Temporal quantum correlations and hidden variable models

Dissertation
zur Erlangung des Grades eines Doktors
der Naturwissenschaften

vorgelegt von
M. Sc. Costantino Budroni

eingereicht bei der Naturwissenschaftlich-Technischen Fakultät
der Universität Siegen
Siegen 2014

Gutachter:

- Prof. Dr. Otfied Gühne
- Prof. Dr. Adán Cabello

Datum der mündlichen Prüfung: 19.09.2014

Prüfer:

- Prof. Dr. Otfied Gühne
- Prof. Dr. Adán Cabello
- Prof. Dr. Gregor Nickel
- Prof. Dr. Peter Buchholz

Abstract

This thesis is devoted to the investigation of the differences between the predictions of classical and quantum theory. More precisely, we shall analyze such differences starting from their consequences on quantities with a clear empirical meaning, such as probabilities, or relative frequencies, that can be directly observed in experiments.

Different kind of classical probability theories, or *hidden variable theories*, corresponding to different physical constraints imposed on the measurement scenario are discussed, namely, locality, noncontextuality and macroscopic realism. Each of these theories predicts bounds on the strength of correlations among different variables, and quantum mechanical predictions violate such bounds, thus revealing a stark contrast with our classical intuition.

Our work starts with the investigation of the set of classical probabilities by means of the correlation polytope approach, which provides a minimal and optimal set of bounds for classical correlations. In order to overcome some of the computational difficulties associated with it, we develop an alternative method that avoid the direct computation of the polytope and we apply it to Bell and noncontextuality scenarios showing its advantages both for analytical and numerical computations.

A different notion of optimality is then discussed for noncontextuality scenarios that provide a state-independent violation: Optimal expression are those maximizing the ratio between the quantum and the classical value. We show that this problem can be formulated as a linear program and solved with standard numerical techniques. Moreover, optimal inequalities for the cases analyzed are also proven to be part of the minimal set described above.

Subsequently, we provide a general method to analyze quantum correlations in the sequential measurement scenario, which allows us to compute the maximal correlations. Such a method has a direct application for computation of maximal quantum violations of Leggett-Garg inequalities, i.e., the bounds for correlation in a macroscopic realist theories, and it is relevant in the analysis of noncontextuality tests, where sequential measurements are usually employed.

Finally, we discuss a possible application of the above results for the construction of dimension witnesses, i.e., as a certification of the minimal dimension of the Hilbert spaces needed to explain the arising of certain quantum correlations.

Zusammenfassung

Diese Doktorarbeit befasst sich mit der Untersuchung der unterschiedlichen Vorhersagen von klassischen Theorien und Quantenmechanik. Dabei untersuchen wir insbesondere die Konsequenzen für diejenigen physikalischen Größen, denen eine klar definierte empirische Bedeutung zugeordnet werden kann. Dies sind zum Beispiel Wahrscheinlichkeiten oder relative Frequenzen, die in Experimenten direkt zu beobachten sind.

Es werden verschiedene klassische Wahrscheinlichkeitstheorien oder Theorien, die auf der Existenz versteckter Variablen basieren, diskutiert und besonders auf ihre Vorhersagen bezüglich der möglichen Stärke der Korrelationen zwischen verschiedenen Variablen eingegangen. Die klassischen Theorien machen dabei unterschiedliche physikalischen Annahmen wie Lokalität, Nichtkontextualität oder makroskopischer Realismus. Für jede dieser Theorien sagt die Quantenmeachnik stärkere Korrelationen voraus, die die klassischen Schranken verletzen und damit im Widerspruch zu unserer klassisch geprägten Intuition stehen.

Unsere Arbeit beginnt mit der Untersuchung der Menge von klassischen Wahrscheinlichkeiten mittels des Korrelations-Polytop-Verfahrens, welches einen minimalen und optimalen Satz an Grenzen für klassische Korrelationen liefert. Um einige der mit diesem Verfahren verbundenen rechnerischen Schwierigkeiten zu überwinden, entwickeln wir eine alternative Methode, die die direkte Berechnung des Polytops umgeht. Angewendet auf Bell- und Kontextualitätsszenarien zeigen wir die Vorteile unserer Methode, sowohl bezüglich analytischer, als auch numerischer Berechnungen.

Danach wird eine andere Möglichkeit betrachtet, Optimalität für Nichtkontextualitätsungleichungen zu definieren, die eine zustandsunabhängige Verletzung aufweisen: Optimale Ungleichungen sind solche, die das Verhältnis zwischen quantenmachanischem und klassischem Wert maximieren. Wir zeigen, dass dieses Problem als lineares Programm formuliert und mit standardmäßigen, numerischen Methoden gelöst werden kann. Darüber hinaus beweisen wir, dass die optimalen Ungleichungen für die betrachteten Fälle jene sind, die Teil des oben beschriebenen minimalen Satzes von Grenzflächen sind.

Anschließend stellen wir eine allgemeine Methode vor mit der man Quantenkorrelationen bei sequentiellen Messungen analysieren kann und die maximalen Korrelationen berechnen kann. Ein solches Verfahren hat als direkte Anwendung die Berechnung maximaler Quantenverletzung von Leggett-Garg Ungleichungen, d.h. der Grenzen für Korrelationen in Theorien, die auf der Annahme des makroskopischem Realismus basieren. Zudem ist diese Methode relevant in der analytischen Betrachtung von Kontextualitätstests, in denen üblicherweise sequentielle Messungen verwendet werden.

Abschließend diskutieren wir für die obigen Resultate Anwendungen bei der Konstruktion von Zeugenoperatoren für die Dimension von Quantensystemen. Damit ist es möglich, die minimale Dimension des Hilbertraums zu zertifizieren, die nötig ist, um das Auftreten von gegebenen Quantenkorrelationen zu erklären.

Contents

Introduction

The notion of probability is intimately related to the notion of uncertainty, and the latter arises in the description of physical systems at various levels and in different ways. In particular, the probabilistic structure arising in quantum mechanics (QM) has been recognized to be of a rather different kind with respect to its classical counterpart.

The formalism of classical mechanics is based on physical quantities which are assumed to have a clear empirical meaning (e.g., position and velocity of a point particle) and uncertainty only arises as a consequence of practical limitations (e.g., finite precision of measurement apparatuses) and can also be included in a systematic way in the description of a physical system, as it happens in statistical mechanics. More precisely, this is done by means of a probability measure, namely, an average over states with precisely determined values for all the physical quantities, describing the relative frequencies for appearance of such values on systems that have been subjected to the same *preparation*.

