalling sets and a possible preparation of a given distribution P_{obs} as a mixture of extremal non-signalling points. Knowledge of the exact preparation of P_{obs} is required to compute its intrinsic randomness, which is thus expected to be lower than its observed randomness (calculated from the observed statistics without regard for the underlying preparation).

The challenge to answering this question in full generality is that the definition of intrinsic randomness in these scenarios (Dfn. 7, Chapter 2) required an optimization over all possible non-signalling preparations of P_{obs}. This is believed to be a hard problem[3]. What we show here is that despite the apparent complexity of calculating the intrinsic randomness in full generality it is possible to choose scenarios carefully in which an *analytic* computation is feasible. We not only demonstrate one such case but also certify that the

[3] For $(2,2,2)$ there are 8 extremal boxes in two inequivalent classes while in $(3,2,2)$ there are 53,856 extremal boxes forming forty six inequivalent classes. A full classification for higher scenarios is non-existent to date.

observed randomness is fully intrinsic in our chosen scenario. What make the result counter-intuitive is that our results are valid for a whole class of non-extremal distributions.

There is a further layer of subtlety which we additionally address in our work. This is related to the particular concern about using the measurement independence assumption in the context of randomness certification using Bell inequalities. This is the assumption of using randomness for the choice of inputs in order to certify randomness of the outcomes. Recently there has been a significant body of work in deriving Bell inequalities with relaxations of this assumption [KPB06, BG10, Gis10, Hal11, GMdlT$^+$12]. A significant feature of our results are that they are valid even under a complete (non-zero) relaxation of the measurement assumption since the conceptual machinery from the previous chapter applies here. For this reason, these results may also be interpreted as an alternative approach for full randomness amplification with the benefit of significantly easier techniques.

6.1 Results

In general, computing the intrinsic randomness of non-extremal distributions belonging to the non-signalling set is infeasible for the lack of characterization of the extremal boxes in all but the simplest cases. However, we demonstrate a criterion that allows the computation of the intrinsic randomness for the class of non-signalling boxes maximally violating the Mermin inequality (See Eqn. 6.4). We do this by choosing a function of the outcomes. Furthermore, for the GHZ correlations in particular, these functions provide a fully random bit asymptotically in the system size. We express our results concisely in the following theorems.

Theorem 14. *Let $P(\mathbf{a}|\mathbf{x})$ be a N-partite (odd N) non-signalling probability distribution maximally violating the corresponding Mermin inequality and $f : \{0,1\}^N \mapsto \{0,1\}$ a boolean function defined as,*

$$f(\mathbf{a}) = \begin{cases} 0 & \sum_{i=1}^{N} a_i = (4j+2); \quad j = 0,1,2,\ldots \\ \\ 1 & \textit{otherwise} \end{cases} \tag{6.1}$$

where the outcomes $a_i \in \{0,1\}$. Then $G_{\text{int}}(f, \mathbf{x}, P) = G_{\text{obs}}(f, \mathbf{x}, P)$ for any input $\mathbf{x}$ appearing in the Mermin inequality.

From this, one can derive the following result for the important family of quantum probability distributions called the GHZ correlations:

Lemma 15. *For the N-partite (odd N) quantum probability probability distribution $P_{\mathrm{ghz}}(\mathbf{a}|\mathbf{x})$, the intrinsic randomness of f for a Mermin input is,*

$$G_{\mathrm{int/obs}}(f, \mathbf{x}, P_{\mathrm{ghz}}) = \frac{1}{2} + \frac{1}{2^{(N+1)/2}}$$

This means that, in this particular scenario, additional knowledge of the particular preparation of P_{ghz} (as a mixture of non-signalling boxes) does not help to predict the the output bit $k = f(\mathbf{a})$ better. Furthermore, output bit k approaches a perfect random bit exponentially fast in the system size.

We prove these results explicitly for $N = 3$ in this Chapter and generalize it for arbitrary N in Appendix B.

6.2 Methods

Although we have already defined our measures of randomness in Chapter 2, it is useful to recall them once again particularly since they are adapted to calculating the randomness of a function of the outcomes. To this end, consider a function $g : \{0,1\}^N \mapsto \{0,1\}^R$ mapping a string of N bits $\mathbf{a} = (a_1, \ldots, a_N)$ to an output of R bits $g(\mathbf{a})$ where $1 \leq R \leq N$. Given an N-partite probability distribution $P_{\mathrm{obs}}(\mathbf{a}|\mathbf{x})$ we define the following two quantities:

Definition. *(*Observed randomness*) Given the correlations $P_{\mathrm{obs}}(\mathbf{a}|\mathbf{x})$, the observed randomness of the function $g(\mathbf{a})$ for a given input $\mathbf{x}$ is quantified by the maximum probability of guessing its outcome:*

$$G_{\mathrm{obs}}(g, \mathbf{x}, P_{\mathrm{obs}}) = \max_{k \in \mathrm{Im}(g)} P_{\mathrm{obs}}(g(\mathbf{a}) = k|\mathbf{x}). \tag{6.2}$$

As before, the randomness is given by the min-entropy defined as, $H_{\infty}^{\mathrm{obs}}(g, \mathbf{x}, P_{\mathrm{obs}}) = -\log_2 G_{\mathrm{obs}}(g, \mathbf{x}, P_{\mathrm{obs}})$.

This is a quantity that can be easily computed from the observed statistics.

Definition. *(*Intrinsic randomness*) Given the correlations $P_{\mathrm{obs}}(\mathbf{a}|\mathbf{x})$, the intrinsic randomness of the function $g(\mathbf{a})$ for a given input $\mathbf{x}$ is quantified by the maximum probability of guessing its outcome optimized over all possible* *non-signalling extremal decompositions of $P_{\text{obs}}(\mathbf{a}|\mathbf{x})$:*

$$G_{\text{int}}(g, \mathbf{x}, P_{\text{obs}}) = \max_{\{p_j, P_j^{\text{ex}}\}} \sum_j p_j G_{\text{obs}}(g, \mathbf{x}, P_j^{\text{ex}})$$

$$\textit{subject to:}$$

$$\sum_j p_j P_j^{\text{ex}}(\mathbf{a}|\mathbf{x}) = P_{\text{obs}}(\mathbf{a}|\mathbf{x}), \tag{6.3}$$

where $G_{\text{obs}}(g, \mathbf{x}, P_j^{\text{ex}}) = \max_{k \in \text{Im}(g)} P_j^{\text{ex}}(g(\mathbf{a}) = k|\mathbf{x})$ *is also the intrinsic randomness of* P_j^{ex} *since it is an extremal point of the non-signalling set.* H_∞^{int} *is defined analogous to the above.*

This definition of intrinsic randomness manifestly depends on the knowledge of the extremal boxes and is operationally relevant to cryptographic scenarios involving a non-signalling eavesdropper potentially possessing such knowledge. It quantifies the randomness perceived by the eavesdropper.

We now make a simple observation that not only allows calculating the intrinsic randomness without reference to the extremal boxes but also guarantees that the intrinsic randomness thus calculated is exactly equal to the observed randomness. We work with the class of distributions $P(\mathbf{a}|\mathbf{x})$ maximally violating a Mermin inequality,

$$M_N = \frac{1}{2} M_{N-1}(A_N + A'_N) + \frac{1}{2} M'_{N-1}(A_N - A'_N), \tag{6.4}$$

where M_2 is the CHSH inequality and M'_{N-1} denotes M_{N-1} with all operators locally swapped.

Let us assume without immediate justification that there exists a boolean function $f : \{0,1\}^N \mapsto \{0,1\}$ satisfying,

$$P(f(\mathbf{a}) = 0|\mathbf{x}) \geq 1/2 \tag{6.5}$$

for the entire class of distributions maximally violating the Mermin inequality.

It is easy to see that the intrinsic randomness of f for some distribution $P_{\text{obs}}(\mathbf{a}|\mathbf{x})$ in this class is equal to its observed randomness. To this end, we first make the observation that all P_j^{ex} in any valid decomposition of $P_{\text{obs}}(\mathbf{a}|\mathbf{x})$ must also necessarily lead to a maximal violation of the Mermin inequality[4]. Hence, they satisfy property (6.5). We use this to first compute

[4] If not, the Mermin violation of the decomposition is incompatible with the maximal violation of P_{obs}

the randomness in any extremal point appearing in Eqn. 6.3,

$$\begin{aligned} G_{\text{obs}}(f,\mathbf{x},P_j^{\text{ex}}) &= \max\{P_j^{\text{ex}}(f(\mathbf{a})=0|\mathbf{x}),P_j^{\text{ex}}(f(\mathbf{a})=1|\mathbf{x})\} \\ &= |P_j^{\text{ex}}(f(\mathbf{a})=0|\mathbf{x})-1/2|+1/2 \\ &= P_j^{\text{ex}}(f(\mathbf{a})=0|\mathbf{x}), \end{aligned} \tag{6.6}$$

for every j. We use this expression to finally compute the intrinsic randomness of the given distribution P_{obs},

$$\begin{aligned} G_{\text{int}}(f,\mathbf{x},P_{\text{obs}}) &= \max_{\{p_j,P_j^{\text{ex}}\}}\sum_j p_j G_{\text{obs}}(f,\mathbf{x},P_j^{\text{ex}}) \\ &= \max_{\{p_j,P_j^{\text{obs}}\}}\sum_j p_j P_j^{\text{ex}}(f(\mathbf{a})=0|\mathbf{x}) \\ &= P_{\text{obs}}(f(\mathbf{a})=0|\mathbf{x}), \end{aligned} \tag{6.7}$$

where the last equality follows from the constraint $\sum_j p_j P_j(\mathbf{a}|\mathbf{x}) = P_{\text{obs}}(\mathbf{a}|\mathbf{x})$. On the other hand the observed randomness given by Eqn. (6.2) for f gives us, $G_{\text{obs}}(f,\mathbf{x},P_{\text{obs}}) = P_{\text{obs}}(f(\mathbf{a})=0|\mathbf{x})$. Thus, the intrinsic and observed randomness are identical for any P maximally violating the Mermin inequality for any function that come equipped with property (6.5).

6.2.1 A function satisfying the required property

The task is now to demonstrate that the function f defined in Eqn. 6.1 satisfies the property 6.5 for all N. Since the general proof involves many details, we prove it for $N=3$ below and relegate the generalization to Appendix B.

Denoting the measurement choices of each party with the labels $\{0,1\}$ for convenience, the tri-partite Mermin inequality may be expressed as,

$$M_3 = \langle 001\rangle + \langle 010\rangle + \langle 100\rangle - \langle 111\rangle \leq 2. \tag{6.8}$$

The quantum and non-signalling maximal violation assigns $M_3 = 4$ which can clearly only occur when the first three correlators in 6.8 take their maximum value of $+1$ and the last takes its minimum of -1. Let $P(\mathbf{a}|\mathbf{x})$ be a tripartite no-signalling distribution maximally violating M_3. We extract randomness from one of the inputs combinations appearing in the inequality. To be concrete, we choose this input to be $\mathbf{x_m} = (0,0,1)$. Notice that the correlator corresponding to this input appears in Eqn. (6.8). For tripartite systems, the definition of the function (6.1) implies that $f(\mathbf{a}) = 0$ if

$\sum_i a_i = 2$. In order that the machinery of and conclusions following Eqn. (6.7) hold, we would like to prove that f has the property,

$$\begin{aligned} & P(f(\mathbf{a}) = 0|\mathbf{x_m}) \\ & = P(0,1,1|\mathbf{x_m}) + P(1,0,1|\mathbf{x_m}) + P(1,1,0|\mathbf{x_m}) \\ & \geq 1/2 \end{aligned} \tag{6.9}$$

for any P satisfying $M_3(P) = 4$ *i.e.* the maximal violation of the inequality.

Maximal violation of the inequality imposes several conditions on the correlators and the marginals. Since, we use these extensively it is useful to state them explicitly for this case. Maximal violation imposes the chain of relationships:

1. $\langle 001\rangle = 1$. This further implies $\langle 0\rangle_x = \langle 01\rangle_{yz}$, $\langle 0\rangle_y = \langle 01\rangle_{xz}$ and $\langle 1\rangle_z = \langle 00\rangle_{xy}$.
2. $\langle 010\rangle = 1$ implying $\langle 0\rangle_x = \langle 10\rangle_{yz}$, $\langle 1\rangle_y = \langle 00\rangle_{xz}$ and $\langle 0\rangle_z = \langle 01\rangle_{xy}$.
3. $\langle 100\rangle = 1$ implying $\langle 1\rangle_x = \langle 00\rangle_{yz}$, $\langle 0\rangle_y = \langle 10\rangle_{xz}$ and $\langle 0\rangle_z = \langle 10\rangle_{xy}$.
4. $\langle 111\rangle = -1$ implying $\langle 1\rangle_x = -\langle 11\rangle_{yz}$, $\langle 1\rangle_y = -\langle 11\rangle_{xz}$ and $\langle 1\rangle_z = -\langle 11\rangle_{xy}$

Using these relations in the usual parametrization of P for input $\mathbf{x_m} = (0,0,1)$ gives,

$$\begin{aligned} & P(a,b,c|0,0,1) \\ =& \frac{1}{8}(1 + a\langle 0\rangle_x + b\langle 0\rangle_y + c\langle 1\rangle_z + ab\langle 00\rangle_{xy} \\ & + ac\langle 01\rangle_{xz} + bc\langle 01\rangle_{yz} + abc\langle 001\rangle_{xyz}) \\ =& \frac{1}{8}(1 + abc + (a+bc)\langle 0\rangle_x + (b+ac)\langle 0\rangle_y \\ & + (c+ab)\langle 1\rangle_z) \end{aligned} \tag{6.10}$$

We are finally ready to expand Eqn. (6.9) using the relations above. Taking care to change the outcome labelling from $\{0,1\} \mapsto \{+1,-1\}$ in order to correctly use the parametrization Eqn. (6.10) we may write,

$$\begin{aligned} & P(f(\mathbf{a}) = 0|\mathbf{x_m}) \\ & = P(1,-1,-1|\mathbf{x_m}) + P(-1,1,-1|\mathbf{x_m}) \\ & + P(-1,-1,1|\mathbf{x_m}) \\ & = \frac{1}{4}(3 - \langle 0\rangle_x - \langle 0\rangle_y - \langle 1\rangle_z) \end{aligned} \tag{6.11}$$

Hence, proving $P(f(\mathbf{a})=0|\mathbf{x_m})\geq 1/2$ is equivalent to proving $1-\langle 0\rangle_x-\langle 0\rangle_y-\langle 1\rangle_z\geq 0$.

This form is very convenient since it immediately reminds one of a positivity condition of probabilities. Our next task is precisely to show that summing appropriate positivity conditions gives us the result.

6.2.2 Positivity conditions from the swapped input

Given that $\mathbf{x_m}=(0,0,1)$ is the input from which we extract randomness, we denote by $\bar{\mathbf{x}}_\mathbf{m}$ the input combination which is a local swapping by each party of their setting in $\mathbf{x_m}$. We call this the swapped input. Thus, in this case $\bar{\mathbf{x}}_\mathbf{m}=(1,1,0)$. Note that this is *not* an input appearing in the Mermin inequality. We can again write the correlator parametrization as before,

$$\begin{aligned}
&P(a,b,c|1,1,0)\\
=&\frac{1}{8}(1+a\langle 1\rangle_x+b\langle 1\rangle_y+c\langle 0\rangle_z+ab\langle 11\rangle_{xy}\\
&+ac\langle 10\rangle_{xz}+bc\langle 10\rangle_{yz}+abc\langle 110\rangle_{xyz})\\
=&\frac{1}{8}(1+a\langle 1\rangle_x+b\langle 1\rangle_y+c\langle 0\rangle_z-ab\langle 1\rangle_z\\
&+ac\langle 0\rangle_y+bc\langle 0\rangle_x+abc\langle 110\rangle_{xyz}),
\end{aligned}\tag{6.12}$$

where the second equality results from the relations $\langle 11\rangle_{xy}=-\langle 1\rangle_z$, $\langle 10\rangle_{xz}=\langle 0\rangle_y$ and $\langle 10\rangle_{yz}=\langle 0\rangle_x$.

It can be easily verified that summing the two positivity conditions $P(1,1,-1|\bar{\mathbf{x}}_\mathbf{m})\geq 0$ and $P(-1,-1,1|\bar{\mathbf{x}}_\mathbf{m})\geq 0$ gives us the result we seek, namely, $1-\langle 0\rangle_x-\langle 0\rangle_y-\langle 1\rangle_z\geq 0$.

This proves that f satisfies $P(f(\mathbf{a})=0|\mathbf{x})\geq 1/2$ for all tripartite distributions maximally violating the Mermin inequality. This guarantees that the observed randomness of f has no classical component, namely that $H_\infty^{class}(f,\mathbf{x},P_{\text{obs}})=0$.

6.3 Discussion

We have seen that for the choice of our function, the observed randomness in distributions maximally violating the Mermin inequality is wholly intrinsic. This includes the physically realizable GHZ correlations. For the latter, the randomness of the function approaches that of a perfect bit exponentially fast in the size of the system. In adversarial terms, this implies that no non-signalling adversary has additional knowledge or can predict the outcome of f better than the parties performing the Bell test.

In the context of the GHZ correlations (being the only correlations in the class we have defined that may be attained by quantum systems), our result bears a resemblance to those in [CR11, CR12a] where the completeness of quantum theory was discussed. These results attempt to show that the predictive power of quantum theory is maximal. However, our scenario departs significantly from the one considered there. For one thing, our results concern a specific class of non-signalling distributions which contains only one quantum point. Besides, we consider a function of the outcomes. Most important of all, our setup allows us to relax the critical measurement independence assumption arbitrarily, as long as it is non-zero. This was not possible in [CR11, CR12a], except perhaps in a limited sense due to the partial randomness amplification results of [CR12b].

Furthermore, our results bear a deep relationship with full randomness amplification [GMdlT$^+$12]. Since the measurement independence can be relaxed and we find our function approaching a perfect random bit with increasing system size for GHZ correlations, this is precisely the task set out to full randomness amplification. The missing link is a protocol to estimate the obtained statistics. Such a protocol is a departure from the theme we pursue here but we point out that one does not in principle require the (technically complex) distillation step used in Chapter 5. Thus, our techniques here may potentially lead to a much simplified randomness amplification procedure.

Our techniques can also be extended in some natural directions. For instance, for a chosen value of N, suppose there exists some function l satisfying $P(l = 0|\mathbf{x}_m) \leq 1/2 + \zeta$ for all P maximally violating the Mermin inequality. Using logic similar to Eqn. 6.6 and the triangle inequality, we find that the intrinsic randomness of l for the GHZ correlations is, $G_{\text{int}}(l, \mathbf{x_m}, P_{ghz}) \leq P_{ghz}(l = 1|\mathbf{x}_m) + 2\zeta_N$. In other words, intrinsic randomness is generally smaller than the observed randomness in such cases.

Future directions of work include exploiting such relations to upper bound the classical randomness where exact relations are not possible. Moreover, an interesting line of work is to extend these techniques for distributions non-maximally violating Bell inequalities. These could lead to experimentally viable tests of device independence randomness certification [PAM$^+$10b].

Chapter 7

Noisy entanglement dynamics in graph states

In the preceding chapters we studied the properties of intrinsic randomness in the different classes of non-local correlations. However, we have not yet discussed the resources required for the generation of such randomness in a physical set-up. In general, behaviour such as spontaneous emission and nuclear decay are believed objectively random. However, there is no known way to *certify* this to be true. It is simply axiomatic in operational quantum theory. Non-locality is the only known certificate of objective randomness and occurs in multipartite systems by definition. In fact, in physical terms, it is only entanglement that can general non-local correlations which in turn is necessary to generate certifiable objective randomness. For this reason, we focus our attention in this last chapter on the resource of entanglement. We do this for a class of special quantum states known as graph states, which are fully multipartite entangled as already pointed out in Chapter 2. In particular this class includes the GHZ states which are used extensively in this work for randomness generation and amplification. They have also been identified as critical resources for several other tasks. We discuss these below in the context of graph states.

Graph states [HEB04, HDE$^+$06] constitute an important family of genuine multiparticle-entangled states with several applications in quantum information. The most popular example of these are arguably the cluster states, which have been identified as a crucial resource for universal measurement-based quantum computation [BBD$^+$09, RB01, RBB03]. Other members of this family were also proven to be potential resources, as codewords for quantum error correction [SW01], to implement secure quantum communication [DCB05, CL07], and to simulate some aspects of the entanglement distribution of random states [DP06a]. Moreover, graph states encompass the GHZ states [GHZ89], whose importance spans several fundamental and applied issues beyond those discussed in the previous chapters. GHZ states can – for large-dimensional systems – be considered as simple models of the *gedanken* Schrödinger-cat states, are crucial for quantum communication protocols [BVK98, HBB99, DP06b], and find applications in quantum metrology [GLM04] and high-precision spectroscopy [BIWH96]. All these reasons explain the great deal of effort made both to theoretically understand the features [HEB04, HDE+06], and to generate and coherently manipulate graph states in the laboratory [WRR+05, KSW+05, LZG+07, CLZ+07, VPDMM08].

For the same reasons, it is crucial to unravel the dynamics of graph states in realistic scenarios, where the system is unavoidably exposed to interactions with its environment and/or experimental imperfections. Previous studies on the robustness of graph-state entanglement in the presence of decoherence showed that the disentanglement times (*i.e.* the time for which the state becomes separable) increases with the system size [SK02, DB04, HDB05]. However the disentanglement time on its own is known not to provide in general a faithful figure of merit for the entanglement robustness: although the disentanglement time can grow with the number N of particles, the amount of entanglement in a given time can decay exponentially with N [ACC+08, ACA+09]. The full dynamical evolution must then be monitored to draw any conclusions on the entanglement robustness.

A big obstacle must be overcome in the study of entanglement robustness in general mixed states: the direct quantification of the entanglement involves optimizations requiring computational resources that increase exponentially with N. The problem thus becomes in practice intractable even for relatively small system sizes, not to mention the direct assessment of entanglement during the entire noisy dynamics. All in all, some progress has been achieved in the latter direction for some very particular cases: For arbitrarily-large linear-cluster states under collective dephasing, it is possible to calculate the exact value of the geometric measure of entanglement [WG03] throughout the evolution [GBB08]. Besides, bounds to the relative entropy and the global robustness of entanglement for two-colorable graph states [HEB04, HDE+06] of any size under local dephasing were obtained [HW10].

In a conceptually different approach, a framework to obtain families of lower and upper bounds to the entanglement evolution of graph, and graph-diagonal, states under decoherence was introduced in Ref. [CCA+09]. The bounds are obtained via a calculation that involves only the *boundary subsystem*, composed of the qubits lying at the boundary of the multipartition under scrutiny. This, very often, considerably reduces the size of the matrices involved in the calculation of entanglement. No optimization on the full system's parameter space is required throughout. Another remarkable feature of the method is that it is not limited to a particular entanglement quantifier but applies to all convex (bi-or multi-partite) entanglement measures that do not increase under local operations and classical communication (LOCC). The latter are indeed two rather natural and general requirements [PV07, HHHH09].

In the case of open-dynamic processes described by Pauli maps, which are defined below, the lower and upper bounds coincide and the method thus allows one to calculate the exact entanglement of the noisy evolving state. Pauli maps encompass popular models of (independent or collective) noise, as depolarization, phase flip, bit flip and bit-phase flip errors, and are defined below. Moreover, one of the varieties of lower bounds is of extremely simple calculation and – despite less tight – depends only on the connectivity of the graph and not on its total size. The latter is a very advantageous property in situations where one wishes to assess the resistance of entanglement with growing system size. For example, the versatility of the formalism has very recently been demonstrated in Ref. [DCA10], where it was applied to demonstrate the robustness of thermal bound entanglement in macroscopic many-body systems of spin-1/2 particles.

7.1 Results

We elaborate on the details of the formalism introduced in [CCA^{+}09] and partially studied in [Dha09]. For Pauli maps we give an explicit formula for the characterization of the effective noise involved in the calculation of the bounds. We use this formula to compute the dynamics of entanglement of a 12 and 14 qubit cluster state under individual depolarizing noise (see Fig. 7.3). Furthermore, we extend the method to the case where each qubit is subjected to the action of independent thermal baths of arbitrary temperature. This is a crucial, realistic type of dynamic process that is not described by Pauli maps. In all cases, we exhaustively compare the different bounds with several concrete examples. Finally, we discuss the main advantages and limitations of our method in comparison with other approaches.

7.2 Basic concepts

In this section, we define graph and graph-diagonal states, introduce the basics of open-system dynamics and the particular noise models used later.

7.2.1 Graph and graph-diagonal states

Graph states are multiqubit quantum states defined from mathematical graphs through the rule described below. First, a mathematical graph $G_{(\mathcal{V},\mathcal{C})} \equiv \{\mathcal{V},\mathcal{C}\}$ is defined by a set $\mathcal{V}$ of N vertices, or nodes, and a set $\mathcal{C}$, of connections, or edges, connecting each node i to some other j. An example of such graph is illustrated in Fig. 7.1. Each vertex $i \in \mathcal{V}$ represents a qubit in the associated physical system, and each edge $\{i,j\} \in \mathcal{C}$ represents a unitary maximally-entangling controlled-Z (CZ) gate, $CZ_{ij} = |0_i 0_j\rangle\langle 0_i 0_j| + |0_i 1_j\rangle\langle 0_i 1_j| + |1_i 0_j\rangle\langle 1_i 0_j| - |1_i 1_j\rangle\langle 1_i 1_j|$, between the qubits i and j connected through the corresponding edge. The N-qubit graph state $|G_{(\mathcal{V},\mathcal{C})_0}\rangle$ corresponding to graph $G_{(\mathcal{V},\mathcal{C})}$ is then operationally defined as follows:

(i) Initialize every qubit i in the superposition $|+_i\rangle = \frac{1}{\sqrt{2}}(|0_i\rangle + |1_i\rangle)$, so that the joint state is in the product state $|g_{(\mathcal{V})_0}\rangle \equiv \bigotimes_{i\in\mathcal{V}} |+_i\rangle$.

(ii) Then, for every connection $\{i,j\} \in \mathcal{C}$ apply the gate CZ_{ij} to $|g_{(\mathcal{V})_0}\rangle$. That is,

$$|G_{(\mathcal{V},\mathcal{C})_0}\rangle = \bigotimes_{\{i,j\}\in\mathcal{C}} CZ_{ij} |g_{(\mathcal{V})_0}\rangle. \tag{7.1}$$

Graph state (7.1) can also be defined in an alternative, non-operational fashion. Associated to each node $i \in \mathcal{V}$ of a given graph $G_{(\mathcal{V},\mathcal{C})}$ we define the operator

$$S_i \equiv X_i \bigotimes_{j\in\mathcal{N}_i} Z_j, \tag{7.2}$$

with X_i and Z_j the usual Pauli operators acting respectively on qubits i and j, and where $\mathcal{N}_i$ denotes the set of neighbours of i, directly connected to it by an edge $\{i,j\}$. Operator (7.2) possess eigenvalues 1 and -1. It is the i-th generator of the stabilizer group and is often called for short *stabilizer operator*. All N stabilizer operators commute and share therefore a common basis of eigenstates. Graph state $|G_{(\mathcal{V},\mathcal{C})_0}\rangle$ in turn has the peculiarity of being the unique common eigenstate of eigenvalue $+1$ [HEB04, HDE$^+$06]. In other words,

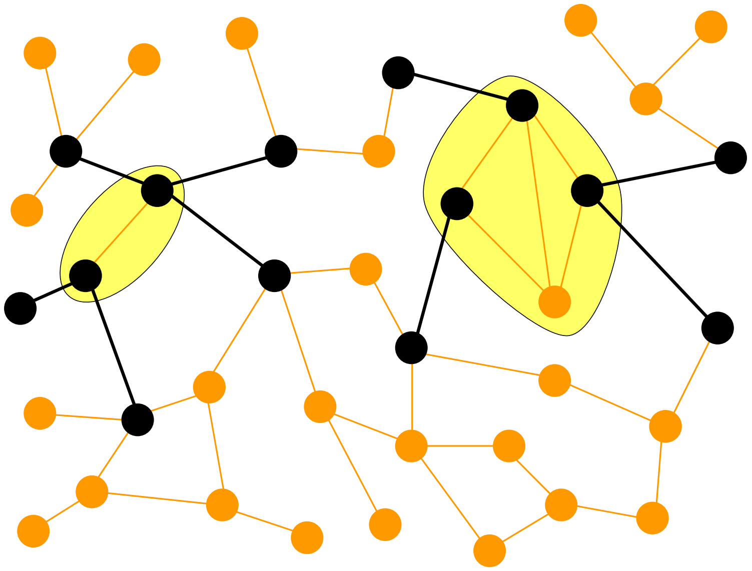

Figure 7.1: Mathematical graph associated to a given physical graph state. An exemplary bipartition divides the system into two subparts: the yellow and white regions. The edges in black are the *boundary-crossing edges* $\mathcal{X}$ and the nodes (also in black) connected by these are the *boundary nodes* $\mathcal{Y}$. Together they compose the *boundary sub-graph* $G(\mathcal{Y}, \mathcal{X})$. The remaining vertices, painted in orange (grey), constitute the *non-boundary subsystem.*

$$S_i|G_{(\mathcal{V},\mathcal{C})_0}\rangle = |G_{(\mathcal{V},\mathcal{C})_0}\rangle \;\forall\; i \in \mathcal{V}.$$

The other $2^N - 1$ common eigenstates $|G_{(\mathcal{V},\mathcal{C})_\nu}\rangle$ are in turn related to (7.1) by a local unitary operation:

$$|G_{(\mathcal{V},\mathcal{C})_\nu}\rangle = \bigotimes_{i\in\mathcal{V}} {Z_i}^{\nu_i}|G_{(\mathcal{V},\mathcal{C})_0}\rangle = Z^\nu|G_{(\mathcal{V},\mathcal{C})_0}\rangle, \tag{7.3}$$

such that $S_i\left|G_{(\mathcal{V},\mathcal{C})\nu}\right\rangle$=$(-1)^{\nu_i}\left|G_{(\mathcal{V},\mathcal{C})\nu}\right\rangle$, where ν is a multi-index representing the binary string $\nu \equiv \nu_1 \ldots \nu_N$, with $\nu_i = 0$ or $1 \;\; \forall\; i \in \mathcal{V}$, and where the short-hand notation $Z^\nu \equiv \bigotimes_{i\in\mathcal{V}} {Z_i}^{\nu_i}$ has been introduced. Therefore, states (7.3) possess all exactly the same entanglement properties and, together with $|G_{(\mathcal{V},\mathcal{C})_0}\rangle$, define the complete orthonormal graph-state basis of $\mathcal{H}$ (corresponding to the graph $G_{(\mathcal{V},\mathcal{C})}$). Any state ρ diagonal in such basis is called a *graph-diagonal state*:

$$\rho_{GD} = \sum_{\nu} P_{\nu} |G_{(\mathcal{V},\mathcal{C})_{\nu}}\rangle\langle G_{(\mathcal{V},\mathcal{C})_{\nu}}|, \tag{7.4}$$

where p_ν is any probability distribution. Interestingly, for any graph, any arbitrary N-qubit state can always be depolarized by some separable map (defined below) into the form (7.4) without changing its diagonal elements in the considered graph basis [DAB03, ADB05].

Two simple identities following from definition (7.2) will be crucial for our purposes. For every eigenstate $|G_{(\mathcal{V},\mathcal{C})_{\nu}}\rangle$ of S_i, with eigenvalues $s_{i\nu} = 1$ or -1

$$\begin{aligned} X_i|G_{(\mathcal{V},\mathcal{C})_{\nu}}\rangle &= S_i \otimes \bigotimes_{j\in\mathcal{N}_i} Z_j |G_{(\mathcal{V},\mathcal{C})_{\nu}}\rangle \\ &= s_{i\nu} \bigotimes_{j\in\mathcal{N}_i} Z_j |G_{(\mathcal{V},\mathcal{C})_{\nu}}\rangle, \end{aligned} \tag{7.5}$$

where definition (7.2) was used, and

$$\begin{aligned} Y_i|G_{(\mathcal{V},\mathcal{C})_{\nu}}\rangle &= (-i) Z_i . S_i \otimes \bigotimes_{j\in\mathcal{N}_i} Z_j |G_{(\mathcal{V},\mathcal{C})_{\nu}}\rangle \\ &= s_{i\nu} (-i) Z_i \otimes \bigotimes_{j\in\mathcal{N}_i} Z_j |G_{(\mathcal{V},\mathcal{C})_{\nu}}\rangle. \end{aligned} \tag{7.6}$$

So, when applied to any pure graph – or mixed graph-diagonal – state, the following operator equivalences hold up to a global phase:

$$X_i \leftrightarrow \bigotimes_{j\in\mathcal{N}_i} Z_j, \tag{7.7a}$$

$$Y_i \leftrightarrow Z_i \otimes \bigotimes_{j\in\mathcal{N}_i} Z_j. \tag{7.7b}$$

7.2.2 Open-system dynamics

As mentioned before, our ultimate goal is to study the behaviour of graph state entanglement in realistic dynamic scenarios where the system evolves during a time interval t according to a generic physical process, which can include decoherence. This process can always be represented by a completely-positive trace-preserving map Λ, that maps any initial state ρ to the evolved one after a time t, $\rho_t \equiv \Lambda(\rho)$. In turn, for every such Λ, there always exists a maximum of D^2 ($D = \dim(\mathcal{H})$) operators K_μ such that the map is expressed in a Kraus form [NC00].

$$\rho_t \equiv \Lambda(\rho) = \sum_\mu K_\mu \rho K_\mu^\dagger. \tag{7.8}$$

Operators K_μ are called the Kraus operators, and decompose the identity operator $\mathbf{1}$ of $\mathcal{H}$ in the following manner: $\sum_\mu K_\mu^\dagger K_\mu = \mathbf{1}$. Conversely, the Kraus representation encapsulates all possible physical dynamics of the system. That is, any map expressible as in (7.8) is automatically completely-positive and trace-preserving. For our case of interest – N-qubit systems –, index μ runs from 0 to $\left(2^N\right)^2 - 1 = 4^N - 1$. For later convenience, we will represent it in base 4, decomposing it as the following multi-index: $\mu \equiv \mu_1 \, ... \, \mu_N$, with $\mu_i = 0$, 1, 2, or 3 $\;\forall \; i \in \mathcal{V}$.

We call Λ a separable map with respect to some multipartition of the system if each and all of its Kraus operators factorize as tensor products of local operators each one with support on only one of the subparts. For example, if we split the qubits associated to the graph shown in Fig. 7.1 into a set $\mathcal{Y}$ of *boundary qubits* (black) and its complement $\overline{\mathcal{Y}} \equiv \mathcal{V}/\mathcal{Y}$ of *non-boundary qubits* (orange), Λ is separable with respect to this partitioning if $K_\mu \equiv K_{\mathcal{Y}_\mu} \otimes K_{\overline{\mathcal{Y}}_\mu}$, with $K_{\mathcal{Y}_\mu}$ and $K_{\overline{\mathcal{Y}}_\mu}$ operators acting non-trivially only on the Hilbert spaces of the boundary and non-boundary qubits, respectively. A separable map cannot increase the entanglement in the considered multipartition [GG08].

In turn, we call Λ an *independent map* with respect to some multipartition of the system if it can be factorized as the composition (tensor product) of individual maps acting independently on each subpart. Otherwise, we say that Λ is a *collective map*. Examples of *fully independent maps* are those in which each qubit i is independently subject to its own local noise channel $\mathcal{E}_i$. By the term *independent map* without explicit mention to any respective multipartition we will refer throughout to fully independent maps. In this case, the global map Λ factorizes completely:

$$\Lambda(\rho) = \mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \ldots \otimes \mathcal{E}_N(\rho). \tag{7.9}$$

It is important to notice that all independent maps are necessarily separable but a general separable map does not need to be factorable as in (7.9) and can therefore be both, either individual or collective.

7.2.3 Pauli maps

A crucial family of fully separable maps is that of the *Pauli maps*, for which every Kraus operator is proportional to a product of individual Pauli and identity operators acting on each qubit. That is, $K_\mu \equiv \sqrt{P_{(\mu_1,\ \dots\ \mu_N)}}\ \sigma_{1\mu_1} \otimes \ldots \otimes \sigma_{N\mu_N} \equiv \sqrt{P_\mu}\ \sigma_\mu$, with $\sigma_{i0} = \mathbf{1}_i$ (the identity operator on qubit i), $\sigma_{i1} = X_i$, $\sigma_{i2} = Y_i$, and $\sigma_{i3} = Z_i$, and $P_{(\mu_1,\ \dots\ \mu_N)} \equiv P_\mu$ any probability distribution. Popular instances are the (collective or independent) depolarization and dephasing (also called phase damping, or phase-flip) maps, and the (individual) bit-flip and bit-phase-flip channels [NC00]. For example, the independent depolarizing (D) channel describes the situation in which the qubit remains untouched with probability $1-p$, or is depolarized - meaning that it is taken to the maximally mixed state (white noise) - with probability p. It is characterized by the fully-factorable probability $P_\mu = p_{1\mu_1} \times \dots p_{N\mu_N}$, with $p_{i0} = 1-p$ and $p_{i1} = p_{i2} = p_{i3} = p/3,\ \forall\ i \in \mathcal{V}$. The independent phase damping (PD) channel in turn induces the complete loss of quantum coherence with probability p, but without any energy (population) exchange. It is also given by a fully-factorable probability with $p_{i0} = 1-p/2$, $p_{i1} = 0 = p_{i2}$, and $p_{i3} = p/2,\ \ \forall\ i \in \mathcal{V}$.

For later convenience, we finally recall that each Pauli operator $\sigma_{i\mu_i}$ can be written in the following way (also called the *chord representation* [AGMS04]): $T_{i(u_i,v_i)} \equiv Z_i^{v_i}.X_i^{u_i}$, with u_i and $v_i = 0$, or 1. Indeed, notice that $\sigma_{i2v_i+|v_i+u_i|_2} = T_{i(u_i,v_i)}$ (up to an irrelevant phase factor for $u_i = 1 = v_i$), where "$|\ |_2$" stands for modulo 2. In this representation, the Kraus decomposition of the above-considered general Pauli map has the following Kraus operators: $K_{C(U,V)} \equiv \sqrt{P_{C(u_1,v_1,\ \dots\ u_N,v_N)}}\ T_{1(u_1,v_1)} \otimes \ldots \otimes T_{N(u_N,v_N)} \equiv \sqrt{P_{C(U,V)}}\ T_{(U,V)}$, where $U \equiv (u_1,\ \dots\ u_N)$ and $V \equiv (v_1,\ \dots\ v_N)$. The probability $P_{C(U,V)} \equiv P_{C(u_1,v_1,\ \dots\ u_N,v_N)}$ in turn is related to the original P_μ by $P_{C(u_1,v_1,\ \dots\ u_N,v_N)} \equiv P_{(2v_1+|v_1+u_1|_2,\ \dots\ ,2v_N+|v_N+u_N|_2))}$.

7.2.4 The thermal bath

An important example of a non-Pauli, independent map is the generalized amplitude-damping channel (GAD) [NC00]. It represents energy diffusion and dissipation with a thermal bath into which each qubit is individually immersed. Its Kraus representation is

$$K_{i\mu_i=0} \equiv \sqrt{\frac{\overline{n}+1}{2\overline{n}+1}}(|0_i\rangle\langle 0_i| + \sqrt{1-p}|1_i\rangle\langle 1_i|), \tag{7.10a}$$

$$K_{i\mu_i=1} \equiv \sqrt{\frac{\overline{n}+1}{2\overline{n}+1}p}|0_i\rangle\langle 1_i|, \tag{7.10b}$$

$$K_{i\mu_i=2} \equiv \sqrt{\frac{\overline{n}}{2\overline{n}+1}}(\sqrt{1-p}|0_i\rangle\langle 0_i| + |1_i\rangle\langle 1_i|), \tag{7.10c}$$

and

$$K_{i\mu_i=3} \equiv \sqrt{\frac{\overline{n}}{2\overline{n}+1}}p|1_i\rangle\langle 0_i|. \tag{7.10d}$$

Here $\overline{n}$ is the average number of quanta in the thermal bath, $p \equiv p(t) \equiv 1-e^{-\frac{1}{2}\gamma(2\overline{n}+1)t}$ is the probability of the qubit exchanging a quantum with the bath after a time t, and γ is the zero-temperature dissipation rate. Channel GAD is actually the extension to finite temperature of the purely dissipative amplitude damping (AD) channel, which is obtained from GAD in the zero-temperature limit $\overline{n} = 0$. In the opposite extreme, the purely diffusive case is obtained from GAD in the composite limit $\overline{n} \to \infty$, $\gamma \to 0$, and $\overline{n}\gamma = \Gamma$, where Γ is the diffusion constant. Note that in the purely-diffusive limit channel GAD becomes a Pauli channel, with defining individual probabilities $p_{i0} = \frac{1}{2}(1 - p/2 + \sqrt{1-p})$, $p_{i1} = \frac{p}{4} = p_{i2}$, and $p_{i3} = \frac{1}{2}(1 - p/2 - \sqrt{1-p})$, $\forall\ i \in \mathcal{V}$.

Finally, the probability p in channels D, PD and GAD above can be interpreted as a convenient parametrization of time, where $p = 0$ refers to the initial time 0 and $p = 1$ refers to the asymptotic $t \to \infty$ limit.

7.3 Methods

As mentioned before, the direct calculation of the entanglement in arbitrary mixed states is a task exponentially hard in the system's size [PV07, HHHH09]. In this section, we elaborate in detail a formalism that dramatically simplifies this task for graph – or graph-diagonal – states undergoing a noisy evolution in a fully general context. Along the way, we also describe carefully which requirements an arbitrary noisy map has to satisfy so that the formalism can be applied.

7.3.1 Evolution of graph-state entanglement under generic noise: the general idea

Consider a system initially in graph state (7.1) that evolves during a time t according to the general map (7.8) towards the evolved state

$$\rho_t \equiv \Lambda(|G_{(\mathcal{V},\mathcal{C})_0}\rangle) = \sum_{\mu} K_\mu |G_{(\mathcal{V},\mathcal{C})_0}\rangle\langle G_{(\mathcal{V},\mathcal{C})_0}| K_\mu^\dagger. \tag{7.11}$$

We would like to follow the entanglement $E(\rho_t)$ of ρ_t during its entire evolution. Here, E is any convex entanglement monotone [PV07, HHHH09] that quantifies the entanglement content in some given multi-partition of the system. An example of such multi-partition is displayed in Fig. 7.1, where the associated graph is split into two subsets, painted respectively in yellow and white in the figure. The edges that connect vertices at different subsets are called the *boundary-crossing edges* and are painted in black in the figure. We call the set of all the boundary-crossing edges $\mathcal{X} \subseteq \mathcal{C}$, and its complement $\overline{\mathcal{X}} \equiv \mathcal{C}/\mathcal{X}$ the set of all non-boundary-crossing edges. All the qubits associated to vertices connected by any edge in $\mathcal{X}$ constitute the set $\mathcal{Y} \subseteq \mathcal{V}$ of *boundary qubits* (or *boundary subsystem*), and its complement $\overline{\mathcal{Y}} \equiv \mathcal{V}/\mathcal{Y}$ is the non-boundary qubit set. We refer to $G_{(\mathcal{Y},\mathcal{X})}$ as the *boundary sub-graph.*

We can use this classification and the operational definition (7.1) to write the initial graph state as

$$|G_{(\mathcal{V},\mathcal{C})_0}\rangle = \bigotimes_{\{i,j\}\in\overline{\mathcal{X}}} CZ_{ij}|G_{(\mathcal{Y},\mathcal{X})_0}\rangle \otimes |g_{(\overline{\mathcal{Y}})_0}\rangle, \tag{7.12}$$

where $|g_{(\overline{\mathcal{Y}})_0}\rangle \equiv \bigotimes_{i\in\overline{\mathcal{Y}}} |+_i\rangle$. In other words, we explicitly factor out all the CZ gates corresponding to non-boundary qubits.