The situation is much more complex in QM. In fact, a problem of interpretation arises as a consequence of the lack of an apparent and unambiguous empirical meaning of the elements of formalism, e.g., self-adjoint operators and state vectors. An interpretation provides a set of rules allowing for a derivation of experimental predictions from the formalism, e.g., rules associating self-adjoint operators with experimental apparatuses, state vectors with preparation procedures, eigenvalues with outcomes of experiments. In particular, the predictions are restricted to well-defined experimental situations, avoiding some of the idealizations implicit in classical mechanics. An illustrative example of this attitude is given by Peres' claim that "quantum phenomena do not occur in a Hilbert space, they occur in a laboratory" [1].

Such a separation between formalism and interpretation is ultimately due to difficulties of a realistic description of quantum phenomena, i.e., a description in terms of an "actual state of affairs". Whether such an interpretative caution is justified, i.e., whether QM can be formulated in terms of classical probability theory, is still an open question, known as the *hidden variable* problem.

The necessity of such a completion of QM with additional hidden variables (HV), able to identify a dispersion free state and thus responsible for the randomness in the measurements results, was firstly advocated by Einstein, Podolsky, and Rosen (EPR) [2]. Their original argument was subsequently developed by Bell who showed that any completion of QM with a hidden variable theory satisfying the *locality assumption* of EPR, i.e., the existence of a finite speed at which causal influences can travel, must also obey some bounds on the possible strength of correlations, bounds nowadays known as *Bell inequalities* [3]. Such bounds are violated in QM, thus showing a contradiction between testable predictions of QM and local hidden variable theories. The possibility of such experimental tests distinguishes Bell's *no-go* theorem from many previous attempts (see discussion in [4]).

Another fundamental result on the impossibility of a HV completion of QM is Kochen-Specker theorem [5]. Here the assumption of locality is substituted by the more general notion of *non-*

contextuality, i.e., independence of the measurement context. To define precisely such a notion, we first need the notion of *compatible* measurements. Two or more measurements are said to be compatible if they can be performed jointly on the same system without disturbing each other. They can be performed jointly or sequentially, in any order, and must always reproduce the same result. In QM this notion correspond to the case of mutually commuting projective measurements. A measurement context is then defined as a set of compatible measurements. Noncontextuality clearly coincides with locality in the case of spacelike separated observables.

Analogously to Bell inequalities, noncontextuality inequalities can be defined as constraints on the possible ranges for classical probabilities under noncontextuality assumption. For a given measurement scenario, it has been proven [6, 7] that there exists a finite set of such constraints giving necessary and sufficient conditions for the existence of a noncontextual hidden variable theory.

A third class of hidden variable theory, is that of *macroscopic realist* theories, introduced by Leggett and Garg [8]. Here the usual assumption of realism, but this time applied only to *macroscopic* quantities, is combined with the assumption of *noninvasive measurability*, namely, the possibility of determining the value of such macroscopic quantities with an arbitrary small perturbation to their subsequent dynamics. The aim of the authors was to investigate and detect macroscopic coherence, i.e., the quantum superposition of macroscopically distinct states.

As can be intuitively seen already from the definition, the noninvasive measurability assumption resembles the assumption of context independence, in the sense that each measurement should not affect the measurement that may be performed subsequently. Here such an assumption is physically motivated by the macroscopic nature of such quantities and our experience with everyday objects. A similar analysis of such hidden variable models in terms of linear inequalities, the Leggett-Garg inequalities [8], is possible also in this case.

Besides the foundational problems, it is well known that the nonclassical features of quantum systems can be exploited to perform information processing tasks in a more efficient way. Two prominent examples are given by Shor's factoring algorithm [9] and quantum simulation algorithms [10]. It is therefore a fundamental question to identify and quantify the resources needed for quantum information processing tasks. For instance, in quantum communication tasks, the impossibility of a local hidden variable description, also named *nonlocality*, has been proven to be a fundamental resource for secure quantum key distribution among distant parties [11]. From the point of view of quantum computation, e.g., in the measurement-based quantum computation model, such a locality restriction is unnecessary. In this framework, *contextuality*, i.e., the impossibility of a noncontextual hidden variable model, has been proven to be a fundamental resource for computation [12, 13].

In this thesis we consider the problem of characterizing the ranges of values for probabilities in different kinds of hidden variable theories as well as in QM, with particular emphasis on the temporal scenario, and discuss possible applications of such result (e.g., dimension witnesses).

More in details, the thesis is structured as follows. In Chapter 1, we first recall the basic definitions and properties of the three main hidden variable theories mentioned above, namely, local, noncontextual, and macroscopic realist theories, and how their corresponding ranges for probabilities can be computed in the unified framework of Pitowsky's correlation polytope [7, 14]. Such a method provides a set of necessary and sufficient conditions for the existence of a hidden variable model, expressed in terms of a system of linear inequalities, as a solution of a geometrical problem known as the *hull problem*. Similarly, we discuss an approach for computing quantum bounds for the Bell and contextuality scenarios.

Despite the full generality of Pitowsky's method, and the existence of algorithms for solution of the hull problem, computing such a minimal set of conditions is a non trivial task. In fact, the complexity of the polytope grows rapidly with the number of measurement settings and outcomes (e.g., the number of vertices is exponential in the number of settings), and a direct computation has been performed only in simple cases.

To overcome this problem, in Chapter 2 we develop an alternative method based on the analysis of probability models for subsets of variables that are subsequently combined imposing some consistency conditions on their intersection. We then proceed to show the advantages of our method, both for analytical and numerical computation, in some non-trivial scenarios.

In Chapter 3, we analyze the case of contextuality scenarios where each state gives rise to the same violation of a given noncontextuality inequality, also known as state-independent contextuality (SIC) scenarios. Due to the high number of measurement settings involved in such scenario, the correlation polytope approach is usually inapplicable. We then define optimal inequalities, for a given SIC scenario, in terms of the maximal ratio between the quantum and the classical value, and show that such an optimization can be solved via linear programming, and thus efficiently with standard numerical techniques and with the optimality of the solution guaranteed. We discuss the most fundamental SIC scenarios and we found that the corresponding optimal inequalities are also facets of the associated correlation polytope.

In Chapter 4, we discuss the computation of quantum bounds for temporal correlations, namely, for sequences of quantum projective measurements. We provide a general method for computing such bounds that is based on semidefinite programming. Analogously to linear programming, such maximization procedure can be efficiently performed with standard numerical techniques and the optimality of the solution guaranteed.

In Chapter 5, we discuss the application of previous results as dimension witnesses, namely as a certification of the minimal dimension of the Hilbert space necessary to reproduce a set of correlations as a measurement on a quantum system. We provide dimension witnesses based on noncontextuality and Leggett-Garg inequalities for different dimensions and we discuss their robustness under noise and imperfections.