Consider now the application of some Kraus operator K_μ of a general map on graph state (7.12): $K_\mu \bigotimes_{\{i,j\}\in\overline{\mathcal{X}}} CZ_{ij}|G_{(\mathcal{Y},\mathcal{X})_0}\rangle \otimes |g_{(\overline{\mathcal{Y}})_0}\rangle$. The latter can always be written as $\bigotimes_{\{i,j\}\in\overline{\mathcal{X}}} CZ_{ij}\tilde{K}_\mu|G_{(\mathcal{Y},\mathcal{X})_0}\rangle \otimes |g_{(\overline{\mathcal{Y}})_0}\rangle$, with

$$\tilde{K}_\mu = \bigotimes_{\{i,j\}\in\overline{\mathcal{X}}} CZ_{ij}\ K_\mu \bigotimes_{\{i',j'\}\in\overline{\mathcal{X}}} CZ_{i'j'},\ \forall\ \mu, \tag{7.13}$$

Now, consider every map Λ such that transformation rule (7.13) yields, for each and all of its Kraus operators, modified Kraus operators of the form

$$\tilde{K}_\mu = \tilde{K}_{\mathcal{Y}\gamma} \otimes \tilde{K}_{\overline{\mathcal{Y}}\omega}, \tag{7.14}$$

where $\tilde{K}_{\mathcal{Y}\gamma}$ and $\tilde{K}_{\overline{\mathcal{Y}}\omega}$ are normalized (modified) Kraus operators acting non-trivially only on the boundary and non-boundary qubits, respectively, and $\gamma = \{\mu_i,\ \ i \in \mathcal{Y}\}$ and $\omega = \{\mu_i,\ \ i \in \overline{\mathcal{Y}}\}$ are multi-indices [1] labeling respectively the alternatives for the boundary and non-boundary subsystems. The modified map $\tilde{\Lambda}$, composed of Kraus operators $\tilde{K}_\mu$ is then clearly bi-separable with respect to the bi-partition “boundary / non-boundary”. For

[1] That μ, γ and δ may be multi-indices is not implicit from the context or the definition of μ in Eqn. (7.11). However, for the purpose of our later examples, we already term them so.

all such maps the calculation of $E(\rho_t)$ can be drastically simplified, as we see in what follows.

In these cases, the evolved state (7.11) can be written as

$$\rho_t \equiv \Lambda(|G_{(\mathcal{V},\mathcal{C})_0}\rangle) = \bigotimes_{\{i,j\}\in\overline{\mathcal{X}}} CZ_{ij}\ \tilde{\rho}_t \bigotimes_{\{k,l\}\in\overline{\mathcal{X}}} CZ_{kl}. \tag{7.15}$$

with

$$\begin{aligned}
\tilde{\rho}_t &= \tilde{\Lambda}(|G_{(\mathcal{Y},\mathcal{X})_0}\rangle \otimes |g_{(\overline{\mathcal{Y}})_0}\rangle) \\
&= \sum_{\mu} \tilde{K}_{\mathcal{Y}\gamma(\mu)} |G_{(\mathcal{Y},\mathcal{X})_0}\rangle\langle G_{(\mathcal{Y},\mathcal{X})_0}| \tilde{K}^{\dagger}_{\mathcal{Y}\gamma(\mu)} \otimes \tilde{K}_{\overline{\mathcal{Y}}\omega(\mu)} |g_{(\overline{\mathcal{Y}})_0}\rangle\langle g_{(\overline{\mathcal{Y}})_0}| \tilde{K}^{\dagger}_{\overline{\mathcal{Y}}\omega(\mu)} \\
&= \sum_{\omega} \tilde{K}_{\overline{\mathcal{Y}}\omega} |g_{(\overline{\mathcal{Y}})_0}\rangle\langle g_{(\overline{\mathcal{Y}})_0}| \tilde{K}^{\dagger}_{\overline{\mathcal{Y}}\omega} \otimes \sum_{\gamma} \tilde{K}_{\mathcal{Y}(\gamma|\omega)} |G_{(\mathcal{Y},\mathcal{X})_0}\rangle\langle G_{(\mathcal{Y},\mathcal{X})_0}| \tilde{K}^{\dagger}_{\mathcal{Y}(\gamma|\omega)},
\end{aligned}$$

where $\tilde{K}_{\mathcal{Y}(\gamma|\omega)}$ is the γ-th modified Kraus operator on the boundary subsystem given that $\tilde{K}_{\overline{\mathcal{Y}}_\omega}$ has been applied to the non-boundary one. Recall that both $\gamma \equiv \gamma(\mu)$ and $\omega \equiv \omega(\mu)$ come from the same single multi-index μ and are therefore in general not independent on one another. In the second equality of (7.16) we have chosen to treat ω as an independent variable for the summation and make γ explicitly depend on ω. This can always be done and will be convenient for our purposes.

The crucial observation now is that the CZ operators explicitly factored out in the evolved state (7.15) correspond to non-boundary-crossing edges. Thus, they act as *local unitary operations* with respect to the multi-partition of interest. For this reason, and since local unitary operations do not change the entanglement content of any state, the equivalence

$$E(\rho_t) = E(\tilde{\rho}_t) \tag{7.16}$$

holds.

In the forthcoming subsections we see how, by exploiting this equivalence in different noise scenarios, the computational effort required for a reliable estimation (and in some cases, an exact calculation) of $E(\rho_t)$ can be considerably reduced. The main idea behind this reduction lies on the fact that, whereas in the general expression (7.11) the entanglement can be distributed among all particles in the graph, in state (7.16) the boundary and non-boundary subsystems are explicitly in a separable state. All the entanglement in the multi-partition of interest is therefore localized exclusively in the boundary subgraph. The situation is graphically represented

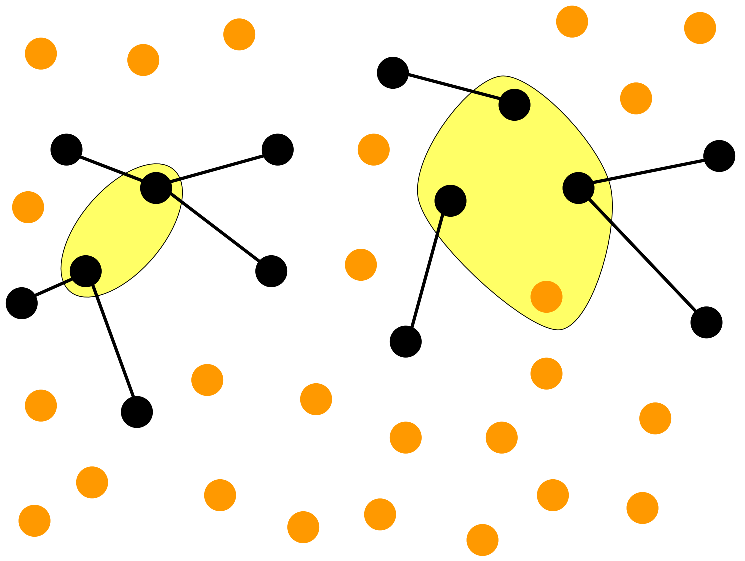

Figure 7.2: Same graph as in Fig. 7.1 but where all non-boundary-crossing edges have been erased, representing the fact that the boundary and non-boundary subsystems are fully unentangled. The entanglement in the whole system is obtained via a calculation involving only the smaller boundary subsystem.

in Fig. 7.2, where the same graph as in Fig. 7.1 is plotted but with all its non-boundary-crossing edges erased.

More precisely, the general approach consists of obtaining lower and upper bounds on $E(\rho_t)$ by bounding the entanglement of state (7.16) from above and below as explained in what follows.

Lower bounds to the entanglement evolution

The property of LOCC monotonicity of E, which means that the average entanglement cannot grow during an LOCC process [Vid00], allows us to derive lower-bounds on $E(\tilde{\rho}_t)$. The ones we consider can be obtained by the following generic procedure:

(i) after bringing the studied state into the form (7.16), apply some local general measurement $M = \{M_{\omega'}\}$, with measurement elements M_ω, on the non-boundary subsystem $\overline{\mathcal{Y}}$;

(ii) for each measurement outcome ω trace out the measured non-boundary subsystem, and finally

(iii) calculate the mean entanglement in the resulting state of the boundary subsystem $\mathcal{Y}$, averaged over all outcomes ω.

Since this procedure constitutes an LOCC with respect to the multipartition under scrutiny, the latter average entanglement can only be smaller than, or equal to, that of the initial state, *i.e.* :

$$E(\rho_t) = E(\tilde{\rho}_t) \geq \sum_{\omega} P_{\omega} E\Big(\sum_{\omega'} \frac{1}{P_{\omega}} \langle g_{(\overline{\mathcal{Y}})_0} | \tilde{K}^{\dagger}_{\overline{\mathcal{Y}}_{\omega'}} M^{\dagger}_{\omega} . M_{\omega} \tilde{K}_{\overline{\mathcal{Y}}_{\omega'}} | g_{(\overline{\mathcal{Y}})_0} \rangle \cdot \\ \sum_{\gamma} \tilde{K}_{\mathcal{Y}(\gamma|\omega')} | G_{(\mathcal{Y},\mathcal{X})_0} \rangle \langle G_{(\mathcal{Y},\mathcal{X})_0} | \tilde{K}^{\dagger}_{\mathcal{Y}(\gamma|\omega')} \Big), \tag{7.17}$$

with $P_{\omega} \equiv \sum_{\omega'} \langle g_{(\overline{\mathcal{Y}})_0} | \tilde{K}^{\dagger}_{\overline{\mathcal{Y}}_{\omega'}} M^{\dagger}_{\omega} . M_{\omega} \tilde{K}_{\overline{\mathcal{Y}}_{\omega'}} | g_{(\overline{\mathcal{Y}})_0} \rangle$ being the probability of outcome ω.

Notice that if the states $\{\tilde{K}_{\overline{\mathcal{Y}}_{\omega'}} | g_{(\overline{\mathcal{Y}})_0} \rangle\}$ of the non-boundary subsystem happen to be orthogonal, then there exists an optimal measurement $M = \{M_{\omega} \equiv \frac{\tilde{K}_{\overline{\mathcal{Y}}_{\omega}} | g_{(\overline{\mathcal{Y}})_0} \rangle \langle g_{(\overline{\mathcal{Y}})_0} | \tilde{K}^{\dagger}_{\overline{\mathcal{Y}}_{\omega}}}{\langle g_{(\overline{\mathcal{Y}})_0} | \tilde{K}^{\dagger}_{\overline{\mathcal{Y}}_{\omega}} \tilde{K}_{\overline{\mathcal{Y}}_{\omega}} | g_{(\overline{\mathcal{Y}})_0} \rangle}\}$ that can distinguish them unambiguously, so that we have $\langle g_{(\overline{\mathcal{Y}})_0} | \tilde{K}^{\dagger}_{\overline{\mathcal{Y}}_{\omega'}} M^{\dagger}_{\omega} . M_{\omega} \tilde{K}_{\overline{\mathcal{Y}}_{\omega'}} | g_{(\overline{\mathcal{Y}})_0} \rangle = \delta_{\omega,\omega'} \times \langle g_{(\overline{\mathcal{Y}})_0} | \tilde{K}^{\dagger}_{\overline{\mathcal{Y}}_{\omega}} \tilde{K}_{\overline{\mathcal{Y}}_{\omega}} | g_{(\overline{\mathcal{Y}})_0} \rangle$ and $P_{\omega} = \langle g_{(\overline{\mathcal{Y}})_0} | \tilde{K}^{\dagger}_{\overline{\mathcal{Y}}_{\omega}} \tilde{K}_{\overline{\mathcal{Y}}_{\omega}} | g_{(\overline{\mathcal{Y}})_0} \rangle$. In these cases an optimal lower bound is achieved as (constant normalization factors omitted again)

$$E(\rho_t) \geq \sum_{\omega} P_{\omega} E\big(\sum_{\gamma} \tilde{K}_{\mathcal{Y}(\gamma|\omega)} | G_{(\mathcal{Y},\mathcal{X})_0} \rangle \langle G_{(\mathcal{Y},\mathcal{X})_0} | \tilde{K}^{\dagger}_{\mathcal{Y}(\gamma|\omega)} \big). \tag{7.18}$$

Full distinguishability of the states in the non-boundary subsystem allows to reduce the mixing in the remaining boundary subsystem. In other words, the measurement outcome ω works as a perfect *flag* that marks which sub-ensemble of states of the boundary subsystem, from all those present in mixture (7.16), corresponds indeed to the obtained outcome.

In the opposite extreme, when states $\{\tilde{K}_{\overline{\mathcal{Y}}_{\omega'}} | g_{(\overline{\mathcal{Y}})_0} \rangle\}$ are all equal, no flagging information can be obtained via any measurement. In this case, the resulting bound is always equal to that obtained had we not made any measurement at all, but just directly taken the partial trace over $\overline{\mathcal{Y}}$ from (7.16):

$$E(\rho_t) \geq E\big(\frac{1}{2^{|\overline{\mathcal{Y}}|}} \sum_{\omega,\gamma} \tilde{K}_{\mathcal{Y}(\gamma|\omega)} | G_{(\mathcal{Y},\mathcal{X})_0} \rangle \langle G_{(\mathcal{Y},\mathcal{X})_0} | \tilde{K}^{\dagger}_{\mathcal{Y}(\gamma|\omega)} \big), \tag{7.19}$$

where $|\overline{\mathcal{Y}}|$ stands for the number of non-boundary qubits and full mixing over variable ω takes place now.

Henceforth we refer to lower bound (7.19) as the *lowest lower bound* (LLB). As its name suggests, its tightness is far from the optimal one given by (7.18). However, as we will see in the forthcoming subsections, due to the partial tracing, it typically does not depend on the total system's size but just on the boundary subsystem's.

This constitutes an appealing, useful property, for it allows one to draw generic conclusions about the robustness of entanglement in certain partitions of graph states, irrespective of their number of constituent particles (see examples below).

Upper bounds to the entanglement evolution

On the other hand, we consider upper-bounds on $E(\rho_t)$ based on the property of convexity of E, which essentially means that the entanglement of the convex sum is lower than, or equal to, the convex sum of the entanglements [PV07, HHHH09]. From (7.16), the latter implies that

$$\begin{aligned} E(\rho_t) = E(\tilde{\rho}_t) \leq \sum_{\omega} P_\omega E\big(\frac{1}{P_\omega}\tilde{K}_{\overline{\mathcal{Y}}_\omega}|g_{(\overline{\mathcal{Y}})_0}\rangle\langle g_{(\overline{\mathcal{Y}})_0}|\tilde{K}^\dagger_{\overline{\mathcal{Y}}_\omega}\otimes \\ \sum_{\gamma}\tilde{K}_{\mathcal{Y}(\gamma|\omega)}|G_{(\mathcal{Y},\mathcal{X})_0}\rangle\langle G_{(\mathcal{Y},\mathcal{X})_0}|\tilde{K}^\dagger_{\mathcal{Y}(\gamma|\omega)}\big), \end{aligned} \tag{7.20}$$

where, once again, $P_\omega = \langle g_{(\overline{\mathcal{Y}})_0}|\tilde{K}^\dagger_{\overline{\mathcal{Y}}_\omega}\tilde{K}_{\overline{\mathcal{Y}}_\omega}|g_{(\overline{\mathcal{Y}})_0}\rangle$. In each term of the last summation the boundary and non-boundary subsystems inside the brackets are in a *product state*. Therefore, as for what the multi-partition of interest concerns, the non-boudary subsystem works as a locally-added ancila (in a state $\frac{1}{P_\omega}\tilde{K}_{\overline{\mathcal{Y}}_\omega}|g_{(\overline{\mathcal{Y}})_0}\rangle\langle g_{(\overline{\mathcal{Y}})_0}|\tilde{K}^\dagger_{\overline{\mathcal{Y}}_\omega}$) and consequently does not have any influence on the amount of entanglement. This leads to the generic upper bound

$$E(\rho_t) \leq \sum_{\omega} P_\omega E\big(\sum_{\gamma}\tilde{K}_{\mathcal{Y}(\gamma|\omega)}|G_{(\mathcal{Y},\mathcal{X})_0}\rangle\langle G_{(\mathcal{Y},\mathcal{X})_0}|\tilde{K}^\dagger_{\mathcal{Y}(\gamma|\omega)}\big). \tag{7.21}$$

Exact entanglement

Notice that upper bound (7.21) and optimal lower bound (7.18) coincide. This means that, in the above-mentioned case when states $\{\tilde{K}_{\overline{\mathcal{Y}}_\omega}|g_{(\overline{\mathcal{Y}})_0}\rangle\}$ are orthogonal, these coincident bounds yield actually the exact value of $E(\rho_t)$:

$$E(\rho_t) = \sum_{\omega} P_\omega E\big(\sum_{\gamma}\tilde{K}_{\mathcal{Y}(\gamma|\omega)}|G_{(\mathcal{Y},\mathcal{X})_0}\rangle\langle G_{(\mathcal{Y},\mathcal{X})_0}|\tilde{K}^\dagger_{\mathcal{Y}(\gamma|\omega)}\big). \tag{7.22}$$

Expression (7.22) is still not an analytic closed formula for the exact entanglement of ρ_t, but reduces its calculation to that of the average entanglement over an ensemble of states of the boundary subsystem alone. More in detail, a brute-force calculation of $E(\rho_t)$ would require in general a convex optimization over the entire $(2^N)^2$-complex-parameter space. Through Eq. (7.22) in turn such calculation is reduced to that of the average entanglement over a sample of $2^{|\overline{\mathcal{Y}}|}$ states (one for each ω) of $|\mathcal{Y}|$ qubits, being $|\mathcal{Y}|$ the number of boundary qubits. The latter involves at most $2^{|\overline{\mathcal{Y}}|}$ independent optimizations over a $(2^{|\mathcal{Y}|})^2$-complex-parameter space. This, from the point of view of computational memory required, accounts for a reduction of resources by a factor of $(2^{|\overline{\mathcal{Y}}|})^2$. Alternatively, when computational memory is not a major restriction – for example if large classical-computer clusters are at hand –, one can take advantage of the fact that the $|\overline{\mathcal{Y}}|$ required optimizations in (7.22) are independent and therefore the calculation comes readily perfectly-suited for parallel computing. In this case, it is in the required computational time where an $|\overline{\mathcal{Y}}|$-exponentially large speed-up is gained.

In the cases where states $\{\tilde{K}_{\overline{\mathcal{Y}}\omega}|g_{(\overline{\mathcal{Y}})_0}\rangle\}$ are not orthogonal and the upper and lower bounds do not coincide, expressions (7.21) and (7.17) still yield highly non-trivial upper and lower bounds, respectively, as we discuss in Sec. 7.4.1.

Finally, it is important to stress that all the bounds derived here are general in the sense that they hold for any function fulfilling the *fundamental properties of convexity and monotonicity under LOCC processes.* This class includes genuine multipartite entanglement measures, as well as several quantities designed to quantify the usefulness of quantum states in the fulfilment of some given task for quantum-information processing or communication [PV07, HHHH09].

7.4 Application and examples

In the present section we apply the ideas of the previous section to some important concrete examples of noise processes. This shows how the method is helpful in the entanglement calculation for systems in natural dynamics physical scenarios. We first discuss the case of Pauli maps and then the generalized amplitude damping channel (thermal reservoir). Explicit calculations for noisy graph states composed of up to fourteen qubits are presented as examples.

7.4.1 Pauli maps on graph states

Pauli maps defined in Sec. 7.2.2 provide the most important and general subfamily of noise types for which expression (7.22) for the exact entanglement of the evolved state applies. In this case, every X_i or Y_i Pauli matrix in the map's Kraus operators is systematically substituted by products of Z_i and $\mathbf{1}_i$ according to rules (7.7). The resulting map $\tilde{\Lambda}$ defined in this way automatically commutes with any CZ gate and is fully separable, so that condition (7.14) is trivially satisfied. Since for every qubit in the system four orthogonal single qubit operators are mapped into products of just two, several different Kraus operators of the original map contribute to the same Kraus operator of the modified one. This allows us to simplify the notation going from indices μ_i, which run over 4 possible values each, to modified indices $\tilde{\mu}_i$ having only two different alternatives. In fact, the original operators K_μ give rise to only 2^N modified ones of the form

$$\tilde{K}_{\tilde{\mu}} = \sqrt{\tilde{P}_{\tilde{\mu}}} Z_1^{\tilde{\mu}_1} \otimes Z_2^{\tilde{\mu}_2} \otimes \cdots \otimes Z_N^{\tilde{\mu}_N} \equiv \sqrt{\tilde{P}_{\tilde{\mu}}} Z^{\tilde{\mu}} \tag{7.23}$$

where multi-index $\tilde{\mu}$ stands for the binary string $\tilde{\mu} = \tilde{\mu}_1 \ \ldots \ \tilde{\mu}_N$, with $\mu_i = 0$ or 1, $\forall i \in \mathcal{V}$. Probability $\tilde{P}_{\tilde{\mu}}$ is given simply by the summation of all P_μ in the original Pauli map over all the different events μ for which σ_μ yields – via rules (7.7) – the same modified operator $Z^{\tilde{\mu}}$ in (7.23).

To compute the latter modified probability we move to the chord notation [AGMS04], mentioned at the end of Sec. 7.2.3. Indeed, under transformation (7.7), we have that $T_{i(u_i,v_i)} \rightarrow Z_i^{v_{i1}} \otimes \bigotimes_{j\in\mathcal{N}_i} Z_j^{u_i}$, so that $T_{(U,V)} \equiv T_{1(u_1,v_1)} \otimes \ \ldots \ T_{N(u_N,v_N)} \rightarrow Z_1^{v_1+\sum_{j\in\mathcal{N}_1} u_j} \otimes \cdots \otimes Z_N^{v_N+\sum_{j\in\mathcal{N}_N} u_j}$. The latter coincides with $Z^{\tilde{\mu}}$ every time $\tilde{\mu}_i = |v_i + \sum_{j\in\mathcal{N}_i} u_j|_2, \ \forall \ i \in \mathcal{V}$. Thus, in this representation, the modified probability $\tilde{P}_{C\tilde{\mu}}$ is obtained from the defining probability $P_{C(U,V)}$ in the original map by the explicit formula

$$\tilde{P}_{C\tilde{\mu}} \equiv \sum_U P_{C(u_1,|\tilde{\mu}_1 - \sum_{j\in\mathcal{N}_1} u_j|_2, \ \ldots \ ,u_N,|\tilde{\mu}_1 - \sum_{j\in\mathcal{N}_N} u_j|_2)}. \tag{7.24}$$

The modified Kraus operators (7.23) in turn are fully separable, thus trivially satisfying factorization condition (7.14). We can express them as $\tilde{K}_{\tilde{\mu}} = \tilde{K}_{\mathcal{Y}\tilde{\gamma}} \otimes \tilde{K}_{\overline{\mathcal{Y}}\tilde{\omega}}$, with

$$\tilde{K}_{\mathcal{Y}\tilde{\gamma}} \equiv \tilde{K}_{\mathcal{Y}(\tilde{\gamma}|\tilde{\omega})} = \sqrt{\tilde{P}_{(\tilde{\gamma}|\tilde{\omega})}} Z^{\tilde{\gamma}} \text{ and } \ \tilde{K}_{\overline{\mathcal{Y}}\tilde{\omega}} = \sqrt{\tilde{P}_{\tilde{\omega}}} Z^{\tilde{\omega}}. \tag{7.25}$$

The new multi-indices are $\tilde{\gamma} = \{\tilde{\mu}_i, \ i \in \mathcal{Y}\}$ and $\tilde{\omega} = \{\tilde{\mu}_i, \ i \in \overline{\mathcal{Y}}\}$, and the corresponding probabilities satisfy $\tilde{P}_{(\tilde{\gamma}|\tilde{\omega})} \tilde{P}_{\tilde{\omega}} \equiv \tilde{P}_{\tilde{\mu}}$.

The states $\{\tilde{K}_{\overline{\mathcal{Y}}\tilde{\omega}'}|g_{(\overline{\mathcal{Y}})_0}\rangle = \sqrt{\tilde{P}_{\tilde{\omega}'}}Z^{\tilde{\omega}'}|+_i\rangle = \sqrt{\tilde{P}_{\tilde{\omega}'}}\bigotimes_{i\in\overline{\mathcal{Y}}}\frac{1}{\sqrt{2}}(|0_i\rangle+(-1)^{\tilde{\mu}_i}|1_i\rangle) \equiv \sqrt{\tilde{P}_{\tilde{\omega}'}}|g_{(\overline{\mathcal{Y}})_{\tilde{\omega}'}}\rangle\}$ are trivially checked to be all orthogonal. Thus, they provide perfect flags that mark each sub-ensemble in the boundary subsystem's ensemble. The perfect flags are revealed by local measurements on the non-boundary qubits in the product basis $\{|g_{(\overline{\mathcal{Y}})_{\tilde{\omega}}}\rangle\}$. Therefore, for Pauli maps the exact entanglement $E(\rho_t)$ can be calculated by expression (7.22), which, in terms of binary indexes $\tilde{\gamma}$ and $\tilde{\omega}$, and using graph-state relationship (7.3), can be finally expressed as

$$E(\rho_t) = \sum_{\tilde{\omega}} \tilde{P}_{\tilde{\omega}} E\big(\sum_{\tilde{\gamma}} \tilde{P}_{(\tilde{\gamma}|\tilde{\omega})}|G_{(\mathcal{Y},\mathcal{X})_{\tilde{\gamma}}}\rangle\langle G_{(\mathcal{Y},\mathcal{X})_{\tilde{\gamma}}}|\big), \tag{7.26}$$

In Fig. 7.3 we have plotted the bipartite entanglement of the exemplary bipartition of one qubit versus the rest fourteen and a twelve qubit graph states evolving under individual depolarization. This map, as said before, is characterized by the one-qubits Kraus operators $\sqrt{1-p}\mathbf{1}$, $\sqrt{p/3}X$, $\sqrt{p/3}Y$, and $\sqrt{p/3}Z$. The parameter p ($0 \leq p \leq 1$) refers to the probability that the map has acted: for $p = 0$ the state is left untouched and for $p = 1$ it is completely depolarized. Once more, p can be also set as a parametrization of time: $p = 0$ referring to the initial time (when nothing has occurred) and $p = 1$ referring to the asymptotic time $t \to \infty$ (when the system reaches its final steady state).

As the quantifier of entanglement, we choose the negativity [VW02] which is the absolute value of the sum of the negative eigenvalues of the density matrix partially transposed with respect to the considered bipartition. It has been defined formally in Section 2.5.3. Negativity, in general, fails to quantify the entanglement in those systems containing bound entanglement *i.e.* those with a positive partial transpose (PPT) [HHH98]. However, none of the examples (graph states) we work with bear any PPT bipartite entanglement. The negativity thus encapsulates all the relevant information about the separability in bipartitions of these states. In other words, null negativity implies separability in the corresponding partition. The most important feature of this quantifier relevant to us is its mathematical property of convexity and the ease of its computation since it involves only matrix diagonalization rather than optimizations. For these reasons, it is well suited for a simple illustration of our ideas.

We emphasize that, for the graph used in Fig. 7.3, a brute-force calculation would involve diagonalizing a $2^{14} \times 2^{14} = 16384 \times 16384$ density matrix for each value of p, whereas with the assistance of expression (7.26) $E(\rho_p)$ is calculated via diagonalization of many $2^3 \times 2^3 = 8 \times 8$ dimensional matrices only.

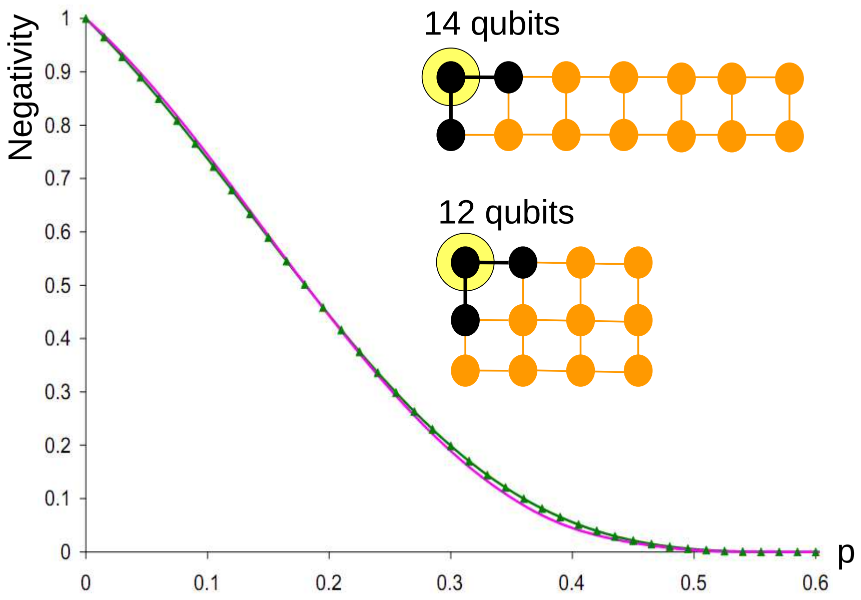

Figure 7.3: Negativity vs p for a 14 qubit (triangulated green curve) and 12 qubit (red curve) cluster state for depolarizing noise and a selected partition (shown in inset). The parameter p can be thought as a parametrization of time (see text).

In the case of Pauli maps the entanglement lower and upper bounds coincide, and provide the exact entanglement. However, this is not the case for general, non-Pauli, noise channels. The upper bound is given, as usual, by convexity. The lower bounds must be optimized by appropriate choice of LOCC operations. Here, we investigate and optimize measurement strategies for channel GAD, defined in Sec. 7.2.4.

Observe that the Kraus operators defined in Eqs. (7.10) satisfy the following: K_{i0} and K_{i2} commute with any CZ operator, while for every $j \in \mathcal{V}$ different from i it holds that $(K_{i1} \otimes \mathbf{1}_j).CZ_{ij} = CZ_{ij}.(K_{i1} \otimes Z_j)$ and $(K_{i3} \otimes \mathbf{1}_j).CZ_{ij} = CZ_{ij}.(K_{i3} \otimes Z_j)$. Based on this, one can perform the factorization in equation (7.13) and apply this way the formalism described in Sec. 7.3.1.

In what follows we focus on two main limits of channel GAD discussed in Sec. 7.2.4: the purely-dissipative limit $\overline{n} = 0$ (amplitude damping), and the purely-difusive limit $\overline{n} \to \infty$, $\gamma \to 0$, and $\overline{n}\gamma = \Gamma$.

7.4.2 Graph states under zero-temperature dissipation

We consider a four qubit linear (1D) cluster state subjected to the AD map and study the decay of entanglement in the partition consisting of the first qubit versus the rest shown in the inset of Fig. (7.4). Along with the exact calculation of entanglement via brute-force diagonalization of the partially-transposed matrices, the lowest lower bound LLB, obtained by tracing out the flags, and the upper bound (7.21), obtained from convexity, are plotted. In addition, the tightness of the lower bounds obtained by the flag measurements can be scanned as a function of the measurement bases.

Based on observations about the behavior of the system under the AD map we can guess good measurement strategies. For example, examination of the initial state reveals that at $p = 0$ each of the non-boundary qubits is in one of the states of the basis $\{|+\rangle, |-\rangle\}$; whereas at $p = 1$, in one of the states of $\{|0\rangle, |1\rangle\}$. We call the lower bound corresponding to measurements in the basis $\{|+\rangle, |-\rangle\}$ LB($\pi/4$), and the one obtained through measurements in $\{|0\rangle, |1\rangle\}$ LB(0). The latter bounds are the two additional curves plotted in Fig. (7.4). We observe that LB(0) provides only a slight improvement over the LLB, whereas LB($\pi/4$) appears to give a significant one. This raises the obvious question of how to optimize the choice of measurement basis at each instant p in the evolution.

As an illustration we consider lower bounds LB(θ) obtained through orthogonal measurements composed ot the projectors $|\theta+\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle$ and $|\theta-\rangle = -\sin\theta|0\rangle + \cos\theta|1\rangle$, and look for the angle θ that gives us approximately the largest value of LB(θ). This is certainly not the most general measurement scenario one may consider, but it gives a clue on how to increase the tightness of the bounds. Figs. (7.5) and (7.6) illustrate this idea.

At discrete values of p, we have varied parameter $0 \leq \theta \leq \pi/2$ through its whole range. The measured entanglement for each value of θ is compared against the exact entanglement at the given p. In physical terms, we are taking a snapshot of the system at discrete time instants. The non-boundary qubits are then measured in a range of different bases and the lower bounds to the entanglement are computed after each measurement. The value of θ at each time interval that gives the maximum entanglement represents

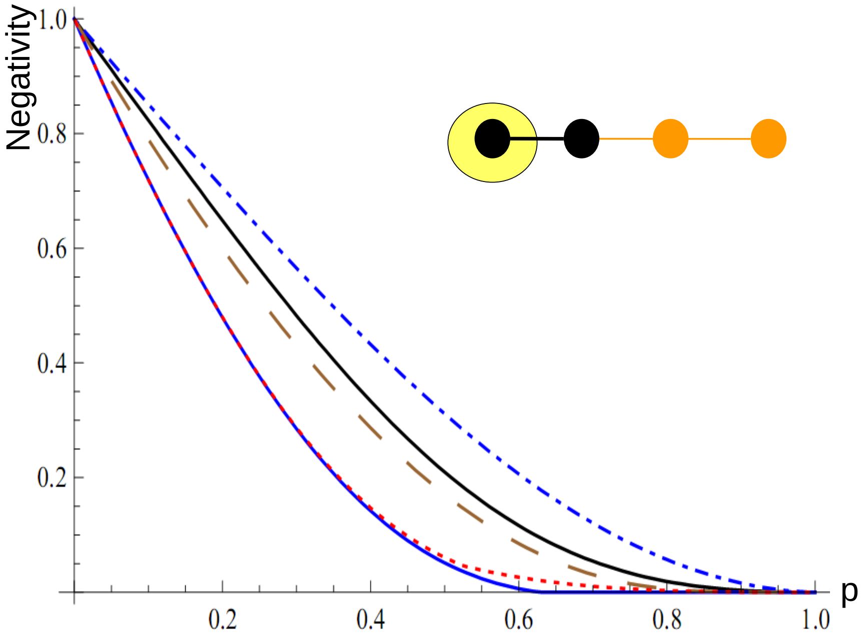

Figure 7.4: Negativity vs p for a 4 qubit linear cluster subjected to the amplitude damping channel and for the partition displayed in the inset. Solid (upper, black) curve: exact entanglement; Solid (lower, blue) curve: lowest lower bound (obtained by tracing out the flags); Dashed-dotted (blue) curve: upper bound (7.21) obtained by convexity; Dotted (red) curve: LB(0) obtained by measuring the flags in the Z basis; Dashed (brown) curve: LB($\pi/4$) obtained by measuring the flags in the X basis.

the optimal basis for measurement at that particular time interval. It is clearly seen that, for small values of p, angles around $\theta = \frac{\pi}{4}$ give the closest approximations to the exact entanglement, in consistence with the significant improvement of LB($\pi/4$) over the LLB observed in Fig. (7.4). For large values of p though, the best approximations tend to be given by the angles $\theta = 0$ or $\theta = \frac{\pi}{2}$, as can be observed in Fig. 7.6. It must still be kept in mind that none of these closest approximations equals the exact entanglement of the state.

7.4.3 Graph states under infinite-temperature difusion

We now consider the purely-diffusive case of the GAD channel, where each qubit is in contact with an independent bath of infinite temperature. In

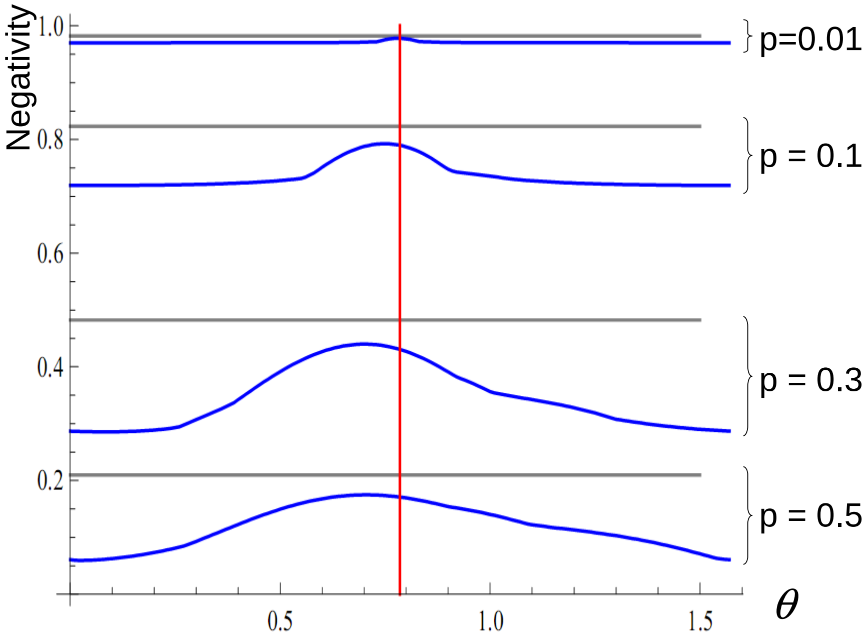

Figure 7.5: Lower bound LB(θ) to the negativity as a function of the angle θ in the measurement basis, for the same situation as in Fig. 7.4, and for fixed values of p. Each value $p = 0.01, 0.1, 0.3, 0.5$ has two curves associated to it. The horizontal (gray) straight line represents the exact entanglement at each p, while the blue (black) curve represents the bound LB(θ) at this p. The red line (vertical) corresponds to $\theta = \frac{\pi}{4}$, *i.e.* measurements in the basis $\{|+\rangle, |-\rangle\}$.

Fig. 7.7 we display the entanglement evolution in a similar way as in Fig. 7.4. Since in the purely-diffusive limit channel GAD becomes a Pauli map, as was mentioned in the end of Sec. 7.2.4, the bound LB($\pi/4$) yields the exact entanglement. LB(0) on the other hand coincides with the lowest lower bound LLB. The fact that in this case LB(θ) reaches the exact entanglement at $\theta = \frac{\pi}{4}$ can also be seen in a clearer way in Fig. 7.8.

In Fig. 7.7 upper bound (7.21) is plotted as well. Since in this case the channel is a Pauli channel, one would expect the upper bound to coincide with the exact entanglement as well. The fact that this does not occur is because, even though the noise itself is describable as a Pauli map, the plotted upper bound has been calculated using the original Kraus decomposition of

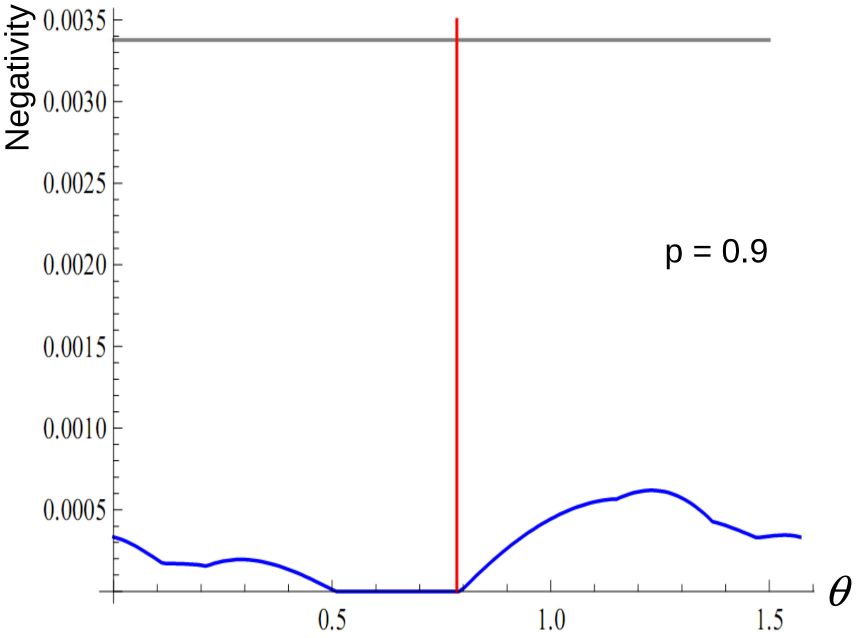

Figure 7.6: Same as fig. (7.5) for $p = 0.9$ plotted separately for clarity.

Eqs. (7.10), which is not in a Pauli-map form. For every given particular Kraus decomposition of a superoperator, the naive application of convexity always yields UB through Eq. (7.21), but this needs not the tightest, for the Kraus decomposition of a superoperator is in general not unique. This observation leads to a whole family of upper bounds for a given map. In the same spirit as with the lower bounds, one could in principle optimize the obtained UBs over all possible Kraus representations of the map.

7.5 Extentions and Limitations

The framework developed here is not restricted to graph states. The crucial ingredient in our formalism is the factorization of entangling operations that act as local unitary transformations in a considered partition and the redefinition of the Kraus operators acting on the state, reducing the entanglement evaluation problem to the boundary system. Given an entangled

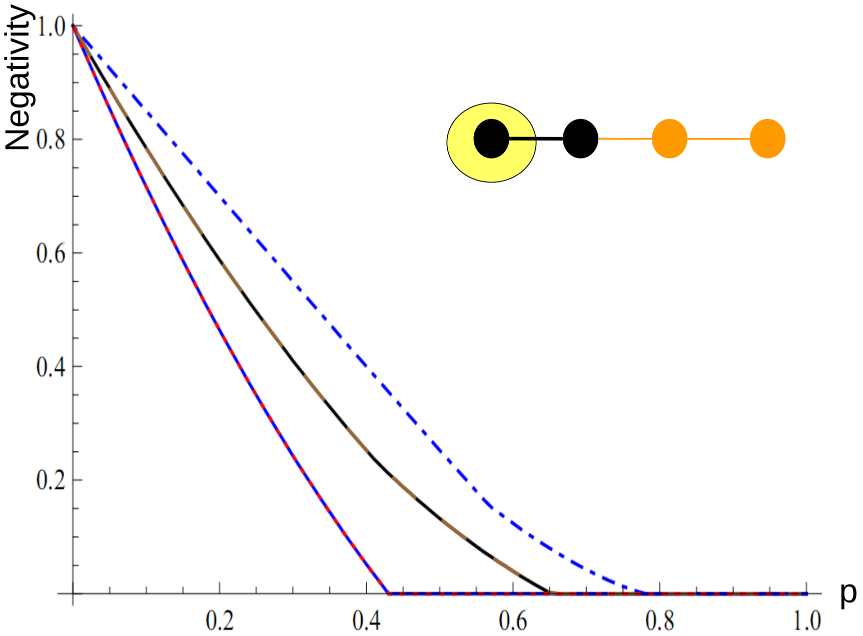

Figure 7.7: Negativity vs p for a 4 qubit linear cluster and a chosen partition (inset) subjected to the generalized amplitude damping channel in the diffusive limit $\overline{n} \to \infty$. The central curve in an exact overlap of the curves for exact entanglement (solid black) and LB($\pi/4$) (dashed brown). The lowest curve is an exact overlap of the curves for lowest lower bound (solid blue) and LB(0) (dotted red). The top most curve is the upper bound (7.21) (dot-dashed blue).

state and a prescription of its construction in terms of entangling operations, useful bounds and exact expressions for the entanglement can be readily obtained. As an example, a GHZ-like state $|\psi\rangle = \alpha\,|0\rangle^{\otimes N} + \beta\,|1\rangle^{\otimes N}$ can be operationally constructed by the sequential application of maximally-entangling operation $CNOT_{ij} = |0_i 0_j\rangle\,\langle 0_i 0_j| + |0_i 1_j\rangle\,\langle 0_i 1_j| + |1_i 0_j\rangle\,\langle 1_i 1_j| + |1_i 1_j\rangle\,\langle 1_i 0_j|$ to the product state $(\alpha|0\rangle + \beta|1\rangle) \otimes |0\rangle \otimes ... \otimes |0\rangle$ such that $|\psi\rangle = \bigotimes_{i=1}^{N-1} CNOT_{i,i+1}(\alpha|0\rangle + \beta|1\rangle) \otimes |0\rangle \otimes ... \otimes |0\rangle$. Using our techniques and the permutation symmetry of the state it can be seen that the entanglement evaluation in any bipartition can be reduced to that of a two qubits system. It is also worth noticing that the techniques presented here can also be extended to higher-dimensional graph states [ZZXS03, TPH$^+$06].