Finally, we present a discussion and outlook of the results of the thesis.

Chapter 1

Preliminary notions

1.1 Hidden variable theories

The necessity for a completion of quantum mechanics advocated by Einsten, Podolsky, and Rosen (EPR) [2] resulted in the *hidden variable* program, namely, the attempt to reinterpret quantum mechanical predictions as averages on a phase space, in a manner reminiscent of classical statistical mechanics. More precisely, the introduction of additional (hidden) variables allows all physical quantities to have a definite value (e.g., position and velocity of a point particle) and uncertainty only arises as a consequence of practical limitations (e.g., finite precision of measurement apparatuses). Uncertainty is thus included in the description of a physical system by means of a probability measure, namely, an average over states with precisely determined values for all the physical quantities, describing the relative frequencies for appearance of such values on systems that have been subjected to the same preparation procedure.

Mathematically, such a representation amounts to a classical probability theory defined by Kolomogorov's axioms [15] and described by a probability space (Λ, Σ, μ), where Λ is a set, Σ its σ-algebra of μ-measurable subsets, with Boolean operations $(\cap, \cup, ^c)$, and μ a normalized measure on Σ, i.e., $\mu(\Lambda) = 1$. In this framework, each point of the probability space determines the value of all the relevant physical quantities, which are described by classical random variables $f : \Lambda \to \sigma$, where σ is the set of their possible values. The randomness only arises as a consequence of practical limitations preventing us from preparing a state with zero uncertainty, i.e., a Dirac δ measure.

In quantum mechanics (QM), we know such a representation is possible for single observables, namely, the expectation value and the single-outcome probabilities can be computed via the Born rule and the spectral theorem. The simplest example is that of a discrete observable A with spectral decomposition $A = \sum_i \lambda_i P_i$, we have

$$\langle A \rangle_\psi = \langle \psi | A | \psi \rangle, \qquad \text{and} \qquad p_i \equiv Prob(A = \lambda_i) = \langle \psi | P_i | \psi \rangle \tag{1.1}$$

where, $p_i \geq 0$ and $\sum_i p_i = 1$, giving rise to the classical probabilistic interpretation of the numbers p_i. The same reasoning can be applied to a pair of commuting observables A, B, with $A = \sum_i \lambda_i P_i$ and $B = \sum_j \mu_j Q_j$ with probabilities defined as

$$p_{ij} \equiv Prob(A = \lambda_i, B = \mu_j) = \langle \psi | P_i Q_j | \psi \rangle, \tag{1.2}$$

where p_{ij} satisfy $p_{ij} \geq 0$ and $\sum_{ij} p_{ij} = 1$, and again can be interpreted as probabilities.

Such an interpretation in terms of classical probabilities actually holds in the general case, i.e., for arbitrary subalgebras of commuting observables with arbitrary spectra, as a consequence of the generality of the spectral theorem [16].

A natural questions is the following: Is it possible to embed such a collection of classical probabilities, arising from the spectral theorem, in a single "global" probability space, i.e., in a single hidden variable theory? There have been various attempts to introduce hidden variable theories [17, 18, 4, 5] and various impossibility proofs of their existence [19, 20, 21, 3, 5] . The difference resides in the conditions one assumes to be satisfied by a "reasonable" hidden variable theory. Two trivial constructions are possible:

(A) The global probability space is defined as a product of single-observable probability spaces.

(B) The global probability space is defined as a product of probability spaces associated with maximal contexts (maximal commuting subalgebras)

A possible objection to the model (A) is that no functional relation (between commuting observables) is satisfied, all observables are represented as independent variables (e.g., $\sigma_z \otimes \mathbb{1}$, $\mathbb{1} \otimes \sigma_x$ and $\sigma_z \otimes \sigma_x$ are represented by three independent random variables). On the other hand, the problem with model (B) is that, even though the functional relations are satisfied, each observable is represented by many random variables, one for each context. The problems associated with such constructions will be apparent in the following, in particular when we will associate a precise physical meaning to QM formal notion of *commuting* observable.

We shall analyse three different kinds of hidden variable theories and their associated impossibility proofs, namely

(*i*) local hidden variable theories (LHV),

(*ii*) non-contextual hidden variable theories (NCHV),

(*iii*) macrorealist hidden variable theories (MRHV).

Each of the above qualities associated with a hidden variable theory (local, non-contextual, macrorealist) refers to specific physical constraints that must be satisfied by the theory and that translate into conditional statistical independence relations among the classical random variables reproducing the measurement outcomes. We shall discuss in detail each of the above theories in the following sections.

Notice that the classical probabilistic description (Λ, Σ, μ), contains both an algebraic part, the algebra Σ, which encodes the logical relations between events, and a measure-theoretic part μ, which encodes the probabilistic structure. With the exception of one class of NCHV theories, where one wants the classical logical structure to reproduce the algebraic relations among commuting projectors, the most unconstrained classical logical structure can be assumed, thus reducing the problem only to the probabilistic description. More details can be found in Sect. 1.3.

In the analysis of the quantum versus classical predictions that follows, we shall consider only the case of projective measurement. The reason is that, while the generalization to positive-operator valued measure (POVM) can be easily done for locality scenarios, for the other two cases it is certainly problematic. In fact, for noncontextuality scenarios the correct notion of *compatible* measurements for POVM is still under debate [22] and for macrorealist theories the use of POVM explicitly contradicts the noninvasive measurability assumption [23, 24, 25].

1.2 Local hidden variables and Bell's theorem

Local hidden variable theories are classical theories that attempt to describe the statistics of measurements performed on distant systems. The locality condition, therefore, amounts to a statistical independence for the probabilities for outcomes on separated systems once conditioned on the hidden variable.

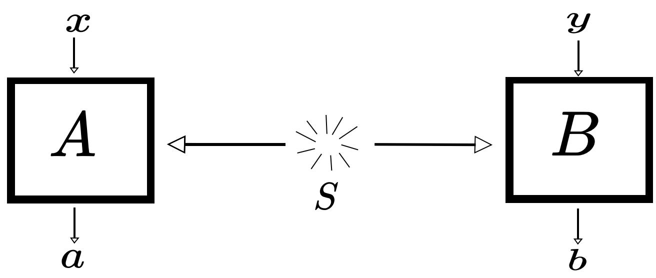


Figure 1.1: Schematic representation of a Bell measurement scenario. A source S produces two entangled particles that travel to two experiment sites, A and B. The two experimenters, Alice and Bob, can choose their measurement settings, respectively, x and y, and get an outcome, respectively, a and b.