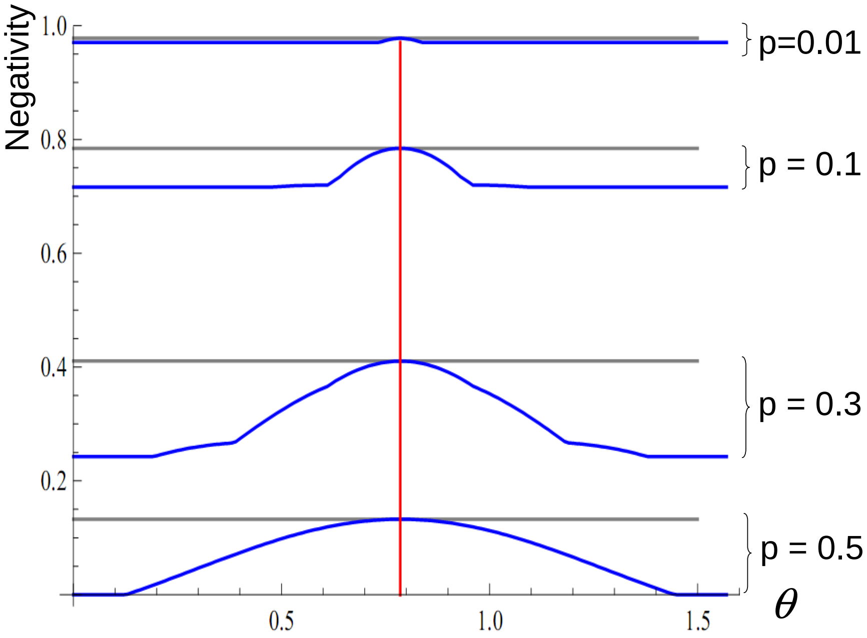

Figure 7.8: Lower bound LB(θ) to the negativity as a function of the angle θ in the measurement basis, for the same situation as in Fig. 7.7, and for fixed values of p. Each value $p = 0.01, 0.1, 0.3, 0.5$ has two curves associated to it. Again, the horizontal (gray) straight line represents the exact entanglement at each p, while the blue (black) curve represents the bounds LB(θ) at the this p. The red line (vertical) corresponds to $\theta = \frac{\pi}{4}$, i.e. measurement in the basis $\{|+\rangle, |-\rangle\}$.

In addition, it is important to mention that all bounds developed so far can in fact also be exploited to follow the entanglement evolution when the system's initial state is a mixed graph-diagonal state. This is simply due to the fact that any graph-diagonal state (7.4) can be thought of as a Pauli map Λ_{GD} acting on a pure graph state:

$$\begin{aligned}\rho_{GD} &= \sum_{\nu} P_{\nu} |G_{(\mathcal{V},\mathcal{C})_{\nu}}\rangle\langle G_{(\mathcal{V},\mathcal{C})_{\nu}}| \\ &= \sum_{\nu} P_{\nu} Z^{\nu} |G_{(\mathcal{V},\mathcal{C})_{0}}\rangle\langle G_{(\mathcal{V},\mathcal{C})_{0}}| Z^{\nu} \\ &= \Lambda_{GD}(|G_{(\mathcal{V},\mathcal{C})_{0}}\rangle).\end{aligned} \tag{7.27}$$

Thus, the entanglement at any time t in a system initially in a mixed graph-diagonal state ρ_{GD}, and evolving under some map Λ, is equivalent to that of an initial pure graph state $|G_{(\mathcal{V},\mathcal{C})_0}\rangle$ whose evolution is ruled by the composed map $\Lambda \circ \Lambda_{GD}$, where Λ_{GD} is defined in (7.27). When Λ is itself a Pauli map, then $\Lambda \circ \Lambda_{GD}$ is also a Pauli map and the expression (7.26) for the exact entanglement can be applied. For the cases where Λ is not a Pauli map but the relations (7.13) are satisfied by its Kraus operators, the relations (7.13) will also be satisfied by the composed map $\Lambda \circ \Lambda_{GD}$, so that all other bounds derived here also hold.

Furthermore, as briefly mentioned before, any arbitrary state can be depolarized by some separable map towards a graph-diagonal state [DAB03, ADB05]. The latter, since the entanglement of any state cannot increase under separable maps, implies that all the lower bounds presented here also provide lower bounds to the decay of the entanglement that, though in general far from tight, apply to *any arbitrary initial state subject to any decoherence process.*

The gain in computational effort provided by the machinery presented here diminishes with the ratio between the number of particles in the boundary subsystem and the total number of particles. For example, for multipartitions such that the boundary system is the total system itself, or for entanglement quantifiers that do not refer to any multi-partition at all, our method yields no gain. An example of the latter are the entanglement measures that treat all parties in a system indistinguishably, some of which, as was mentioned in the introduction, have been studied in Refs. [GBB08, HW10]. These methods naturally complement with ours to offer a rather general and versatile toolbox for the study of the open-system dynamics of graph-state entanglement.

7.6 Discussion

We have studied in detail a general framework for computing the entanglement of a graph state under decoherence introduced in [CCA$^+$09]. This framework allows to drastically reduce the effort in computing the entanglement evolution of graph states in several physical scenarios. We have given an explicit formula for the construction of the effective noise involved in the calculation of the entanglement for Pauli maps and extended the formalism to the case of independent baths at arbitrary temperature. Also, we have elaborated the formalism to construct non-trivial lower and upper bounds to the entanglement decay where exact results cannot be obtained from the formalism itself.

Finally we would like to add that the necessary requirements on the noise channels for the method to apply do not prevent us from obtaining general results for a wide variety of realistic decoherence processes. Furthermore, the conditions required on the entanglement measures are satisfied by most quantifiers.

Summary and outlook

In this thesis, we have explored several aspects of randomness involving its generation, quantification and amplification (the latter two being device independent). We have discussed at length an explanation of upper bounds on randomness in non-local theories based on symmetries. Not only do these arguments help identify maximal randomness (or the lack of it) in both the quantum and non-signalling sets but also shed light on the relationship between the resource of non-locality and that of randomness. We have also shown for the first time that randomness can be quantified for arbitrary system size given specific correlations and can be fully amplified given even the lowest quality initial source of randomness. The criteria and techniques developed can be extended to many other scenarios.

In all this work, we have extensively used the so-called GHZ correlations. These were already known to be of great utility in quantum information since they display both genuine N-partite entanglement and non-locality. The present work demonstrates that these correlations also encode robust randomness. Even allowing non-signalling preparations, one can distil randomness in a Bell experiment allowing an arbitrary (but not absolute) weakening of the freedom of choice assumption.

This work raises several questions. Uniqueness of quantum correlations maximally violating Bell inequalities is a conjecture we have shown to be very useful to detect maximal randomness. While a formal proof in full generality appears to be difficult, the significance of such a result makes an attempt worthwhile. The NPA semi-definite hierarchy bounding quantum correlations is potentially a significant tool to attack this problem. Concerning the bounds to randomness in the non-signalling set, it would be of great interest to find tighter ones than the ones we derive. Furthermore, there is the outstanding question if full randomness amplification can be achieved in the bipartite scenario.

Most of our results are derived for ”pure” correlations free of noise. From an experimental point of view this is a significant drawback. Laboratory preparations of such correlations involve noise in either the preparation or measurement or both. This relates to our final study of the entanglement properties of GHZ correlations (as an example of the much larger class of graph states) evolving under noise. This suggests several avenues for further work, some of which we are currently working on. One is the characterization of intrinsic randomness for GHZ correlations subject to noise. Work in progress suggests that the randomness content in such states decreases linearly with the visibility under white noise. Analytical results, where possible, for arbitrary system size would be a very useful demonstration of the relationship between not just non-locality and randomness but also with entanglement, which has been computed in the final chapter. Moreover, such results would allow experimental generation of verifiable randomness with the advantage of a weakened freedom of choice assumption. This has not yet been achieved in any experimental scenario to our knowledge.

Appendix A

Proof of full randomness amplification

Before entering the details of the Theorem in the main text, let us introduce a convenient notation. In what follows, we sometimes treat conditional probability distributions as vectors. To avoid ambiguities, we explicitly label the vectors describing probability distributions with the arguments of the distributions in upper case. Thus, for example, we denote by $P(\mathbf{A}|\mathbf{X})$ the $(2^5 \times 2^5)$-dimensional vector with components $P(\mathbf{a}|\mathbf{x})$ for all $\mathbf{a}, \mathbf{x} \in \{0,1\}^5$. With this notation, the five-partite Mermin inequality can be written as the scalar product

$$I \cdot P(\mathbf{A}|\mathbf{X}) = \sum_{\mathbf{a},\mathbf{x}} I(\mathbf{a},\mathbf{x}) P(\mathbf{a}|\mathbf{x}) \geq 6 \ . \tag{A.1}$$

Any probability distribution $P(\mathbf{a}|\mathbf{x})$ satisfies $C \cdot P(\mathbf{A}|\mathbf{X}) = 1$, where C is the vector with components $C(\mathbf{a},\mathbf{x}) = 2^{-5}$. We also use this scalar-product notation for full blocks, as in

$$I^{\otimes N_d} \cdot P(B|Y) = \sum_{\mathbf{a}_1,\ldots\mathbf{a}_{N_d}} \sum_{\mathbf{x}_1,\ldots\mathbf{x}_{N_d}} \left[\prod_{i=1}^{N_d} I(\mathbf{a}_i,\mathbf{x}_i)\right] P(\mathbf{a}_1,\ldots\mathbf{a}_{N_d}|\mathbf{x}_1,\ldots\mathbf{x}_{N_d}) \ .$$

Following our upper/lower-case convention, the vector $P(B|Y,e,z)$ has components $P(b|y,e,z)$ for all b, y but fixed e, z.

Recall also that the bits produced by the source $\mathcal{S}$ are such that the probability $P\left(x_j|rest\right)$ that bit j takes a given value x_j, conditioned on e and the rest of the bits from the source is bounded by,

$$\epsilon \leq P\left(x_j|rest\right) \leq 1-\epsilon, \tag{A.2}$$

where $0 < \epsilon \leq 1/2$. The bound, when applied to n-bit strings produced by the ϵ-source, implies that

$$\epsilon^n \leq P\left(x_1, \ldots, x_n | e\right) \leq (1-\epsilon)^n. \tag{A.3}$$

The proof of the Theorem in the main text relies on two crucial lemmas, which are stated and proven in Sections A.1.1 and A.1.2, respectively. The first lemma bounds the distinguishability between the distribution distilled from a block of N_d quintuplets and the ideal free random bit as function of the Bell violation (A.1) in each quintuplet. In particular, it guarantees that, if the correlations of all quintuplets in a given block violate inequality (A.1) sufficiently much, the bit distilled from the block will be indistinguishable from an ideal free random bit. The second lemma is required to guarantee that, if the statistics observed in all blocks but the distilling one are consistent with a maximal violation of inequality (A.1), the violation of the distilling block will be arbitrarily large.

A.1 Proof of the Theorem

We begin with the identity

$$P(\text{guess}) = P(g=0)P(\text{guess}|g=0) + P(g=1)P(\text{guess}|g=1)\ . \tag{A.4}$$

As discussed, when the protocol is aborted ($g = 0$) the distribution generated by the protocol and the ideal one are indistinguishable. In other words,

$$P(\text{guess}|g=0) = \frac{1}{2}\ . \tag{A.5}$$

If $P(g = 0) = 1$ then the protocol is secure, though in a trivial fashion. Next we address the non-trivial case where $P(g = 1) > 0$.

From the definition of the guessing probability, we have

$$\begin{aligned}
& P(\text{guess}|g=1) \\
= \;& \frac{1}{2}+\frac{1}{4}\sum_{k,\tilde{y},t}\max_{z}\sum_{e}\left|P(k,\tilde{y},t,e|z,g=1)-\frac{1}{2}P(\tilde{y},t,e|z,g=1)\right| \\
= \;& \frac{1}{2}+\frac{1}{4}\sum_{\tilde{y},t}P(\tilde{y},t|g=1)\sum_{k}\max_{z}\sum_{e}\left|P(k,e|z,\tilde{y},t,g=1)-\frac{1}{2}P(e|z,\tilde{y},t,g=1)\right| \\
\leq \;& \frac{1}{2}+\frac{1}{4}\sum_{\tilde{y},t}P(\tilde{y},t|g=1)\,6\sqrt{N_d}\;(\alpha C+\beta I)^{\otimes N_d}\cdot P(\tilde{B}|\tilde{Y},t,g=1) \\
= \;& \frac{1}{2}+\frac{3\sqrt{N_d}}{2}\,(\alpha C+\beta I)^{\otimes N_d}\cdot\sum_{\tilde{y},t}P(\tilde{y},t|g=1)P(\tilde{B}|\tilde{Y},t,g=1) \\
= \;& \frac{1}{2}+\frac{3\sqrt{N_d}}{2}\,(\alpha C+\beta I)^{\otimes N_d}\cdot\sum_{t}P(t|g=1)P(\tilde{B}|\tilde{Y},t,g=1) \\
= \;& \frac{1}{2}+\frac{3\sqrt{N_d}}{2}\,(\alpha C+\beta I)^{\otimes N_d}\cdot\sum_{t}P(\tilde{B},t|\tilde{Y},g=1) \\
= \;& \frac{1}{2}+\frac{3\sqrt{N_d}}{2}\,(\alpha C+\beta I)^{\otimes N_d}\cdot P(\tilde{B}|\tilde{Y},g=1)
\end{aligned} \tag{A.6}$$

where the inequality is due to Lemma 16 in Section A.1.1, we have used the no-signaling condition through $P(\tilde{y},t|z,g=1)=P(\tilde{y},t|g=1)$, in the second equality, and Bayes rule in the second and sixth equalities. From (A.6) and Lemma 17 in Section A.1.2, we obtain

$$P(\text{guess}|g=1)\;\leq\;\frac{1}{2}+\frac{3\sqrt{N_d}}{2}\left[\alpha^{N_d}+\frac{2\,N_b^{\log_2(1-\epsilon)}}{P(g=1)}\left(32\beta\epsilon^{-5}\right)^{N_d}\right]\;. \tag{A.7}$$

Finally, substituting bound (A.7) and equality (A.5) into (A.4), we obtain

$$P(\text{guess})\;\leq\;\frac{1}{2}+\frac{3\sqrt{N_d}}{2}\left[P(g=1)\,\alpha^{N_d}+2\,N_b^{\log_2(1-\epsilon)}\left(32\beta\epsilon^{-5}\right)^{N_d}\right]\;, \tag{A.8}$$

which, together with $P(g=1)\leq 1$, implies

$$P(\text{guess})\;\leq\;\frac{1}{2}+\frac{3\sqrt{N_d}}{2}\left[\alpha^{N_d}+2\,N_b^{\log_2(1-\epsilon)}\left(32\beta\epsilon^{-5}\right)^{N_d}\right]\;. \tag{A.9}$$

and in turn, proves the Theorem in the main text.

A.1.1 Statement and proof of Lemma 16

As mentioned, Lemma 16 provides a bound on the distinguishability between the probability distribution obtained after distilling a block of N_d quintuplets and an ideal free random bit in terms of the Bell violation (A.1) in each quintuplet. The proof of Lemma 16, in turn, requires two more lemmas, Lemma 18 and Lemma 19, stated and proven in Section A.1.3.

Lemma 16. *For each integer* $N_d \geq 130$ *there exists a function* $f : \{0,1\}^{N_d} \to \{0,1\}$ *such that, for any given* $(5N_d+1)$*-partite non-signaling distribution* $P(\mathbf{a}_1, \dots \mathbf{a}_{N_d}, e|\mathbf{x}_1, \dots \mathbf{x}_{N_d}, z) = P(b, e|y, z)$*, the random variable given by* $k = f(\mathrm{maj}(\mathbf{a}_1), \dots \mathrm{maj}(\mathbf{a}_{N_d}))$ *satisfies*

$$\sum_k \max_z \sum_e \left| P(k, e|y, z) - \frac{1}{2} P(e|y, z) \right| \ \leq \ 6\sqrt{N_d}\,(\alpha C + \beta I)^{\otimes N_d} \cdot P(B|Y) \tag{A.10}$$

for all inputs $y = (\mathbf{x}_1, \dots \mathbf{x}_{N_d}) \in \mathcal{X}^{N_d}$*, and where* α *and* β *are real numbers such that* $0 < \alpha < 1 < \beta$*.*

Proof of Lemma 16. For any $\mathbf{x}_0 \in \mathcal{X}$ let $M_w^{\mathbf{x}_0}$ be the vector with components $M_w^{\mathbf{x}_0}(\mathbf{a}, \mathbf{x}) = \delta^w_{\mathrm{maj}(\mathbf{a})} \delta^{\mathbf{x}_0}_{\mathbf{x}}$. The probability of getting $\mathrm{maj}(\mathbf{a}) = w$ when using $\mathbf{x}_0$ as input can be written as $P(w|\mathbf{x}_0) = M_w^{\mathbf{x}_0} \cdot P(\mathbf{A}|\mathbf{X})$. Note that this probability can also be written as $P(w|\mathbf{x}_0) = \Gamma_w^{\mathbf{x}_0} \cdot P(\mathbf{A}|\mathbf{X})$, where $\Gamma_w^{\mathbf{x}_0} = M_w^{\mathbf{x}_0} + \Lambda_w^{\mathbf{x}_0}$ and $\Lambda_w^{\mathbf{x}_0}$ is any vector orthogonal to the no-signaling subspace, that is, such that $\Lambda_w^{\mathbf{x}_0} \cdot P(\mathbf{A}|\mathbf{X}) = 0$ for all non-signaling distribution $P(\mathbf{A}|\mathbf{X})$. We can then write the left-hand side of (A.10) as

$$\begin{aligned}
& \sum_k \max_z \sum_e \left| P(k, e|y, z) - \frac{1}{2} P(e|y, z) \right| \\
= \ & \sum_k \max_z \sum_e P(e|y, z) \left| \sum_{\mathbf{w}} \left(\delta^k_{f(\mathbf{w})} - \frac{1}{2} \right) P(\mathbf{w}|y, e, z) \right| \\
= \ & \sum_k \max_z \sum_e P(e|z) \left| \sum_{\mathbf{w}} \left(\delta^k_{f(\mathbf{w})} - \frac{1}{2} \right) \left(\bigotimes_{i=1}^{N_d} \Gamma^{\mathbf{x}_i}_{w_i} \right) \cdot P(B|Y, e, z) \right|
\end{aligned} \tag{A.11}$$

where in the last equality we have used no-signaling through $P(e|y, z) = P(e|z)$ and the fact that the probability of obtaining the string of majorities $\mathbf{w}$ when inputting $y = (\mathbf{x}_1, \dots \mathbf{x}_{N_d}) \in \mathcal{X}^{N_d}$ can be written as

$$P(\mathbf{w}|y) = \left(\bigotimes_{i=1}^{N_d} \Gamma^{\mathbf{x}_i}_{w_i} \right) \cdot P(B|Y). \tag{A.12}$$

In what follows, the absolute value of vectors is understood to be componentwise. Bound (A.11) can be rewritten as

$$\begin{aligned}
&\sum_k \max_z \sum_e \left| P(k,e|y,z) - \frac{1}{2}P(e|y,z) \right| \\
\leq\ & \sum_k \max_z \sum_e P(e|z) \left| \sum_{\mathbf{w}} \left(\delta^k_{f(\mathbf{w})} - \frac{1}{2} \right) \bigotimes_{i=1}^{N_d} \Gamma^{\mathbf{x}_i}_{w_i} \right| \cdot P(B|Y,e,z) \\
=\ & \sum_k \max_z \left| \sum_{\mathbf{w}} \left(\delta^k_{f(\mathbf{w})} - \frac{1}{2} \right) \bigotimes_{i=1}^{N_d} \Gamma^{\mathbf{x}_i}_{w_i} \right| \cdot \left(\sum_e P(e|z) P(B|Y,e,z) \right) \\
=\ & \sum_k \left| \sum_{\mathbf{w}} \left(\delta^k_{f(\mathbf{w})} - \frac{1}{2} \right) \bigotimes_{i=1}^{N_d} \Gamma^{\mathbf{x}_i}_{w_i} \right| \cdot P(B|Y), \qquad \text{(A.13)}
\end{aligned}$$

where the inequality follows from the fact that all the components of the vector $P(B|Y,e,z)$ are positive and no-signaling has been used again through $P(B|Y,z) = P(B|Y)$ in the last equality. The bound applies to any function f and holds for any choice of vectors $\Lambda^{\mathbf{x}_i}_w$ in $\Gamma^{\mathbf{x}_i}_w$. In what follows, we compute this bound for a specific choice of these vectors and function f.

Take $\Lambda^{\mathbf{x}_i}_w$ to be equal to the vectors $\Lambda^{\mathbf{x}_0}_w$ in Lemma 18. These vectors then satisfy the bounds (A.26) and (A.35) in the same Lemma. Take f to be equal to the function whose existence is proven in Lemma 19. Note that the conditions needed for this Lemma to apply are satisfied because of bound (A.26) in Lemma 18, and because the free parameter $N_d \geq 130$ satisfies $\left(3\sqrt{N_d}\right)^{-1/N_d} \geq \gamma = 0.9732$. With this choice of f and $\Lambda^{\mathbf{x}_i}_w$, bound (A.13) becomes

$$\begin{aligned}
&\sum_k \max_z \sum_e \left| P(k,e|y,z) - \frac{1}{2}P(e|y,z) \right| \\
\leq\ & \sum_k 3\sqrt{N_d} \left(\bigotimes_{i=1}^{N_d} \Omega^{\mathbf{x}_i} \right) \cdot P(B|Y) \\
\leq\ & 6\sqrt{N_d}\,(\alpha C + \beta I)^{\otimes N_d} \cdot P(B|Y)\ , \qquad \text{(A.14)}
\end{aligned}$$

where we have used $\Omega^{\mathbf{x}_i} = \sqrt{(\Gamma^{\mathbf{x}_i}_0)^2 + (\Gamma^{\mathbf{x}_i}_1)^2}$, $\sum_k 3 = 6$, bound (A.26) in Lemma 18 and bound (A.35) in Lemma 19. □

A.1.2 Statement and proof of Lemma 17

In this section we prove Lemma 17. This Lemma bounds the Bell violation in the distillation block in terms of the probability of not aborting the protocol in step 4 and the number and size of the blocks, N_b and N_d.