1.2.1 Local hidden variables

To better introduce the main ideas involved, let us discuss the simplest measurement scenario. Consider two experimenters, Alice and Bob, performing measurements on two distant system. Alice can chose between two measurements, let us denote them as $x \in \{0,1\}$, with outcome $a \in \{-1,1\}$, and similarly for Bob, i.e., two measurements $y \in \{0,1\}$ with outcome $b \in \{-1,1\}$. A schematic representation of the measurement scenario is given in Fig. 1.1.

A LHV model is defined as probability distribution for the joint probabilities $P(ab|xy)$, i.e., the probability of getting the outcomes a and b given that Alice measure x and Bob measure y, of the form

$$P(ab|xy) = \int_\Lambda P(\lambda)P(a|x,\lambda)P(b|y,\lambda)\ d\lambda. \tag{1.3}$$

Once the hidden variable is known, the joint probability for outcomes a and b factorizes, implying that the two variables are independent once conditioned on λ. Without loss of generality, since λ can be chosen arbitrarily, all the indeterminacy left in the variables a and b after conditioning on λ can be removed by redefining the variable λ to include it. As a consequence, $P(a|x,\lambda)$ and $P(b|y,\lambda)$ can be seen as deterministic functions of λ.

Notice also that in Eq. (1.3) it is implicitly assumed that the probability distribution for the hidden variable λ does not depend on the choice of the measurement settings x, y, an assumption is usually called *free will.* The origin of the name can be easily understood by noticing that, by the definition of conditional probability,

$$P(\lambda|x,y) = P(\lambda) \text{ for all } \lambda, x, y \Longleftrightarrow P(x,y|\lambda) = P(x,y) \text{ for all } \lambda, x, y. \tag{1.4}$$

Eq. (1.4) implies that the experimenter is free to choose the to measure x and y, i.e., her choice is not "influenced" by the hidden variable λ. The free will assumption implies that λ, and consequently the measurement outcomes a and b, must be interpreted as statistical properties of the system that are (partially) revealed by the measurement apparatus.

To summarize, the assumptions defining a LHV theory are the following

R Realism: Observables represent well defined properties of the system, which are just revealed by the measurement process. In the probabilistic description of Eq. (1.3), they are fixed once the hidden variable λ is fixed.

Loc Locality: There is a maximum speed at which information propagates. Events in space-like separated regions cannot be in a relation of causal influence. In the probabilistic description of Eq. (1.3), probabilities for measurements on distant systems are statistically independent once conditioned on the hidden variable λ.

FW Free will: The experimenter is able to choose the measurement settings "freely", or, in simpler terms, the source of randomness used for the choice of the measurement settings is independent of the source of randomness of the system preparation. In the probabilistic description of Eq. (1.3), the probability distribution of the hidden variable λ is independent of the choice of the measurement settings.

1.2.2 CHSH inequality and Bell's theorem

The form (1.3) for the probability distribution allows us to compute bounds, usually expressed as linear inequalities, on the correlations among different outcomes. The most celebrated is the Clauser-Horne-Shimony-Holt (CHSH) inequality [26]. Let us denote by A_0, A_1 the $-1, 1$-valued measurement settings for Alice, and by B_0, B_1 the $-1, 1$-valued measurement settings for Bob, the CHSH inequality reads

$$\langle \mathcal{B} \rangle = \langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle \leq 2 \tag{1.5}$$

where $\langle A_i B_j \rangle$ denotes the correlation between A_i and B_j, i.e., the expectation value of the product of their outcomes.

Such a bound can be easily proven as follows. Let us define for Alice's measurements $f_{A_x}(\lambda) = P(+1|x,\lambda) - P(-1|x,\lambda)$, and similarly f_{B_y} for Bob's measurement. The functions f_{A_x}, f_{B_y} are, therefore, deterministic functions of λ that fix the measurement outcomes ± 1 (note that $P(+1|x,\lambda) + P(-1|x,\lambda) = 1$, and similarly for Bob). Eq. (1.5), can therefore be rewritten using Eq. (1.3) as

$$\begin{aligned} \langle A_0 B_0 \rangle + \langle A_0 B1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle &= \int_\Lambda P(\lambda) f_{A_0}(\lambda) f_{B_0}(\lambda)\ d\lambda + \int_\Lambda P(\lambda) f_{A_0}(\lambda) f_{B_1}(\lambda) d\lambda \\ &+ \int_\Lambda P(\lambda) f_{A_1}(\lambda) f_{B_0}(\lambda)\ d\lambda - \int_\Lambda P(\lambda) f_{A_1}(\lambda) f_{B_1}(\lambda) d\lambda \\ = \int_\Lambda P(\lambda) &\left[f_{A_0}(\lambda) f_{B_0}(\lambda) + f_{A_0}(\lambda) f_{B_1}(\lambda) + f_{A_1}(\lambda) f_{B_0}(\lambda) - f_{A_1}(\lambda) f_{B_1}(\lambda) \right]\ d\lambda \\ = \int_\Lambda P(\lambda) &\left[f_{A_0}(\lambda)(f_{B_0}(\lambda) + f_{B_1}(\lambda)) + f_{A_1}(\lambda)(f_{B_0}(\lambda) - f_{B_1}(\lambda)) \right] d\lambda \\ \leq \int_\Lambda P(\lambda) \max_\lambda \{ f_{A_0}(\lambda)(f_{B_0}(\lambda) + f_{B_1}(\lambda)) &+ f_{A_1}(\lambda)(f_{B_0}(\lambda) - f_{B_1}(\lambda)) \} d\lambda = 2 \int_\Lambda P(\lambda)\ d\lambda = 2 \end{aligned} \tag{1.6}$$

We can finally state the following

Theorem (Bell 1964). *No local hidden variable theory can reproduce all the predictions of quantum mechanics.*

Proof. Since Eq. (1.5) has been derived from the assumption of a local hidden variable theory, it is sufficient to provide some quantum mechanical correlations violating the bound.

Consider two spin-1/2 particles in the singlet state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle), \tag{1.7}$$

where $|0\rangle, |1\rangle$ denote the eigenstate of σ_z. By defining Alice's observables $A_0 = \sigma_z^{(1)}$ and $A_1 = \sigma_x^{(1)}$, where the superscript $^{(1)}$ denotes the action of the operator on the first particle, and Bob's observables $B_0 = \frac{1}{\sqrt{2}}(\sigma_z^{(2)} + \sigma_x^{(2)})$ and $B_1 = \frac{1}{\sqrt{2}}(\sigma_z^{(2)} - \sigma_x^{(2)})$, we obtain

$$\langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle = \frac{4}{\sqrt{2}} = 2\sqrt{2} > 2 \tag{1.8}$$

□

1.2.3 Experimental tests

The experimental progress in quantum optics during the 1960s, in particular the possibility of created pairs of photons entangled in polarization using atomic cascades, allowed for the first tests of Bell inequalities. In 1972, three year after Clauser-Horne-Shimony-Holt original proposal [26], Freedman and Clauser performed the first test and reported a violation of the CHSH inequality by six standard deviations [27].

Freedman and Clauser's experiment was followed by others [28, 29, 30] that share all the same problem: The experiments were performed with a static setup in which the polarized were held fixed. In this scenario, one can design a local hidden variable model where the detector on one side is "aware" of the measurement setting chosen on the other side (see the discussion on the free will assumption above). This possibility, preventing a definite answer to the LHV problem, has been named *locality loophole*.

To overcome this problem, Aspect *et al.* [31] introduced time-varying polarization analyzers in the experiment. With this setup, the settings were changed during the flight of the particle in such a way that the change of orientation on one side and the detection event on the other side were separated by a spacelike interval. This, together with the use of independent source of randomness for the change of the settings, justifies the free will assumption and close the locality loophole.

All the above experiments, however, were subjected to the *detection loophole*, namely, the possibility of a local hidden variable model explaining the observed correlations in terms of the statistics of the undetected events. More precisely, given the low efficiency of photon detectors (typically around 10%-20%), one can refute local hidden variable theories only by assuming that the fraction of detected events is a valid representative of the whole sample (the so-called *fair sampling* assumption), or, equivalently, that the probability of detecting is independent of choice of the measurement settings.

The detection loophole in Bell experiments has been first closed by Rowe *et al.* with entangled trapped ions [32], however, such an experiment was still subjected to the locality loophole. Recently, by using highly efficient photon detectors, the detection loophole has been closed in a photon experiment [33], thus showing that photons can, in principle, allow for a loophole-free Bell test, albeit such a test has not been performed yet.

1.3 Noncontextual hidden variables and Kochen-Specker theorem

Bell's theorem strongly constraints the interpretation of measurements as revealing preexisting properties of physical systems. A natural question is whether such a behaviour of quantum correlations appears also in more general measurement scenario, where measurements are not necessarily performed on separated systems. As previously discussed, QM allows joint measurements also for commuting, or *compatible*, observables and the corresponding predictions are described by a classical probability theory, but a much stronger property holds: Commuting measurements can be performed in sequence in any order and repeated multiple times, and the outcomes of each measurement are confirmed by the subsequent ones. This phenomenon suggests the idea that compatible measurements *do not disturb* each other and that each measurement apparatus should behave in the same way, independently of which other compatible measurements are performed together.

We already know, from Bell's theorem, that despite such properties a description in terms of noncontextual hidden variable is, in general, impossible. However, such an approach allows to investigate new phenomena arising from single systems, with potential new applications [12, 13].

In mathematical terms, QM predictions for each set of compatible observables $\mathcal{C} = \{B_1, \dots, B_k\}$ have a classical probabilistic representation given by $((\Lambda_{\mathcal{C}}, \Sigma_{\mathcal{C}}, \mu_{\mathcal{C}}), f_{B_1}, \dots, f_{B_k})$, where the functions $f_{B_i} : \Lambda_{\mathcal{C}} \to \sigma(B_i)$, with $\sigma(B_i)$ the spectrum of B_i, are classical random variables reproducing the expectation values for $\{B_1, \dots, B_k\}$, namely,

$$\langle B_1 \dots B_k \rangle = \int_\Lambda f_{B_1}(\lambda) \dots f_{B_k}(\lambda) d\mu(\lambda). \tag{1.9}$$

The whole set of QM predictions can be therefore seen as a collection of classical probabilities, one for each measurement context. For the sake of simplicity, and since our analysis will always involve only a finite number of events (e.g., a finite number of measurements and outcomes), it is sufficient to take a finite set Λ and Σ the finite Boolean algebra of its subsets.

Equation (1.9) resembles Eqs.(1.3),(1.6), in fact, a similar formal definitions can be given for NCHV in terms of assumptions **R**, **FW**, and **NC** (noncontextuality) substituting **LOC**, as shown below. However, this approach assumes an unconstrained logical structure for the HV theory, whereas, historically, quantum contextuality was introduced by Kochen and Specker [5] as the impossibility of the embedding of the logical structure, i.e., Boolean algebras, of subsets of commuting projectors into a single global logical structure.

We shall first discuss the unconstrained approach to the problem, the one followed by, e.g., Klyachko, Can, Binicioğlu, and Shumovsky (KCBS) [34], which is simpler to introduce in analogy with Bell's approach. Then, the original Kochen and Specker problem and the relation among the two approaches will be discussed. Finally, we shall discuss a new phenomenon, absent in Bell scenarios, which is that of state-independent contextuality (SIC).

1.3.1 Noncontextual hidden variable theories

In the case in which one assumes no constraint on the logical structure of the HV theory, a definition of NCHV theory similar to the one presented above for LHV theory can be given as follows

R Realism: Observables represent well defined properties of the system, which are just revealed by the measurement process.

NC Noncontextuality: The value of an observable is independent of the measurement context, compatible measurements cannot be in a relation of causal influence.

FW Free will: The experimenter is able to choose the measurement settings freely, i.e., the probability distribution of the hidden variable is independent of the choice of the measurement settings.

In the above framework, (KCBS) [34], proposed the following inequality

$$\langle A_0 A_1 \rangle + \langle A_1 A_2 \rangle + \langle A_2 A_3 \rangle + \langle A_3 A_4 \rangle + \langle A_4 A_0 \rangle \geq -3 \tag{1.10}$$

where A_i are measurements with outcomes -1 and 1, and the measurements in the same mean value $\langle \ \rangle$ are compatible, i.e., are represented in quantum mechanics by commuting operators. The