Lemma 17. *Let $P(b_1, \dots b_{N_b}|y_1, \dots y_{N_b})$ be a $(5N_dN_b)$-partite non-signaling distribution, $y_1, \dots y_{N_b}$ and l the variables generated in steps 2 and 3 of the protocol, respectively, and α and β real numbers such that $0 < \alpha < 1 < \beta$; then*

$$(\alpha C + \beta I)^{\otimes N_d} \cdot P(\tilde{B}|\tilde{Y}, g = 1) \ \le \ \alpha^{N_d} + \frac{2\, N_b^{\log_2(1-\epsilon)}}{P(g=1)} \left(32\beta\epsilon^{-5}\right)^{N_d} \ . \quad \text{(A.15)}$$

Proof of Lemma 17. According to definition (see main text)

$$r[b,y] = \left\{ \begin{array}{ll} 1 & \text{if } \ I(\mathbf{a}_1, \mathbf{x}_1) = \cdots = I(\mathbf{a}_{N_d}, \mathbf{x}_{N_d}) = 0 \\ 0 & \text{otherwise} \end{array} \right. \quad \text{(A.16)}$$

we have $I(\mathbf{a}_i, \mathbf{x}_i) \le \delta^0_{r[b,y]}$ for all values of $b = (\mathbf{a}_1, \dots \mathbf{a}_{N_d})$ and $y = (\mathbf{x}_1, \dots \mathbf{x}_{N_d})$. This also implies $I(\mathbf{a}_i, \mathbf{x}_i) I(\mathbf{a}_j, \mathbf{x}_j) \le \delta^0_{r[b,y]}$ and so on. Due to the property $0 < \alpha < 1 < \beta$, one has that $(\alpha\, 2^{-5})^{N_d - i} \beta^i \le \beta^{N_d}$ for any $i = 1, \dots N_d$. All this in turn implies

$$\begin{aligned}
& \prod_{i=1}^{N_d} \left[\alpha\, 2^{-5} + \beta I_i\right] \\
= \quad & \left(\alpha\, 2^{-5}\right)^{N_d} + \left(\alpha\, 2^{-5}\right)^{N_d - 1} \beta \sum_i I_i + \left(\alpha\, 2^{-5}\right)^{N_d - 2} \beta^2 \sum_{i \ne j} I_i I_j + \cdots \\
\le \quad & \left(\alpha\, 2^{-5}\right)^{N_d} + \beta^{N_d} \left(\sum_i I_i + \sum_{i \ne j} I_i I_j + \cdots \right) \\
\le \quad & \left(\alpha\, 2^{-5}\right)^{N_d} + \beta^{N_d} \left(\sum_i \delta^0_{r[b,y]} + \sum_{i \ne j} \delta^0_{r[b,y]} + \cdots \right) \\
\le \quad & \left(\alpha\, 2^{-5}\right)^{N_d} + \beta^{N_d} \left(2^{N_d} - 1\right) \delta^0_{r[b,y]} \ \le \ \left(\alpha\, 2^{-5}\right)^{N_d} + \left(\beta\, 2\right)^{N_d} \delta^0_{r[b,y]} \ \text{(A.17)}
\end{aligned}$$

where $I_i = I(\mathbf{a}_i, \mathbf{x}_i)$. This implies that

$$
\begin{aligned}
&(\alpha C+\beta I)^{\otimes N_d}\cdot P(B|Y,g=1)\\
=\;&\sum_{\mathbf{a}_1,\dots\mathbf{a}_{N_d}}\;\sum_{\mathbf{x}_1,\dots\mathbf{x}_{N_d}}\;\prod_{i=1}^{N_d}\left[\alpha\,2^{-5}+\beta I(\mathbf{a}_i,\mathbf{x}_i)\right]P(\mathbf{a}_1,\dots\mathbf{a}_{N_d}|\mathbf{x}_1,\dots\mathbf{x}_{N_d},g=1)\\
\le\;&\sum_{b,y}\left[\left(\alpha\,2^{-5}\right)^{N_d}+(2\beta)^{N_d}\delta^0_{r[b,y]}\right]P(b|y,g=1)\\
=\;&\alpha^{N_d}\sum_{y}2^{-5N_d}+(2\beta)^{N_d}\sum_{y}P(r=0|y,g=1)\\
=\;&\alpha^{N_d}+(2\beta)^{N_d}\sum_{y}P(r=0|y,g=1)\\
=\;&\alpha^{N_d}+(2\beta)^{N_d}\sum_{y}\frac{P(r=0,y|g=1)}{P(y|g=1)}\;.
\end{aligned}
\tag{A.18}
$$

We can now bound $P(y|g=1)$ taking into account that y denotes a $5N_d$-bit string generated by the ϵ-source $\mathcal{S}$ that remains after step 2 in the protocol. Note that only half of the 32 possible 5-bit inputs $\mathbf{x}$ generated by the source belong to $\mathcal{X}$ and remain after step 2. Thus, $P((\mathbf{x}_1,\dots,\mathbf{x}_{N_d})\in\mathcal{X}^{N_d}|g=1)\le 16^{N_d}(1-\epsilon)^{5N_d}$, where we used (A.3). This, together with $P((\mathbf{x}_1,\dots,\mathbf{x}_{N_d})|g=1)\ge\epsilon^{5N_d}$ implies that

$$
P(y|g=1)\ge\left(\frac{\epsilon^5}{16(1-\epsilon)^5}\right)^{N_d}\;. \tag{A.19}
$$

Substituting this bound in (A.18), and summing over y, gives

$$
(\alpha C+\beta I)^{\otimes N_d}\cdot P(B|Y,g=1)\;\le\;\alpha^{N_d}+(2\beta)^{N_d}\left(\frac{16(1-\epsilon)^5}{\epsilon^5}\right)^{N_d}P(r=0|g=1)\;. \tag{A.20}
$$

In what follows we use the notation

$$
P(1_1,0_2,1_3,1_4,\dots)=P(r[b_1,y_1]=1,r[b_2,y_2]=0,r[b_3,y_3]=1,r[b_4,y_4]=1,\dots)\;.
$$

According to step 4 in the protocol described in the main text, it aborts ($g=0$) if there is at least a “not right” block ($r[b_j,y_j]=0$ for some $j\neq l$). While abortion also happens if there are more than one “not right” block, in what follows we lower-bound $P(g=0)$ by the probability that there is only one "not right" block:

$$\begin{aligned}
1 &\geq P(g=0) \\
&\geq \sum_{l=1}^{N_b} P(l) \sum_{l'=1,\, l'\neq l}^{N_b} P(1_1, \dots 1_{l-1}, 1_{l+1}, \dots 1_{l'-1}, 0_{l'}, 1_{l'+1}, \dots 1_{N_b}) \\
&\geq \sum_{l} P(l) \sum_{l'\neq l} P(1_1, \dots 1_{l-1}, 1_l, 1_{l+1}, \dots 1_{l'-1}, 0_{l'}, 1_{l'+1}, \dots 1_{N_b}) \\
&= \sum_{l'} \left[\textstyle\sum_{l\neq l'} P(l)\right] P(1_1, \dots 1_{l-1}, 1_l, 1_{l+1}, \dots 1_{l'-1}, 0_{l'}, 1_{l'+1}, \dots 1_{N_b}) \\
&= \sum_{l'} [1-P(l')]\, P(1_1, \dots 1_{l'-1}, 0_{l'}, 1_{l'+1}, \dots 1_{N_b}), \qquad \text{(A.21)}
\end{aligned}$$

where, when performing the sum over l, we have used that $P(1_1, \dots 1_{l-1}, 1_l, 1_{l+1}, \dots 1_{l'-1}, 0_{l'}, 1_{l'+1}, \dots 1_{N_b}) \equiv P(1_1, \dots 1_{l'-1}, 0_{l'}, 1_{l'+1}, \dots 1_{N_b})$ does not depend on l. Bound (A.3) implies

$$\frac{1-P(l)}{P(l)} \;\geq\; \frac{1-(1-\epsilon)^{\log_2 N_b}}{(1-\epsilon)^{\log_2 N_b}} \;=\; N_b^{\log_2 \frac{1}{1-\epsilon}} - 1 \;\geq\; \frac{N_b^{\log_2 \frac{1}{1-\epsilon}}}{2}\;, \qquad \text{(A.22)}$$

where the last inequality holds for sufficiently large N_b. Using this and (A.21), we obtain

$$\begin{aligned}
1 &\geq \frac{1}{2} \sum_{l'} N_b^{\log_2 \frac{1}{1-\epsilon}}\, P(l')\, P(1_1, \dots 1_{l'-1}, 0_{l'}, 1_{l'+1}, \dots 1_{N_b}) \\
&\geq \frac{1}{2} N_b^{\log_2 \frac{1}{1-\epsilon}}\, P(\tilde{r}=0, g=1)\;, \qquad \text{(A.23)}
\end{aligned}$$

where $\tilde{r} = r[b_l, y_l]$. This together with (A.20) implies

$$\begin{aligned}
&(\alpha C + \beta I)^{\otimes N_d} \cdot P(\tilde{B}|\tilde{Y}, g=1) \\
&\leq \alpha^{N_d} + (2\beta)^{N_d} \left(\frac{16(1-\epsilon)^5}{\epsilon^5}\right)^{N_d} P(\tilde{r}=0|g=1) \qquad \text{(A.24)} \\
&\leq \alpha^{N_d} + \frac{2}{P(g=1)} \left(\frac{32\beta(1-\epsilon)^5}{\epsilon^5}\right)^{N_d} N_b^{\log_2 (1-\epsilon)}\;, \qquad \text{(A.25)}
\end{aligned}$$

where, in the second inequality, Bayes rule was again invoked. Inequality (A.25), in turn, implies (A.15). □

A.1.3 Statement and proof of the additional Lemmas

Lemma 18. *For each* $\mathbf{x}_0 \in \mathcal{X}$ *there are three vectors* $\Lambda_0^{\mathbf{x}_0}, \Lambda_1^{\mathbf{x}_0}, \Lambda_2^{\mathbf{x}_0}$ *orthogonal to the non-signaling subspace such that for all* $w \in \{0,1\}$ *and* $\mathbf{a}, \mathbf{x} \in \{0,1\}^5$ *they satisfy*

$$\begin{aligned} & \sqrt{[M_0^{\mathbf{x}_0}(\mathbf{a},\mathbf{x}) + \Lambda_0^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})]^2 + [M_1^{\mathbf{x}_0}(\mathbf{a},\mathbf{x}) + \Lambda_1^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})]^2} \\ \leq\ & \alpha C(\mathbf{a},\mathbf{x}) + \beta I(\mathbf{a},\mathbf{x}) + \Lambda_2^{\mathbf{x}_0}(\mathbf{a},\mathbf{x}) \end{aligned} \tag{A.26}$$

and

$$\begin{aligned} & |M_w^{\mathbf{x}_0}(\mathbf{a},\mathbf{x}) + \Lambda_w^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})| \\ \leq\ & \gamma\sqrt{[M_0^{\mathbf{x}_0}(\mathbf{a},\mathbf{x}) + \Lambda_0^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})]^2 + [M_1^{\mathbf{x}_0}(\mathbf{a},\mathbf{x}) + \Lambda_1^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})]^2} \end{aligned} \tag{A.27}$$

where $\alpha = 0.8842$*,* $\beta = 1.260$ *and* $\gamma = 0.9732$*.*

Proof of Lemma 18. The proof of this lemma is numeric but rigorous. It is based on two linear-programming minimization problems, which are carried for each value of $\mathbf{x}_0 \in \mathcal{X}$. We have repeated this process for different values of γ, finding that $\gamma = 0.9732$ is roughly the smallest value for which the linear-programs described below are feasible.

The fact that the vectors $\Lambda_0^{\mathbf{x}_0}, \Lambda_1^{\mathbf{x}_0}, \Lambda_2^{\mathbf{x}_0}$ are orthogonal to the non-signaling subspace can be written as linear equalities

$$D \cdot \Lambda_w^{\mathbf{x}_0} = \mathbf{0} \tag{A.28}$$

for $w \in \{0,1,2\}$, where $\mathbf{0}$ is the zero vector and D is a matrix whose rows constitute a basis of non-signaling probability distributions. A geometrical interpretation of constraint (A.26) is that the point in the plane with coordinates $[M_0^{\mathbf{x}_0}(\mathbf{a},\mathbf{x}) + \Lambda_0^{\mathbf{x}_0}(\mathbf{a},\mathbf{x}), M_1^{\mathbf{x}_0}(\mathbf{a},\mathbf{x}) + \Lambda_1^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})] \in \mathbb{R}^2$ is inside a circle of radius $\alpha C(\mathbf{a},\mathbf{x}) + \beta I(\mathbf{a},\mathbf{x}) + \Lambda_2^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})$ centered at the origin. All points inside an octagon inscribed in this circle also satisfy constraint (A.26). The points of such an inscribed octagon are the ones satisfying the following set of linear constraints:

$$\begin{aligned} & [M_0^{\mathbf{x}_0}(\mathbf{a},\mathbf{x}) + \Lambda_0^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})]\,\eta\cos\theta + [M_1^{\mathbf{x}_0}(\mathbf{a},\mathbf{x}) + \Lambda_1^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})]\,\eta\sin\theta \\ \leq\ & \alpha C(\mathbf{a},\mathbf{x}) + \beta I(\mathbf{a},\mathbf{x}) + \Lambda_2^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})\ , \end{aligned} \tag{A.29}$$

for all $\theta \in \{\frac{\pi}{8}, \frac{3\pi}{8}, \frac{5\pi}{8}, \frac{7\pi}{8}, \frac{9\pi}{8}, \frac{11\pi}{8}, \frac{13\pi}{8}, \frac{15\pi}{8}\}$, where $\eta = (\cos\frac{\pi}{8})^{-1} \approx 1.082$. In other words, the eight conditions (A.29) imply constraint (A.26). From now on, we only consider these eight linear constraints (A.29). With a bit of algebra, one can see that inequality (A.27) is equivalent to the two almost linear inequalities,

$$\pm\left[M_w^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})+\Lambda_w^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})\right] \;\leq\; \sqrt{\frac{\gamma^2}{1-\gamma^2}}\,\left|M_{\bar{w}}^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})+\Lambda_{\bar{w}}^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})\right| \;, \tag{A.30}$$

for all $w \in \{0,1\}$, where $\bar{w} = 1-w$. Clearly, the problem is not linear because of the absolute values. The computation described in what follows constitutes a trick to make a good guess for the signs of the terms in the absolute value of (A.30), so that the problem can be made linear by adding extra constraints.

The first computational step consists of a linear-programming minimization of α subject to the constraints (A.28), (A.29), where the minimization is performed over the variables $\alpha, \beta, \Lambda_0^{\mathbf{x}_0}, \Lambda_1^{\mathbf{x}_0}, \Lambda_2^{\mathbf{x}_0}$. This step serves to guess the signs

$$\sigma_w(\mathbf{a},\mathbf{x}) \;=\; \mathrm{sign}[M_w^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})+\Lambda_w^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})] \;, \tag{A.31}$$

for all $w, \mathbf{a}, \mathbf{x}$, where the value of $\Lambda_w^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})$ corresponds to the solution of the above minimization. Once we have identified all these signs, we can write the inequalities (A.30) in a linear fashion:

$$\sigma_w(\mathbf{a},\mathbf{x})\,[M_w^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})+\Lambda_w^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})] \geq 0 \;, \tag{A.32}$$

$$\sigma_w(\mathbf{a},\mathbf{x})\,[M_w^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})+\Lambda_w^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})] \leq \sqrt{\frac{\gamma^2}{1-\gamma^2}}\,\sigma_{\bar{w}}(\mathbf{a},\mathbf{x})\,[M_{\bar{w}}^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})+\Lambda_{\bar{w}}^{\mathbf{x}_0}(\mathbf{a},\mathbf{x})] \;, \tag{A.33}$$

for all $w \in \{0,1\}$.

The second computational step consists of a linear-programming minimization of α subjected to the constraints (A.28), (A.29), (A.32), (A.33), over the variables $\alpha, \beta, \Lambda_0^{\mathbf{x}_0}, \Lambda_1^{\mathbf{x}_0}, \Lambda_2^{\mathbf{x}_0}$. Clearly, any solution to this problem is also a solution to the original formulation of the Lemma. The minimization was performed for any $\mathbf{x}_0 \in \mathcal{X}$ and the values of α, β turned out to be independent of $\mathbf{x}_0 \in \mathcal{X}$. These obtained numerical values are the ones appearing in the formulation of the Lemma. □

Note that Lemma 18 allows one to bound the predictability of maj($\mathbf{a}$) by a linear function of the 5-party Mermin violation. This can be seen by computing $\Gamma_w^{\mathbf{x}_0} \cdot P(\mathbf{A}|\mathbf{X})$ and applying the bounds in the Lemma. In principle, one expects this bound to exist, as the predictability is smaller than one at the point of maximal violation, as proven in Result 1 in the main text, and equal to one at the point of no violation. However, we were unable to find it. This is why we had to resort to the linear optimization technique given above, which moreover provides the bounds (A.26) and (A.27) necessary for the security proof.

Lemma 19. *Let N_d be a positive integer and let $\Gamma^i_w(\mathbf{a},\mathbf{x})$ be a given set of real coefficients such that for all $i \in \{1,\dots N_d\}$, $w \in \{0,1\}$ and $\mathbf{a},\mathbf{x} \in \{0,1\}^5$ they satisfy*

$$\left|\Gamma^i_w(\mathbf{a},\mathbf{x})\right| \leq \left(3\sqrt{N_d}\right)^{-1/N_d} \Omega_i(\mathbf{a},\mathbf{x}) \ , \tag{A.34}$$

where $\Omega_i(\mathbf{a},\mathbf{x}) = \sqrt{\Gamma^i_0(\mathbf{a},\mathbf{x})^2 + \Gamma^i_1(\mathbf{a},\mathbf{x})^2}$. There exists a function $f : \{0,1\}^{N_d} \to \{0,1\}$ such that for each sequence $(\mathbf{a}_1,\mathbf{x}_1),\dots(\mathbf{a}_{N_d},\mathbf{x}_{N_d})$ we have

$$\left|\sum_{\mathbf{w}} \left(\delta^k_{f(\mathbf{w})} - \frac{1}{2}\right) \prod_{i=1}^{N_d} \Gamma^i_{w_i}(\mathbf{a}_i,\mathbf{x}_i)\right| \ \leq \ 3\sqrt{N_d} \prod_{i=1}^{N_d} \Omega_i(\mathbf{a}_i,\mathbf{x}_i) \ , \tag{A.35}$$

where the sum runs over all $\mathbf{w} = (w_1,\dots w_{N_d}) \in \{0,1\}^{N_d}$.

Proof of Lemma 19. First, note that for a sequence $(\mathbf{a}_1,\mathbf{x}_1),\dots(\mathbf{a}_{N_d},\mathbf{x}_{N_d})$ for which there is at least one value of $i \in \{1,\dots N_d\}$ satisfying $\Gamma^i_0(\mathbf{a}_i,\mathbf{x}_i) = \Gamma^i_1(\mathbf{a}_i,\mathbf{x}_i) = 0$, both the left-hand side and the right-hand side of (A.35) are equal to zero, hence, inequality (A.35) is satisfied independently of the function f. Therefore, in what follows, we only consider sequences $(\mathbf{a}_1,\mathbf{x}_1),\dots(\mathbf{a}_{N_d},\mathbf{x}_{N_d})$ for which either $\Gamma^i_0(\mathbf{a}_i,\mathbf{x}_i) \neq 0$ or $\Gamma^i_1(\mathbf{a}_i,\mathbf{x}_i) \neq 0$, for all $i = 1,\dots N_d$. Or, equivalently, we consider sequences such that

$$\prod_{i=1}^{N_d} \Omega_i(\mathbf{a}_i,\mathbf{x}_i) > 0 \ . \tag{A.36}$$

The existence of the function f satisfying (A.35) for all such sequences is shown with a probabilistic argument. We consider the situation where f is picked from the set of all functions mapping $\{0,1\}^{N_d}$ to $\{0,1\}$ with uniform probability, and upper-bound the probability that the chosen function does not satisfy the constraint (A.35) for all k and all sequences $(\mathbf{a}_1,\mathbf{x}_1),\dots(\mathbf{a}_{N_d},\mathbf{x}_{N_d})$ satisfying (A.36). This upper bound is shown to be smaller than one. Therefore there must exist at least one function satisfying (A.35).

For each $\mathbf{w} \in \{0,1\}^{N_d}$ consider the random variable $F_{\mathbf{w}} = (\delta^0_{f(\mathbf{w})} - \frac{1}{2}) \in \{\frac{1}{2}, -\frac{1}{2}\}$, where f is picked from the set of all functions mapping $\{0,1\}^{N_d} \to \{0,1\}$ with uniform distribution. This is equivalent to saying that the 2^{N_d} random variables $\{F_{\mathbf{w}}\}_{\mathbf{w}}$ are independent and identically distributed according to $\Pr\{F_{\mathbf{w}} = \pm\frac{1}{2}\} = \frac{1}{2}$. For ease of notation, let us fix a sequence $(\mathbf{a}_1, \mathbf{x}_1), \dots (\mathbf{a}_{N_d}, \mathbf{x}_{N_d})$ satisfying (A.36) and use the short-hand notation $\Gamma^i_{w_i} = \Gamma^i_{w_i}(\mathbf{a}_i, \mathbf{x}_i)$.