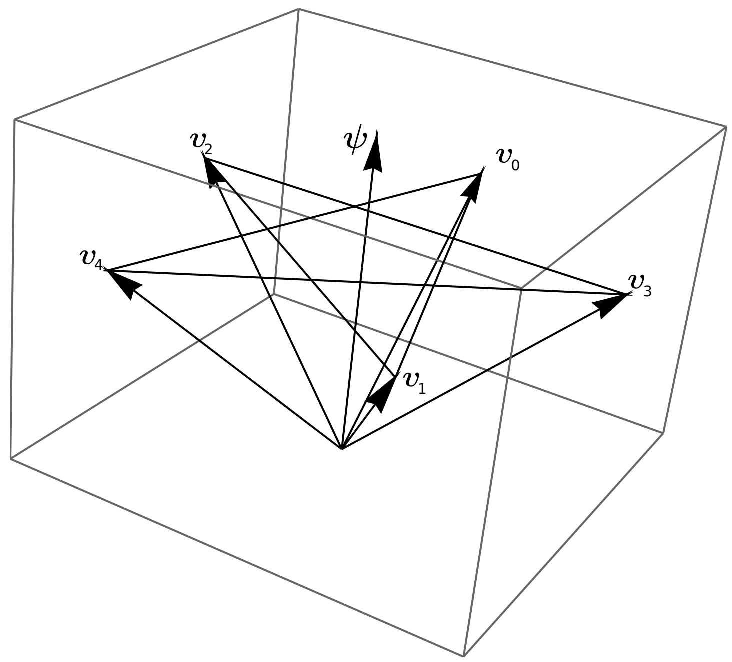


Figure 1.2: The set of vectors v_j giving the dichotomic observables $A_j = 2|v_j\rangle\langle v_j| - \mathbb{1}$ form a pentagram, with orthogonal vectors connected by an edge, and the state ψ is directed along its symmetry axis.

classical bound 3 can be proven, in analogy with the CHSH case above, just by trying all possible ± 1 noncontextual assignments to the observables A_i.

As opposed to Bell inequalities, there is no bipartition of the set of observables such that every observable in one part is compatible with every observable of the other. Consequently, Eq. (1.10) cannot be interpreted as a Bell inequality: The measurements must be performed on a single system.

On a three-level system Eq. (1.10) can be violated up to $5 - 4\sqrt{5} \approx -3.94$ with state $|\psi\rangle = (1, 0, 0)$ and measurement settings $A_j = 2|v_j\rangle\langle v_j| - \mathbb{1}$ as depicted in Fig. 1.2, namely, $|v_j\rangle = (\cos\theta, \sin\theta\cos[j\pi 4/5], \sin\theta\sin[j\pi 4/5])$ with $\cos^2\theta = \cos(\pi/5)/(1 + \cos(\pi/5))$.

1.3.2 Kochen and Specker's original problem

Kochen and Specker's original approach [5] focused on a more strict notion of NCHV. More precisely, it focused on reproducing also the state-independent predictions of QM, namely, those given by functional relations between commuting quantum observables.

As opposed to Birkhoff and von Neumann's approach to *quantum logic* [35], but rather following the same approach as Gleason [21], Kochen and Specker discussed the possibility of reproducing just the logical relations between compatible measurements, since such relations can be tested in joint measurement scenario and have a clear experimental meaning. A discussion of this point can be found in [4].

In mathematical terms, the above notion of NCHV is captured by Kochen and Specker's definition of *partial Boolean algebra* [5], and its subsequent extension to include probabilistic predictions [36]. Without loss of generality, we can consider only prediction for projectors, since the outcome probabilities for any observable can be recovered from those of its spectral decomposition. We refer to the definitions given in [36].

A *partial Boolean algebra (PBA)* is a set X together with a non-empty family $\mathcal{F}$ of Boolean algebras, $\mathcal{F} \equiv \{\mathfrak{B}_i\}_{i\in I}$, such that $\bigcup_i \mathfrak{B}_i = X$, that satisfy

(P_1) for every $\mathfrak{B}_i, \mathfrak{B}_j \in \mathcal{F}$, $\mathfrak{B}_i \cap \mathfrak{B}_j \in \mathcal{F}$ and the Boolean operations $(\cap_i, \cup_i, {}^{c_i})$, $(\cap_j, \cup_j, {}^{c_j})$ of $\mathfrak{B}_i$ and $\mathfrak{B}_j$ coincide on it.

Without loss of generality we can also assume the property

(P_2) for all $\mathfrak{B}_i \in \mathcal{F}$, each Boolean subalgebra of $\mathfrak{B}_i$ belongs to $\mathcal{F}$.

By (P_1), Boolean operations, when defined, are unique and will be denoted by $(\cap, \cup, {}^c)$; we shall denote a partial Boolean algebra by $(X, \{\mathfrak{B}_i\}_{i\in I})$, or simply by $\{\mathfrak{B}_i\}_{i\in I}$. In the following we shall consider only *finite* partial Boolean algebras.

Given a partial Boolean algebra $(X, \{\mathfrak{B}_i\})$, a *state* is defined as a map $f : X \longrightarrow [0,1]$, such that $f_{|\mathfrak{B}_i}$ is a normalized measure on the Boolean algebra $\mathfrak{B}_i$ for all i. Equivalently, a state is given by a collection of *compatible probability measures* $\{\mu_i\}$, i.e., measures coinciding on intersections of Boolean algebras, one for each $\mathfrak{B}_i$.

A *partial probability theory (PPT)* is a pair $((X, \{\mathfrak{B}_i\}); f)$, where $(X, \{\mathfrak{B}_i\})$ is a partial Boolean algebra and f is a state defined on it. Equivalently, a partial probability theory can be denoted with $((X, \{\mathfrak{B}_i\}); \{\mu_i\})$, where $\mu_i = f_{|\mathfrak{B}_i}$, or simply by $(\{\mathfrak{B}_i\}; \{\mu_i\})$.

So far, such a definition just constrain PPTs to behave as classical probabilities when restricted to contexts, and to have a noncontextual definition of their elements. Such a definition is basically the same as that of *nonsignalling theories* for Bell scenario [37], or, *nondisturbing* for noncontextuality scenario [38]. In their original formulation Kochen and Specker [39, 5] proposed the following additional property as a definition of PBA

(P_S) if $A_1, \ldots, A_n$ are elements of X such that any two of them belong to a common algebra $\mathfrak{B}_i$, then there is a $\mathfrak{B}_k \in \mathcal{F}$ such that $A_1, \ldots, A_n \in \mathfrak{B}_k$.

In other words, if n elements are mutually compatible, then they are also globally compatible. Such a property, afterwards named *Specker's principle* [40], has been shown to play a fundamental role in ruling out possible post-quantum theories [41, 43, 44].

It can be easily checked that the above properties are satisfied by the set of all orthogonal projections in a Hilbert space of arbitrary dimension, with Boolean operations defined by

$$P \cap Q \equiv PQ, \quad P \cup Q \equiv P + Q - PQ, \quad P^c \equiv 1 - P, \tag{1.11}$$

for all pairs P, Q of commuting projections. If one considers a finite set of projections, the result of the iteration of the above Boolean operations (on commuting projections) is still a finite set and a partial Boolean algebra.