We proceed using the same ideas as in the derivation of the exponential Chebyshev's Inequality. For any $\mu, \nu \geq 0$, we have

$$\begin{aligned}
& \Pr\left\{\sum_{\mathbf{w}} F_{\mathbf{w}} \prod_{i=1}^{N_d} \Gamma^i_{w_i} \geq \mu\right\} \\
= \;& \Pr\left\{\nu\left(-\mu + \sum_{\mathbf{w}} F_{\mathbf{w}} \prod_{i=1}^{N_d} \Gamma^i_{w_i}\right) \geq 0\right\} \\
= \;& \Pr\left\{\exp\left(-\nu\mu + \nu \sum_{\mathbf{w}} F_{\mathbf{w}} \prod_{i=1}^{N_d} \Gamma^i_{w_i}\right) \geq 1\right\} \\
\leq \;& \mathbb{E}\left[\exp\left(-\nu\mu + \nu \sum_{\mathbf{w}} F_{\mathbf{w}} \prod_{i=1}^{N_d} \Gamma^i_{w_i}\right)\right] && \text{(A.37)} \\
= \;& \mathbb{E}\left[\mathrm{e}^{-\nu\mu} \prod_{\mathbf{w}} \exp\left(\nu F_{\mathbf{w}} \prod_{i=1}^{N_d} \Gamma^i_{w_i}\right)\right] \\
= \;& \mathrm{e}^{-\nu\mu} \prod_{\mathbf{w}} \mathbb{E}\left[\exp\left(\nu F_{\mathbf{w}} \prod_{i=1}^{N_d} \Gamma^i_{w_i}\right)\right] && \text{(A.38)} \\
\leq \;& \mathrm{e}^{-\nu\mu} \prod_{\mathbf{w}} \mathbb{E}\left[1 + \nu F_{\mathbf{w}} \prod_{i=1}^{N_d} \Gamma^i_{w_i} + \left(\nu F_{\mathbf{w}} \prod_{i=1}^{N_d} \Gamma^i_{w_i}\right)^2\right]. && \text{(A.39)}
\end{aligned}$$

Here $\mathbb{E}$ stands for the average over all $F_{\mathbf{w}}$. In (A.37) we have used that any positive random variable X satisfies $\Pr\{X \geq 1\} \leq \mathbb{E}[X]$. In (A.38) we have used that the $\{F_{\mathbf{w}}\}_{\mathbf{w}}$ are independent. Finally, in (A.39) we have used that $\mathrm{e}^\eta \leq 1 + \eta + \eta^2$, which is only valid if $\eta \leq 1$. Therefore, we must show that

$$\left|\frac{\nu}{2} \prod_{i=1}^{N_d} \Gamma^i_{w_i}\right| \leq 1, \tag{A.40}$$

which is done below, when setting the value of ν. In what follows we use the chain of inequalities (A.39), the fact that $\mathbb{E}[F_{\mathbf{w}}] = 0$ and $\mathbb{E}[F_{\mathbf{w}}^2] = 1/4$, bound $1+\eta \leq \mathrm{e}^{\eta}$ for $\eta \geq 0$, and the definition $\Omega_i^2 = (\Gamma_0^i)^2 + (\Gamma_1^i)^2$:

$$\begin{aligned}
& \Pr\left\{\sum_{\mathbf{w}} F_{\mathbf{w}} \prod_{i=1}^{N_d} \Gamma_{w_i}^i \geq \mu\right\} \\
\leq\ & \mathrm{e}^{-\nu\mu} \prod_{\mathbf{w}} \left(1 + \mathbb{E}[F_{\mathbf{w}}]\,\nu \prod_{i=1}^{N_d} \Gamma_{w_i}^i + \mathbb{E}[F_{\mathbf{w}}^2]\,\nu^2 \prod_{i=1}^{N_d} \left(\Gamma_{w_i}^i\right)^2\right) \\
=\ & \mathrm{e}^{-\nu\mu} \prod_{\mathbf{w}} \left(1 + \frac{\nu^2}{4} \prod_{i=1}^{N_d} \left(\Gamma_{w_i}^i\right)^2\right) \\
\leq\ & \mathrm{e}^{-\nu\mu} \prod_{\mathbf{w}} \exp\left(\frac{\nu^2}{4} \prod_{i=1}^{N_d} \left(\Gamma_{w_i}^i\right)^2\right) \\
=\ & \exp\left(-\nu\mu + \sum_{\mathbf{w}} \frac{\nu^2}{4} \prod_{i=1}^{N_d} \left(\Gamma_{w_i}^i\right)^2\right) \\
=\ & \exp\left(-\nu\mu + \frac{\nu^2}{4} \prod_{i=1}^{N_d} \Omega_i^2\right)
\end{aligned} \tag{A.41}$$

In order to optimize this upper bound, we minimize the exponent over ν. This is done by differentiating with respect to ν and equating to zero, which gives

$$\nu = 2\,\mu \prod_{i=1}^{N_d} \Omega_i^{-2}\ . \tag{A.42}$$

Note that constraint (A.36) implies that the inverse of Ω_i exists. Since we assume $\mu \geq 0$, the initial assumption $\nu \geq 0$ is satisfied by the solution (A.42). By substituting (A.42) in (A.41) and rescaling the free parameter μ as

$$\tilde{\mu} = \frac{\mu}{\prod_{i=1}^{N_d} \Omega_i}\ , \tag{A.43}$$

we obtain

$$\Pr\left\{\sum_{\mathbf{w}} F_{\mathbf{w}} \prod_{i=1}^{N_d} \Gamma_{w_i}^i \geq \tilde{\mu} \prod_{i=1}^{N_d} \Omega_i\right\} \leq \mathrm{e}^{-\tilde{\mu}^2}\ , \tag{A.44}$$

for any $\tilde{\mu} \geq 0$ consistent with condition (A.40). We now choose $\tilde{\mu} = 3\sqrt{N_d}$, see Eq. (A.35), getting

$$\Pr\left\{\sum_{\mathbf{w}} F_{\mathbf{w}} \prod_{i=1}^{N_d} \Gamma_{w_i}^i \geq 3\sqrt{N_d} \prod_{i=1}^{N_d} \Omega_i\right\} \leq \mathrm{e}^{-9N_d}\ . \tag{A.45}$$

With this assignment, and using (A.42) and (A.43), condition (A.40), yet to be fulfilled, becomes

$$3\sqrt{N_d}\prod_{i=1}^{N_d}\frac{|\Gamma^i_{w_i}|}{\Omega_i} \leq 1\ , \tag{A.46}$$

which now holds because of the initial premise (A.34).

Bound (A.45) applies to each of the sequences $(\mathbf{a}_1, \mathbf{x}_1), \dots (\mathbf{a}_{N_d}, \mathbf{x}_{N_d})$ satisfying (A.36), and there are at most 4^{5N_d} of them. Hence, the probability that the random function f does not satisfy the bound

$$\sum_{\mathbf{w}} F_{\mathbf{w}} \prod_{i=1}^{N_d} \Gamma^i_{w_i} \geq 3\sqrt{N_d}\prod_{i=1}^{N_d}\Omega_i, \tag{A.47}$$

for at least one of such sequences, is at most $4^{5N_d}\mathrm{e}^{-9N_d}$, which is smaller than 1/2 for any value of N_d. A similar argument proves that the probability that the random function f does not satisfy the bound

$$\sum_{\mathbf{w}} F_{\mathbf{w}} \prod_{i=1}^{N_d} \Gamma^i_{w_i} \leq -3\sqrt{N_d}\prod_{i=1}^{N_d}\Omega_i, \tag{A.48}$$

for at least one sequence satisfying (A.36) is also smaller than 1/2. The lemma now easily follows from these two results. □

A.2 Final remarks

The main goal was to prove full randomness amplification. In this Appendix, we have shown how our protocol, based on quantum non-local correlations, achieves this task. Unfortunately, we are not able to provide an explicit description of the function $f : \{0,1\}^{N_d} \to \{0,1\}$ which maps the outcomes of the black boxes to the final random bit k; we merely show its existence. Such function may be obtained through an algorithm that searches over the set of all functions until it finds one satisfying (A.35). The problem with this method is that the set of all functions has size 2^{N_d}, which makes the search computationally costly. However, this problem can be fixed by noticing that the random choice of f in the proof of Lemma 19 can be restricted to a four-universal family of functions, with size polynomial in N_d. This observation will be developed in future work.

A more direct approach could consist of studying how the randomness in the measurement outcomes for correlations maximally violating the Mermin inequality increases with the number of parties. We solved linear optimization problems similar to those used in Result 1 in the main text, which showed that for 7 parties Eve's predictability is 2/3 for a function of 5 bits defined by $f(00000) = 0$, $f(01111) = 0$, $f(00111) = 0$ and $f(\mathbf{x}) = 1$ otherwise. Note that this value is lower than the earlier 3/4 and also that the function is different from the majority-vote. We were however unable to generalize these results for an arbitrary number of parties, which forced us to adopt a less direct approach. Note in fact that our protocol can be interpreted as a huge multipartite Bell test from which a random bit is extracted by classical processing of some of the measurement outcomes.

We conclude by stressing again that the reason why randomness amplification becomes possible using non-locality is because the randomness certification is achieved by a Bell inequality violation. There already exist several protocols, both in classical and quantum information theory, in which imperfect randomness is processed to generate perfect (or arbitrarily close to perfect) randomness. However, all these protocols, e.g. two-universal hashing or randomness extractors, always require additional good-quality randomness to perform such distillation. On the contrary, if the initial imperfect randomness has been certified by a Bell inequality violation, the distillation procedure can be done with a deterministic hash function (see [Mas09] or Lemma 16 above). This property makes Bell-certified randomness fundamentally different from any other form of randomness, and is the key for the success of our protocol.

Appendix B

Proof for vanishing classical randomness for arbitrary N

Here we prove the main theorem of Chapter 6. It is basically a generalization of the the proof for $N = 3$. We would like to prove that the function f defined in the main text, satisfies the property $P(f(\mathbf{a}) = 0|\mathbf{x_m}) \geq 1/2$ for any N-partite distribution (odd N) that maximally violates the Mermin inequality. As before we will express this condition in terms of correlators and use positivity conditions from the swapped input to complete the proof. One complication for larger N is that conditional on N, the function satisfies either of $P(f(\mathbf{a}) = 0|\mathbf{x_m}) \geq 1/2$ *or* $P(f(\mathbf{a}) = 1|\mathbf{x_m}) \geq 1/2$ for P maximally violating the N-partite Mermin inequality. For instance, it turns out that the former property is satisfied for $N = 3, 5$ and the latter for $N = 7, 9$ etc. Thus, we instead prove that $h_N.(P(f(\mathbf{a}) = 0|\mathbf{x_m}) - 1/2) \geq 0$ where the dependence on N in the equation is now contained in the new parameter h_N taking values $+1$ or -1 conditional on N.The exact expression for h_N is stated later as required.

A N-partite no-signalling probability distribution $P(\mathbf{a}|\mathbf{x})$ with inputs $\mathbf{x} \in \{0,1\}^N$ and outputs $\mathbf{a} \in \{+1,-1\}^N$ can be parameterized in terms of correlators as

$$P(\mathbf{a}|\mathbf{x}) = \frac{1}{2^N}(1 + \sum_{i=1}^{N} a_i \langle x_i \rangle + \sum_{i<j} a_i a_j \langle x_i x_j \rangle + \\ \sum_{i<j<k} a_i a_j a_k \langle x_i x_j x_k \rangle + \dots + a_1 a_2 \dots a_N \langle x_1 x_2 \dots x_N \rangle). \tag{B.1}$$

Restricting $P(\mathbf{a}|\mathbf{x})$ to those maximally violating the N-partite Mermin inequality is equivalent to requiring all correlators of input strings of odd parity to take their extremal values. Namely, we have,

$$\langle x_1 x_2 \dots x_N \rangle = (-1)^{(-1+\sum_{i=1}^{N} x_i)/2}, \tag{B.2}$$

for all N-point correlators satisfying $\sum_{i=0}^{N} x_i = 1 \mod 2$. For instance, $\langle 0, 0, \dots, 1\rangle = 1$ and similarly for all permutations. Also, $\langle 0, 0, \dots, 0, 1, 1, 1\rangle = -1$ as well as for for all permutations, etc. In the following we will use the notation $\langle . \rangle_k$ to denote a k-pt correlator. The input combination used to extract randomness from is a generalization of the tripartite case and denoted by $\mathbf{x_m} = (0, 0, \dots, 0, 1)$. The corresponding N-point correlator satisfies $\langle 00\dots01\rangle = 1$ for all N. The latter implies two useful relations:

1. Half the total outcomes vanish. In particular these are the terms for which the product of outcomes is -1 *i.e.* $P(\prod_{i=1}^{N} a_i = -1|\mathbf{x_m}) = 0$.
2. $\langle . \rangle_{N-k} = \langle . \rangle_k$ for all $1 \le k \le (N-1)/2$ where the correlators $\langle . \rangle_{N-k}$ and $\langle . \rangle_k$ are complementary in the input $\mathbf{x_m}$.

One can use these in Eqn. B.1 to express $P(\mathbf{a}|\mathbf{x_m})$ in terms of only the first $(N-1)/2$-pt. correlators as,

$$\begin{aligned} P(\mathbf{a}|\mathbf{x_m}) = \frac{1}{2^{N-1}}(1 &+ \sum a_i \langle A_i \rangle + \sum a_i a_j \langle A_i A_j \rangle \\ &+ \dots + \sum a_i a_j \cdots a_p \langle A_i A_j \cdots A_p \rangle_{(N-1)/2}). \end{aligned} \tag{B.3}$$

where $a_1 \cdot a_2 \cdot a_3 \dots a_N = +1$.

We denote by $\mathcal{S}^k$ the set of all $\langle . \rangle_k$ correlators that appear in the input $\mathbf{x_m}$.

B.1 Property to be satisfied by f

Now we can calculate the value of $h_N(P(f(\mathbf{a}) = 0|\mathbf{x}_m) - 1/2)$. Taking into account that f is permutationally symmetric, it will prove convenient to express our quantity of interest as

$$\begin{aligned} &(P(f(\mathbf{a}) = 0|\mathbf{x}_m) - 1/2) \cdot h_N = 2^{-(N-1)} \boldsymbol{\alpha}' \cdot \mathbf{c} \\ &\text{where} \\ &\boldsymbol{\alpha}' = h_N \cdot (\alpha_0 - 2^{N-2}, \alpha_1, \alpha_2, \dots, \alpha_{(N-1)/2}) \\ &\mathbf{c} = \left(1, \sum_{\mathcal{S}^1} \langle . \rangle_1, \sum_{\mathcal{S}^2} \langle . \rangle_2, \dots, \sum_{\mathcal{S}^{(N-1)/2}} \langle . \rangle_{(N-1)/2}\right) \end{aligned} \tag{B.4}$$

Recall that function f is such that $f(\mathbf{a}) = 0$ if $\sum_{i=1}^{N} a_i = 2 \mod 4$. By inspection, the explicit values of α_i can be written as

$$\alpha_i = \sum_{r=0}^{i} (-1)^r \binom{i}{r} \sum_{j\geq 0} \binom{n-i}{4j+2-r}. \tag{B.5}$$

For example, $\alpha_0 = \sum_{j\geq 0} \binom{n}{4j+2}$ as one would expect since α_0 simply counts the total number of terms $P(\mathbf{a}|\mathbf{x}_m)$ being summed to obtain $P(f(\mathbf{a}) = 0|\mathbf{x}_m)$.

Making use of the closed formula $\sum_{j\geq 0} \binom{n}{rj+a} = \frac{1}{r} \sum_{k=0}^{r-1} \omega^{-ka}(1+\omega^k)^n$ [BCK10], where $\omega = e^{i2\pi/r}$ is the r^{th} root of unity we can simplify the second sum appearing in Eqn. B.5. Finally the exact form of the phase h_N as a function of N is $h_N = \sqrt{2}\cos{(N+4)\pi/4}$. Putting all this together and performing the first sum in Eqn. (B.5) gives us,

$$\alpha_i' = 2^{\frac{N-3}{2}} \left(-2\cos\frac{(N-2i)\pi}{4} \cos\frac{(N+4)\pi}{4} \right) \tag{B.6}$$

Notice that the term in the parenthesis is a phase taking values in the set $\{+1, -1\}$ since N is odd while the amplitude is independent of N. Thus, we can simplify Eqn. (B.6) for even and odd values of i as,

$$\alpha_i' = \begin{cases} 2^{(N-3)/2}(-1)^{\frac{N-i}{2}} & i \text{ odd} \\ \\ 2^{(N-3)/2}(-1)^{\frac{i}{2}} & i \text{ even} \end{cases} \tag{B.7}$$

Thus, to prove that f possesses the property $(P(f(\mathbf{a}) = 0|\mathbf{x_m}) - 1/2)h_N \geq 0$ is equivalent to proving

$$\boldsymbol{\alpha}' \cdot \mathbf{c} \geq 0, \tag{B.8}$$

for $\mathbf{c}$ as defined in Eqn. (B.4) and for the values of $\boldsymbol{\alpha}'$ given by Eqn. (B.7). This is the task of the following section.

B.2 Swapped Input

We show that positivity conditions derived from the swapped input $\bar{\mathbf{x}} = (1, 1, \ldots, 1, 0)$ may be used to show $\boldsymbol{\alpha}' \cdot \mathbf{c} \geq 0$. In the following we will repeatedly use the Mermin conditions of Eqn. (B.2).

We start by summing the positivity conditions $P(+ + + \cdots + -|\bar{\mathbf{x}}) \geq 0$ and $P(- - - \cdots - +|\bar{\mathbf{x}}) \geq 0$. Using Eqn. (B.1), one can easily see that upon summing, all k-point correlators for odd k cancel while those for even k add up. Thus, we are left with an expression of the form

$$\begin{aligned} 1+\sum_{i<j} a_i a_j \langle A_i A_j\rangle + \sum_{i<j<k<l} a_i a_j a_k a_l \langle A_i A_j A_k A_l\rangle + \\ \cdots + \sum a_i \dots a_p \underbrace{\langle A_i \dots A_p\rangle}_{\text{(N−1)-pt. corr}} \geq 0. \end{aligned} \tag{B.9}$$

We would like to cast the above in a form that can be compared directly with Eqn. (B.4). Hence, we need to convert to an expression of the form,

$$\begin{aligned} (\beta_0, \beta_1, \dots \beta_{(N-1)/2}). \\ \Big(1, \sum \langle\cdot\rangle_1, \dots, \sum \langle\cdot\rangle_{(N-1)/2}\Big) \geq 0 \end{aligned} \tag{B.10}$$

Notice however that the correlators appearing in Eqn. (B.10) are locally swapped relative to the ones appearing in Eqn. (B.4). In order to compare the two expressions, we would like to convert all $2m$-pt. correlators (for $2m \geq (N-1)/2$) to $(N-2m)$-pt. correlators with the entries appearing in Eqn. (B.4). Moreover, we also need to convert $2m$-pt. correlators to the "correct" entries *i.e.* their locally swapped forms that appear in $\mathbf{c}$ in Eqn. B.4.

We show that this may be achieved by systematic use again of the Mermin conditions Eqn. (B.2).

B.2.1 Even-point correlators

Consider a $2k$-pt. correlator where $2k \leq (N-1)/2$. The correlators are of two forms and we show how they are transformed in each case:

- $\langle 11\dots1\rangle_{2k}$. The corresponding correct correlator entry is $\langle 00\dots0\rangle_{2k}$. We achieve the mapping by completing each to the corresponding Mermin full-correlators $\langle\underbrace{11\dots1}_{2k}\underbrace{100\dots0}_{(N-2k)}\rangle_N = (-1)^k$ and $\langle\underbrace{00\dots0}_{2k}\underbrace{100\dots0}_{(N-2k)}\rangle_N = (-1)^0 = 1$. From the signs, we have the relation, $\langle 11\dots1\rangle_{2k} = (-1)^k\langle 00\dots0\rangle_{2k}$
- $\langle 11\dots10\rangle_{2k}$, which we would like to map to $\langle 00\dots01\rangle_{2k}$. Using the same ideas we get $\langle\underbrace{11\dots10}_{2k}\underbrace{110\dots0}_{(N-2k)}\rangle_N = (-1)^k$ and $\langle\underbrace{00\dots01}_{2k}\underbrace{110\dots0}_{(N-2k)}\rangle_N = (-1)^1 = -1$. Thus, giving us the relation $\langle 11\dots10\rangle_{2k} = (-1)^{k+1}\langle 00\dots01\rangle_{2k}$.

Finally, a little thought shows that

$$\underbrace{a_1 a_2 \dots a_{2k}}_{\text{even}} \underbrace{\langle A_1 A_2 \dots A_{2k} \rangle}_{\text{wrong cor}} = (-1)^k \underbrace{\langle A_1 A_2 \dots A_{2k} \rangle}_{\text{correct cor}}$$

for correlators of either form discussed above on multiplying with their corresponding coefficients. Since we have finally converted to the correct entries of the correlator, we can read off β_i as the corresponding phase. Thus, $\beta_i = (-1)^{i/2}$ for even i.

B.2.2 Odd-point correlators

Consider now a $2k$-pt. correlator where $2k \geq (N-1)/2$. The correlators are again of two forms and may be transformed to the required $(N-2k)$-pt correlators in each case. The only difference from before is that the the two correlators are now complementary to each other in the swapped input. Since the details are similar, we simply state the final result $\beta_i = (-1)^{(N-i)/2}$ for odd i.

The final expression thus reads,

$$\beta_i = \begin{cases} (-1)^{\frac{N-i}{2}} & i \text{ odd} \\ \\ (-1)^{\frac{i}{2}} & i \text{ even} \end{cases} \tag{B.11}$$

These expressions exactly match the ones for α'_i (up to the constant factor) given in Eqn. B.7. Together with the correlators matching those in $\mathbf{c}$, it proves that f satisfies the required $\boldsymbol{\alpha}' \cdot \mathbf{c} \geq 0$ and hence the full result.

Bibliography

[ABG$^+$07] Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, Stefano Pironio, and Valerio Scarani. Device-independent security of quantum cryptography against collective attacks. *Phys. Rev. Lett.*, 98(23):230501–, 2007.

[ACA$^+$09] Leandro Aolita, Daniel Cavalcanti, Antonio Acín, Alejo Salles, Markus Tiersch, Andreas Buchleitner, and Fernando de Melo. Scalability of greenberger-horne-zeilinger and random-state entanglement in the presence of decoherence. *Phys. Rev. A*, 79(3):032322–, 2009.

[ACC$^+$08] L. Aolita, R. Chaves, D. Cavalcanti, A. Acín, and L. Davidovich. Scaling laws for the decay of multiqubit entanglement. *Phys. Rev. Lett.*, 100(8):080501–, 2008.

[ACC$^+$10] Leandro Aolita, Daniel Cavalcanti, Rafael Chaves, Chirag Dhara, Luiz Davidovich, and Antonio Acín. Noisy evolution of graph-state entanglement. *Phys. Rev. A*, 82(3):032317–, 2010.

[ADB05] H. Aschauer, W. Dür, and H.-J. Briegel. Multiparticle entanglement purification for two-colorable graph states. *Phys. Rev. A*, 71(1):012319–, 2005.

[ADR82] Alain Aspect, Jean Dalibard, and Gérard Roger. Experimental test of bell's inequalities using time- varying analyzers. *Phys. Rev. Lett.*, 49(25):1804–1807, 1982.

[AGCA12] Leandro Aolita, Rodrigo Gallego, Adán Cabello, and Antonio Acín. Fully nonlocal, monogamous, and random genuinely multipartite quantum correlations. *Phys. Rev. Lett.*, 108(10):100401–, 2012.

[AGMS04] Mario Leandro Aolita, Ignacio García-Mata, and Marcos Saraceno. Noise models for superoperators in the chord representation. *Phys. Rev. A*, 70(6):062301–, 2004.

[AMP12] Antonio Acín, Serge Massar, and Stefano Pironio. Randomness versus nonlocality and entanglement. *Phys. Rev. Lett.*, 108(10):100402–, 2012.