Moreover, given a set of projections, the corresponding predictions given by a QM state define a PPT on the generated PBA. In fact, given a PBA of projections on a Hilbert space $\mathcal{H}$, by the spectral theorem, a quantum mechanical state ψ defines a state f_ψ on it, given by $f_\psi(P) = (\psi, P\psi)$. The generalization to density matrices is obvious.

In this framework, given a partial probability theory $(\{\mathfrak{B}_i\}; \{\mu_i\})$, a NCHV theory extending it is given by a Boolean algebra $\mathfrak{B}$ together with a normalized measure μ such that $\mathfrak{B}_i$ is a subalgebra of $\mathfrak{B}$ for all i, and $\mu_i = \mu_{|\mathfrak{B}_i}$.

The above implies that such an embedding must not only reproduce the probability structure of QM predictions, i.e., the measure μ, but also the logical structure, i.e., the Boolean algebra of events Σ. The impossibility of embedding of QM PBAs into a single Boolean algebra, as we shall discuss below, is precisely the statement of the Kochen-Specker theorem.

For a better understanding of the above notions, let us consider a simple example. Consider a three dimensional Hilbert space $\mathcal{H}$ and three projectors associated with three orthogonal directions, say P_1, P_2, P_3, and let us denote the generated Boolean algebra $\mathfrak{B}_{123}$. Their Boolean relations can be written as follows

a) $P_i \cap P_j = 0$ for any $i \neq j$,

b) $P_1^c \cap P_2^c \cap P_3^c = 0$.

In terms of truth-value assignments, they read

a') P_i and P_j cannot be simultaneously "true"

b') Not all three can be simultaneously "false"

In the next section, we shall see how such conditions cannot be simultaneously satisfied for some particular sets QM projectors.

1.3.3 Kochen-Specker theorem

Consider a collection of orthogonal triads $\{(ijk)\}$, and the associated partial Boolean algebra, $\{\mathfrak{B}_{ijk}\}$. Independently of the probabilistic predictions induced by a quantum state, a necessary condition for the existence of a NCHV theory is the existence of an embedding $\{\mathfrak{B}_{ijk}\} \hookrightarrow \mathfrak{B}$, into a single Boolean algebra $\mathfrak{B}$.

Kochen and Specker proved (Ref. [5] Th. 0) that a necessary and sufficient condition for the existence of such an embedding is the existence of "enough" consistent truth-value assignments, e.g., respecting rules $(a'), (b')$ above, to the set of propositions. They then proceed to show a partial Boolean algebra admitting no consistent truth-value assignment.

Kochen and Specker's original proof has been subsequently simplified [45, 46, 47]. A simple proof, based on Peres' argument for dimension 4, was then proposed by Peres and Mermin [48, 49]. By translating logical relations $(a'), (b')$ into multiplicative rules for ± 1-valued assignments, the argument is greatly simplified. The proof is based on the following set of observables known as the *Peres-Mermin (PM) square*

$$\begin{array}{lll} A = \sigma_z \otimes \mathbb{1}, & B = \mathbb{1} \otimes \sigma_z, & C = \sigma_z \otimes \sigma_z, \\ a = \mathbb{1} \otimes \sigma_x, & b = \sigma_x \otimes \mathbb{1}, & c = \sigma_x \otimes \sigma_x, \\ \alpha = \sigma_z \otimes \sigma_x, & \beta = \sigma_x \otimes \sigma_z, & \gamma = \sigma_y \otimes \sigma_y. \end{array} \tag{1.12}$$

Each observable commutes, and it is therefore jointly measurable, with the other observables in the same row and with those in the same column. Moreover, the product of the observables on each row, i.e., ABC, abc, $\alpha\beta\gamma$, gives the identity $\mathbb{1}$, the same for the columns, except for the last one $Cc\gamma$, which gives $-\mathbb{1}$. This implies that in each joint or sequential measurement of the observables in a row or a column, the outcomes must also satisfy similar rules. For instance, let us denote the outcomes of a measurement as $v(A), v(B), \ldots, v(\gamma)$, then such values satisfy $v(A)v(B)v(C) = 1$, $v(C)v(c)v(\gamma) = -1$ and so on.

Let us assume it is possible to assign a value ± 1 to each observable independently of the measurement contexts, i.e., independently of whether it is measured together with the other observables in the same row or in the same column. Then by multiplying the values along the rows we get

$$[v(A)v(B)v(C)] \times [v(a)v(b)v(c)] \times [v(\alpha)v(\beta)v(\gamma)] = 1 \times 1 \times 1 = 1, \tag{1.13}$$

whereas along the columns we get

$$[v(A)v(a)v(\alpha)] \times [v(B)v(b)v(\beta)] \times [v(C)v(c)v(\gamma)] = 1 \times 1 \times (-1) = -1, \qquad (1.14)$$

which gives us a contradiction since the $v(\cdot)$ are numbers.

1.3.4 State-independent contextuality

The interpretation of rules $a), a'), b), b')$ above in terms of logical constraints and truth-value assignments for propositions in the corresponding HV model is justified by the fact that such relations holds for any quantum state, i.e., for any possible preparation. One might argue that it is a strong constraint to require a NCHV theory to reproduce also such a quantum logical structure, and, as discussed above, one might relax such an assumption and attempt to reproduce only the probabilistic predictions for a given quantum state assuming for the NCHV the most unconstrained classical logical structure, i.e., a free Boolean algebra [50].

Such a procedure is common in Bell scenarios, where one can easily show that the corresponding PBA of projectors can be embedded in a free Boolean algebra [36]. In this framework, noncontextuality for K-S scenarios can be tested experimentally in the same way as the KCBS inequality. In fact, each *KS set*, i.e., a set of projectors not admitting a noncontextual truth-assignment and thus giving rise to a proof of KS theorem, also provides a violation of a specific NC inequality for any quantum state [51, 52]. This phenomenon has been named *state-independent contextuality* (SIC). Conversely, it has been proven that if a NCHV model, in the above unconstrained sense, i.e., assuming only a free Boolean algebraic structure, exists for "enough" quantum states, then also the embedding of the initial PBA of projectors can be obtained via a quotient induced on the free Boolean algebra by the classical probability assignments [36].

One simple example of SIC is given by the PM-square. Let us consider the expression [51],

$$\langle \chi_{\mathrm{PM}} \rangle = \langle ABC \rangle + \langle abc \rangle + \langle \alpha\beta\gamma \rangle + \langle Aa\alpha \rangle + \langle bB\beta \rangle - \langle Cc\gamma \rangle,$$

where the measurements in each of the six sequences are compatible. Then, for NCHV theories the bound

$$\langle \chi_{\mathrm{PM}} \rangle \overset{\mathrm{NCHV}}{\leq} 4 \qquad (1.15)$$

holds. This can be easily proven by trying all 2^9 noncontextual ± 1-value assignments to the above observables. In a four-dimensional quantum system, however, one can take the observables in Eq. (1.12). These observables lead for any quantum state to a value of $\langle \chi_{\mathrm{PM}} \rangle = 6$, demonstrating state-independent contextuality. The quantum violation of Eq. (1.15) has been observed in several recent experiments [53, 54, 55].