[BBD$^+$09] H. J. Briegel, D. E. Browne, W. Dur, R. Raussendorf, and M. Van den Nest. Measurement-based quantum computation. *Nat Phys*, pages 19–26, 2009.

[BBP$^+$96] Charles H. Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A. Smolin, and William K. Wootters. Purification of noisy entanglement and faithful teleportation via noisy channels. *Phys. Rev. Lett.*, 76(5):722–, 1996.

[BC90] S.L. Braunstein and C.M. Caves. Wringing out better bell inequalities. *Annals of Physics*, 202:22–, 1990.

[BCH$^+$02] Jonathan Barrett, Daniel Collins, Lucien Hardy, Adrian Kent, and Sandu Popescu. Quantum nonlocality, bell inequalities, and the memory loophole. *Phys. Rev. A*, 66(4):042111–, 2002.

[BCK10] Arthur T. Benjamin, Bob Chen, and Kimberly Kindred. Sums of evenly spaced binomial coefficients. *Mathematics Magazine*, 83:370–373, 2010.

[BDSW96] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54:3824–3851, 1996.

[Bel64] John Bell. On the einstein podolsky rosen paradox. *Physics*, 1:195–200, 1964.

[Bel66] John Bell. On the problem of hidden variables in quantum mechanics. *Rev. Mod. Phys.*, 38:447, 1966.

[Bel87] John Bell. *Speakable and Unspeakable in Quantum Mechanics.* Cambridge University Press, 1987.

[BG10] J. Barrett and N. Gisin. How much measurement independence is needed in order to demonstrate nonlocality? *Arxiv*, arXiv:1008.3612, 2010.

[BHK05] Jonathan Barrett, Lucien Hardy, and Adrian Kent. No signaling and quantum key distribution. *Phys. Rev. Lett.*, 95(1):010503–, 2005.

[BIWH96] J. J . Bollinger, Wayne M. Itano, D. J. Wineland, and D. J. Heinzen. Optimal frequency measurements with maximally correlated states. *Phys. Rev. A*, 54(6):R4649–R4652, 1996.

[BKP06] Jonathan Barrett, Adrian Kent, and Stefano Pironio. Maximally nonlocal and monogamous quantum correlations. *Phys. Rev. Lett.*, 97(17):170409–4, 2006.

[BLM$^+$05] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts. Nonlocal correlations as an information-theoretic resource. *Phys. Rev. A*, 71:022101, 2005.

[Boh35] Niels Bohr. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 48:696–702, 1935.

[Boh52] David Bohm. A suggested interpretation of the quantum theory in terms of "hidden" variables. i. *Phys. Rev.*, 85(2):166–179, 1952.

[BP05] Jonathan Barrett and Stefano Pironio. Popescu-rohrlich correlations as a unit of nonlocality. *Phys. Rev. Lett.*, 95(14):140401–, 2005.

[BR01] Hans J. Briegel and Robert Raussendorf. Persistent entanglement in arrays of interacting particles. *Phys. Rev. Lett.*, 86(5):910–, 2001.

[BVK98] S. Bose, V. Vedral, and P. L. Knight. Multiparticle generalization of entanglement swapping. *Phys. Rev. A*, 57(2):822–829, 1998.

[Cal04] C.S Calude. Algorithmic randomness, quantum physics and incompleteness. In *CDMTCS Research Reports CDMTCS-248.* 2004.

[Can01] R. Canetti. Universally composable security: a new paradigm for cryptographic protocols. *Proc. 42nd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 136–145, 2001.

[CCA$^+$09] Daniel Cavalcanti, Rafael Chaves, Leandro Aolita, Luiz Davidovich, and Antonio Acín. Open-system dynamics of graph-state entanglement. *Phys. Rev. Lett.*, 103(3):030502–, 2009.

[CG04] Daniel Collins and Nicolas Gisin. A relevant two qubit bell inequality inequivalent to the chsh inequality. *Journal of Physics A: Mathematical and General*, 37(5):1775, 2004.

[Cha87] Gregory Chaitin. *Algorithmic Information Theory.* Cambridge University Press, 1987.

[Che04] Zeqian Chen. Bell-klyshko inequalities to characterize maximally entangled states of n qubits. *Phys. Rev. Lett.*, 93:110403, 2004.

[CHSH69] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23(15):880–884, 1969.

[CL07] K. Chen and H.-K. Lo. Multi-partite quantum cryptographic protocols with noisy ghz states. *Quant. Inf. and Comp.*, 7:689, 2007.

[CLZ$^+$07] Kai Chen, Che-Ming Li, Qiang Zhang, Yu-Ao Chen, Alexander Goebel, Shuai Chen, Alois Mair, and Jian-Wei Pan. Experimental realization of one-way quantum computing with two-photon four-qubit cluster states. *Phys. Rev. Lett.*, 99(12):120503–, 2007.

[Col07] R. Colbeck. *Quantum and Relativistic Protocols for Secure Multi-Party Computation.* PhD thesis, Univ. of Cambridge, 2007.

[CR11] Roger Colbeck and Renato Renner. No extension of quantum theory can have improved predictive power. *Nat Commun*, 2:411–, 2011.

[CR12a] R. Colbeck and R. Renner. The completeness of quantum theory for predicting measurement outcomes. *arXiv*, arXiv:1208.4123 [quant-ph], 2012.

[CR12b] Roger Colbeck and Renato Renner. Free randomness can be amplified. *Nat Phys*, 8(6):450–454, 2012.

[CT91] T. M. Cover and J. A. Thomas. *Elements of information theory.* Wiley, New York, 1991.

[CW12] Eric G. Cavalcanti and Howard M. Wiseman. Bell nonlocality, signal locality and unpredictability (or what bohr could have told einstein at solvay had he known about bell experiments). *Foundations of Physics*, 42(10):1329–1338, 2012.

[DAB03] W. Dür, H. Aschauer, and H.-J. Briegel. Multiparticle entanglement purification for graph states. *Phys. Rev. Lett.*, 91(10):107903–, 2003.

[Dam05] W. V. Dam. Implausible consequences of superstrong nonlocality. *Arxiv*, arXiv:quant-ph/0501159, 2005.

[DB04] W. Dür and H.-J. Briegel. Stability of macroscopic entanglement under decoherence. *Phys. Rev. Lett.*, 92(18):180403–, 2004.

[DCA10] A. Ferraro A. García-Saez D. Cavalcanti, L. Aolita and A. Acín. Macroscopic bound entanglement in thermal graph states. *New Journal of Physics*, 12:025011, 2010.

[DCB05] W. Dür, J. Calsamiglia, and H.-J. Briegel. Multipartite secure state distribution. *Phys. Rev. A*, 71(4):042336–, 2005.

[Dha09] Chirag Dhara. Entanglement in the graph-state formalism. Master's thesis, UPC, UAB, UB and ICFO, Barcelona, 2009.

[DMPA11] Chirag Dhara, Lluis Masanes, Stefano Pironio, and Antonio Acín. Security of device-independent quantum key distribution protocols. In *Proc. of Theory of Quantum Computation, Communication and Cryptography 2011*, 2011.

[DP06a] Oscar Dahlsten and Martin B. Plenio. Exact entanglement probability distribution of bi-partite randomised stabilizer states. *Quant. Inf. Comp.*, 6:527, 2006.

[DP06b] E. D'Hondt and P. Panangaden. The computataional power of the w and ghz states. *Quant. Inf. Comp.*, 6:173–183, 2006.

[DPA12] Chirag Dhara, Giuseppe Prettico, and Antonio Acín. Maximal quantum randomness in bell tests. *arXiv*, arXiv:1211.0650v2 [quant-ph], 2012.

[Eke91] A. K. Ekert. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, 67:661, 1991.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, 1935.

[ER85] R. Eisberg and R. Resnick. *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles.* Wiley, 1985.

[FFW11] T. Franz, F. Furrer, and R. F. Werner. Extremal quantum correlations and cryptographic security. *Phys. Rev. Lett.*, 106:250502, 2011.

[FGS11] S. Fehr, R. Gelles, and C. Schaffner. Security and composability of randomness expansion from bell inequalities. *Arxiv*, arXiv:1111.6052, 2011.

[Fin82] A. Fine. Hidden variables, joint probability, and the bell inequalities. *Phys. Rev. Lett.*, 48:291, 1982.

[Fri12] Tobias Fritz. Polyhedral duality in bell scenarios with two binary observables. *J. Math. Phys.*, 53:072202, 2012.

[Fro81] M. Froissart. Constructive generalization of bell's inequalities. *Nuov. Cim. B*, 64:241–251, 1981.

[FSA$^+$12] Tobias Fritz, Belén Sainz, Remigiusz Augusiak, Jonatan Bohr Brask, Rafael Chaves, Anthony Leverrier, and Antonio Acín. Local orthogonality: a multipartite principle for correlations. *arXiv*, arXiv:1210.3018 [quant-ph], 2012.

[GBB08] O. Gühne, F. Bodoky, and M. Blaauboer. Multiparticle entanglement under the influence of decoherence. *Phys. Rev. A*, 78(6):060301–, 2008.

[GG08] Vlad Gheorghiu and Robert B. Griffiths. Separable operations on pure states. *Phys. Rev. A*, 78(2):020304–, 2008.

[GHZ89] D. M. Greenberger, M. A. Horne, and A. Zeilinger. *Bell's Theorem, Quantum Theory, and Conceptions of the Universe.* Kluwer, 1989.

[Gis10] Nicolas Gisin. Is realism compatible with true randomness? *arXiv*, arXiv:1012.2536 [quant-ph], 2010.

[GLB$^+$12] Basile Grandjean, Yeong-Cherng Liang, Jean-Daniel Bancal, Nicolas Brunner, and Nicolas Gisin. Bell inequalities for three systems and arbitrarily many measurement outcomes. *Phys. Rev. A*, 85(5):052113–, 2012.

[GLM04] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum-enhanced measurements: Beating the standard quantum limit. *Science*, 306(5700):1330–1336, 2004.

[GMdlT$^+$12] Rodrigo Gallego, Lluis Masanes, Gonzalo de la Torre, Chirag Dhara, Leandro Aolita, and Antonio Acín. Full randomness from arbitrarily deterministic events. *arXiv*, arXiv:1210.6514 [quant-ph], 2012.

[GMR$^+$12] Marissa Giustina, Alexandra Mech, Sven Ramelow, Bernhard Wittmann, Johannes Kofler, Adriana Lita, Brice Calkins, Thomas Gerrits, Sae Woo Nam, Rupert Ursin, and Anton Zeilinger. Bell violation with entangled photons, free of the fair-sampling assumption. *arXiv*, arXiv:1212.0533 [quant-ph], 2012.

[Gol01] Oded Goldreich. *Foundations of Cryptography - Vol. 1.* Cambridge University Press, 2001.

[Gol04] Oded Goldreich. *Foundations of Cryptography - Vol. 2.* Cambridge University Press, 2004.

[GRTZ02] Nicolas Gisin, Grégoire Ribordy, Wolfgang Tittel, and Hugo Zbinden. Quantum cryptography. *Rev. Mod. Phys.*, 74(1):145–, 2002.

[Hal10] Michael J. W. Hall. Local deterministic model of singlet state correlations based on relaxing measurement independence. *Phys. Rev. Lett.*, 105(25):250404–, 2010.

[Hal11] Michael J. W. Hall. Relaxed bell inequalities and kochen-specker theorems. *Phys. Rev. A*, 84(2):022102–, 2011.

[HBB99] Mark Hillery, Vladimír Bužek, and André Berthiaume. Quantum secret sharing. *Phys. Rev. A*, 59(3):1829–1834, 1999.

[HCLB11] M. J. Hoban, E. Campbell, K. Loukopoulos, and D. Browne. Non-adaptive measurement-based quantum computation and multi-party bell inequalities. *New J. Phys.*, 13:023014, 2011.

[HDB05] M. Hein, W. Dür, and H.-J. Briegel. Entanglement properties of multipartite entangled states under the influence of decoherence. *Phys. Rev. A*, 71(3):032350–, 2005.

[HDE$^+$06] M. Hein, W. Dür, J. Eisert, R. Raussendorf, M. Van den Nest, and H.-J. Briegel. Entanglement in graph states and its applications. *Arxiv*, arXiv:quant-ph/0602096:99, 2006.

[HEB04] M. Hein, J. Eisert, and H. J. Briegel. Multiparty entanglement in graph states. *Phys. Rev. A*, 69(6):062311–, 2004.

[HHH98] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Mixed-state entanglement and distillation: Is there a "bound" entanglement in nature? *Phys. Rev. Lett.*, 80:5239–5242, 1998.

[HHHH09] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81(2):865–942, 2009.

[Hob] Matty J. Hoban. Private communication.

[HW10] M. B. Plenio H. Wunderlich, S. Virmani. Highly-efficient estimation of entanglement measures for large experimentally created graph states via simple measurements. *New Journal of Physics*, 12:083026, 2010.

[JM05] Nick S. Jones and Lluis Masanes. Interconversion of nonlocal correlations. *Phys. Rev. A*, 72:052312, 2005.

[Kan98] Robert Kane. *The Significance of Free Will.* Oxford University Press, 1998.

[KHS$^+$12] D. E. Koh, M. J. W. Hall, Setiawan, J. E. Pope, C. Marletto, A. Kay, V. Scarani, and A. Ekert. The effects of reduced "free will" on bell-based randomness expansion. *Arxiv*, arXiv:1202.3571, 2012.

[KPB06] Johannes Kofler, Tomasz Paterek, and Caslav Brukner. Experimenter'ss freedom in bell's theorem and quantum cryptography. *Phys. Rev. A*, 73(2):022104–, 2006.

[KRS09] R. Konig, R. Renner, and C. Schaffner. The operational meaning of min- and max-entropy. *IEEE Trans Inf Theory*, 55(9):4337–4347, 2009.

[KSW+05] Nikolai Kiesel, Christian Schmid, Ulrich Weber, Géza Tóth, Otfried Gühne, Rupert Ursin, and Harald Weinfurter. Experimental analysis of a four-qubit photon cluster state. *Phys. Rev. Lett.*, 95(21):210502–, 2005.

[KTB11] Dirk Kroese, Thomas Taimre, and Zdravko Botev. *Handbook of Monte Carlo Methods*. Wiley, 2011.

[Lap40] P. S. Laplace. *A philosophical essay on probabilities*. 1840.

[LV08] Ming Li and Paul Vitanyi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer Verlag, 2008.

[LZG+07] Chao-Yang Lu, Xiao-Qi Zhou, Otfried Guhne, Wei-Bo Gao, Jin Zhang, Zhen-Sheng Yuan, Alexander Goebel, Tao Yang, and Jian-Wei Pan. Experimental entanglement of six photons in graph states. *Nat Phys*, 3(2):91–95, 2007.

[MAG06] Ll. Masanes, A. Acin, and N. Gisin. General properties of nonsignaling theories. *Phys. Rev. A*, 73(1):012112–, 2006.

[Mas09] Lluis Masanes. Universally composable privacy amplification from causality constraints. *Phys. Rev. Lett.*, 102(14):140501–, 2009.

[Mer90] N. David Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Phys. Rev. Lett.*, 65(15):1838–1840, 1990.

[ML66] P. Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619, 1966.

[NC00] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, Cambridge, 2000.

[NPA07] Miguel Navascués, Stefano Pironio, and Antonio Acín. Bounding the set of quantum correlations. *Phys. Rev. Lett.*, 98(1):010401–, 2007.

[NPA08] Miguel Navascués, Stefano Pironio, and Antonio Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics*, 10(7):073013–, 2008.

[NW10] Miguel Navascués and Harald Wunderlich. A glance beyond the quantum model. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 466(2115):881–890, 2010.

[PAM$^+$10a] S. Pironio, A. Acín, S. Massar, A. Boyer de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe. Random numbers certified by bell's theorem. *Nature*, 464:1021, 2010.

[PAM$^+$10b] S. Pironio, A. Acín, S. Massar, A. Boyer de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe. Random numbers certified by bell's theorems. *Nature*, 464(7291):1021–1024, 2010.

[PBS11] Stefano Pironio, Jean-Daniel Bancal, and Valerio Scarani. Extremal correlations of the tripartite no-signaling polytope. *Journal of Physics A: Mathematical and Theoretical*, 44(6):065303–, 2011.

[Per96] Asher Peres. Separability criterion for density matrices. *Phys. Rev. Lett.*, 77:1413–1415, 1996.

[Pir05] S. Pironio. Lifting bell inequalities. *Journal of Mathematical Physics*, 46:062112, 2005.

[PM11] S. Pironio and S. Massar. Security of practical private randomness generation. *Arxiv*, arXiv:1111.6056, 2011.

[PPK$^+$09] M. Pawlowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Zukowski. A new physical principle: Information causality. *arXiv*, arXiv:0905.2292v1 [quant-ph], 2009.

[PR94] Sandu Popescu and Daniel Rohrlich. Quantum nonlocality as an axiom. *Foundations of Physics*, 24(3):379–385, 1994.

[PV07] M. B. Plenio and S. Virmani. An introduction to entanglement measures. *Quant. Inf. Comp.*, 7:1–51, 2007.

[R6́1] Alfréd Rényi. On measures of entropy and information. In *Proc. Fourth Berkeley Symp. on Math. Statist. and Prob.*, volume 1, pages 547–561. Univ. of Calif. Press, 1961.

[Rao07] Anup Rao. *Randomness extractors for independent sources and applications.* PhD thesis, The University of Texas at Austin, 2007.

[RB01] Robert Raussendorf and Hans J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86(22):5188–5191, 2001.

[RBB03] Robert Raussendorf, Daniel E. Browne, and Hans J. Briegel. Measurement-based quantum computation on cluster states. *Phys. Rev. A*, 68(2):022312–, 2003.

[Ren05] Renato Renner. *Security of Quantum Key Distribution.* PhD thesis, Swiss Federal Institute of Technology, Zurich, 2005.

[Row01] D; Meyer V; Sackett CA; Itano WM; Monroe C; Wineland DJ. Rowe, MA; Kielpinski. Experimental violation of a bell's inequality with efficient detection. *Nature*, 409:791–794, 2001.

[Ś03] Cezary Śliwa. Symmetries of the bell correlation inequalities. *Physics Letters A*, 317(3,Äì4):165–168, 2003.

[Sch35] E. Schrödinger. Die gegenwä rtige situation in der quantenmechanik. *Naturwissenschaften*, 23:807, 1935.

[Sch44] Erwin Schrödinger. *What is Life?* Cambridge University Press, 1944.

[Sha48] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656, 1948.

[SK02] Christoph Simon and Julia Kempe. Robustness of multiparty entanglement. *Phys. Rev. A*, 65(5):052327–, 2002.

[SPG05] Tony Short, Sandu Popescu, and Nicolas Gisin. Entanglement swapping for generalized non-local correlations. *arXiv*, arXiv:quant-ph/0508120, 2005.

[SV86] Miklos Santha and Umesh V. Vazirani. Generating quasi-random sequences from semi-random sources. *J. Comput. Syst. Sci.*, 33(1):75–87, 1986.

[SW01] D. Schlingemann and R. F. Werner. Quantum error-correcting codes associated with graphs. *Phys. Rev. A*, 65(1):012308–, 2001.

[TB03] B. F. Toner and D. Bacon. Communication cost of simulating bell correlations. *Phys. Rev. Lett.*, 91(18):187904–, 2003.

[TBZG98] W. Tittel, J. Brendel, H. Zbinden, and N. Gisin. Violation of bell inequalities by photons more than 10 km apart. *Phys. Rev. Lett.*, 81(17):3563–3566, 1998.

[TPH$^+$06] M. S. Tame, M. Paternostro, C. Hadley, S. Bose, and M. S. Kim. Decoherence-based exploration of d-dimensional one-way quantum computation: Information transfer and basic gates. *Phys. Rev. A*, 74(4):042330–, 2006.

[Tsi87] B.S. Tsirelson. Quantum analogues of the bell inequalities. the case of two spatially separated domains. *Journal of Soviet Mathematics*, 36:557, 1987.

[Ver] Tamas Vertesi. private communication.

[Vid00] G. Vidal. Entanglement monotones. *J. Mod. Opt.*, 47(2-3):355–376, 2000.

[VPDMM08] Giuseppe Vallone, Enrico Pomarico, Francesco De Martini, and Paolo Mataloni. Active one-way quantum computation with two-photon four-qubit cluster states. *Phys. Rev. Lett.*, 100(16):160502–, 2008.

[VV12a] U. Vazirani and T. Vidick. Certifiable quantum dice: or, true random number generation secure against quantum adversaries. *Proceedings of the ACM Symposium on the Theory of Computing*, 2012.

[VV12b] Umesh V. Vazirani and Thomas Vidick. Fully device independent quantum key distribution. *arXiv*, arXiv:1210.1810 [quant-ph], 2012.

[VW02] G. Vidal and R. F. Werner. Computable measure of entanglement. *Phys. Rev. A*, 65(3):032314–, 2002.

[WCPG11] Michael M. Wolf, Toby S. Cubitt, and David Perez-Garcia. Are problems in quantum information theory (un)decidable? *arXiv*, arXiv:1111.5425 [quant-ph]:11, 2011.

[Wer89] Reinhard F. Werner. Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277–4281, 1989.

[WG03] Tzu-Chieh Wei and Paul M. Goldbart. Geometric measure of entanglement and applications to bipartite and multipartite quantum states. *Phys. Rev. A*, 68(4):042307–, 2003.

[WJS$^+$98] Gregor Weihs, Thomas Jennewein, Christoph Simon, Harald Weinfurter, and Anton Zeilinger. Violation of bell's inequality under strict einstein locality conditions. *Phys. Rev. Lett.*, 81(23):5039–5043, 1998.

[WRR$^+$05] P. Walther, K. J. Resch, T. Rudolph, E. Schenck, H. Weinfurter, V. Vedral, M. Aspelmeyer, and A. Zeilinger. Experimental one-way quantum computing. *Nature*, 434(7030):169–176, 2005.

[WW01] R. F. Werner and M. M. Wolf. All-multipartite bell-correlation inequalities for two dichotomic observables per site. *Phys. Rev. A*, 64:032112, 2001.

[ZHSL98] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein. Volume of the set of separable states. *Phys. Rev. A*, 58:883, 1998.

[ZZXS03] D. L. Zhou, B. Zeng, Z. Xu, and C. P. Sun. Quantum computation based on d-level cluster state. *Phys. Rev. A*, 68(6):062303–, 2003.