However, SIC has been proven not to be an exclusive property of KS proofs. A first preliminary result was given in Ref. [36] where was shown that SIC can also appear for sets of projectors that admit some noncontextual truth-assignments, and thus do not provide a proof of KS theorem, but still are not embeddable into a Boolean algebra. A stronger statement was then proven by Yu and Oh [56], that provided a PBA of projectors embeddable into a Boolean algebra, but that also provide SIC.

Yu and Oh's argument uses the vectors in $\mathbb{C}^3$ listed (not normalized for simplicity) in Tab. 1.1, and the corresponding set of projectors $|v\rangle\langle v|$ and ± 1-valued observables $A_i \equiv 2|v_i\rangle\langle v_i| - \mathbb{1}$. The compatibility relations among observables A_i follows from the orthogonality relations of the corresponding vectors, and are summarized in the graph in Fig. 1.3.

$v_1 = (1,0,0)$	$v_5 = (1,0,-1)$	$v_A = (-1,1,1)$
$v_2 = (0,1,0)$	$v_6 = (1,-1,0)$	$v_B = (1,-1,1)$
$v_3 = (0,0,1)$	$v_7 = (0,1,1)$	$v_C = (1,1,-1)$
$v_4 = (0,1,-1)$	$v_8 = (1,0,1)$	$v_D = (1,1,1)$
	$v_9 = (1,1,0)$	

Table 1.1: Set of vectors giving rise to observables of the Yu and Oh's scenario.

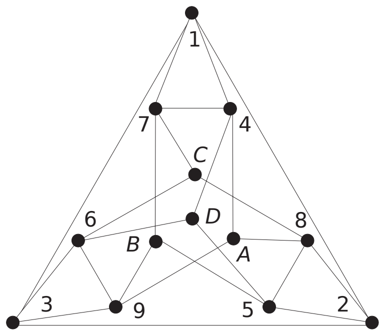


Figure 1.3: Graph of the orthogonality relations among the vectors listed in Tab. 1.1.

Each pair of observables A_iA_j such that ij is an edge of the graph, is therefore jointly measurable. One can therefore write the following NC inequality

$$\sum_i \langle a_i \rangle - \frac{1}{2} \sum_{\text{edges}} \langle a_i a_j \rangle \leq 8, \tag{1.16}$$

where a_i are classical noncontextual random variables and the NCHV bound 8 is computed by trying all possible 2^{13} noncontextual value assignments for a_i. However, using the explicit expression in Tab. 1.1, one can easily compute the quantum value for the operator

$$L = \sum_i A_i - \frac{1}{2} \sum_{\text{edges}} A_i A_j = \frac{25}{3}\mathbb{1}, \tag{1.17}$$

giving

$$\langle L \rangle_\rho = \frac{25}{3} > 8, \tag{1.18}$$

for any quantum state ρ.

The arising of a state-independent violation of a noncontextuality inequality, even for a PBA admitting an embedding into a Boolean algebra can be understood as follows. In quantum mechanics, the projectors associated with nodes A, B, C, D sum up to a multiple of the identity, namely

$$|v_A\rangle\langle v_A| + |v_B\rangle\langle v_B| + |v_C\rangle\langle v_C| + |v_D\rangle\langle v_D| = \frac{4}{3}\mathbb{1}. \tag{1.19}$$

This implies that for any quantum state the sum of their probabilities is $\frac{4}{3} > 1$. On the other hand, from the orthogonality relations among the vectors $\{v_i\}$, which corresponds to exclusivity of the corresponding proposition, i.e., empty intersection of the corresponding Boolean elements, implies that the proposition associated with nodes A, B, C, D are also exclusive.

This can be easily proven as follows. Consider the nodes A and B, and denotes the corresponding Boolean elements as X_A and X_B. From the graph in Fig. 1.3, one can easily see that $X_A \cap X_8 =$

$X_A \cap X_4 = \emptyset$, or equivalently, that $X_A \subset X_8^c$ and $X_A \subset X_4^c$. Similarly, $X_B \subset X_5^c$ and $X_B \subset X_7^c$. We, thus, have that

$$\begin{aligned} X_A \cap X_B \subset X_4^c \cap X_7^c = X_1, \\ X_A \cap X_B \subset X_8^c \cap X_5^c = X_2, \\ \Longrightarrow X_A \cap X_B \subset X_1 \cap X_2 = \emptyset, \end{aligned} \tag{1.20}$$

where we used that the PBA elements represented in the graph as nodes in a triangle sum up to the Boolean algebra identity, as in rules $(b), (b')$ of Sect. 1.3.2.

To summarize, even if the nodes A, B, C, D are not connected in the graph in Fig. 1.3, the Boolean relations with other compatible elements imply that such elements must be disjoint and thus the sum of their probabilities is bounded by one, whereas is QM such a bound does not hold.

1.4 Macrorealist theories and Leggett-Garg inequalities

An approach analogous to Bell and Kochen-Specker has been proposed by Leggett and Garg [8] to investigate the possibility of realization and detection of *macroscopic coherence*, i.e., the quantum superposition of macroscopically distinct states.

More precisely, the two authors introduce the notion of a *macrorealist* hidden variable theory via a list of properties that we expect to be satisfied for a classical macroscopic system. They then proceed to derive a Bell-like inequality for sequential measurements of a single property of the system evolving in time, and show its violation by QM predictions.

1.4.1 Macrorealist theories

The first step is, as always, to define precisely the properties we intuitively expect from a theory describing a classical macroscopic system. Leggett and Garg proposed the following:

MR Macroscopic realism: A system with two or more macroscopically distinct states available to it will at all times be in one of them,

NIM Non-invasive measurability: It is possible, in principle, to determine the state of the system with an arbitrary small perturbation to its subsequent dynamics.

As the previous examples, LHV and NCHV, we have the hypothesis of realism, namely, that the measurement reveals a well defined property of the system. The second assumption plays a role similar to locality and noncontextuality in the derivation of Leggett-Garg inequality, but it is peculiar of a macroscopic system, namely, the possibility of measuring its properties with an arbitrary small perturbation.

The measurement scenario is depicted in Fig. 1.4. A dichotomic variable $Q = \pm 1$, representing a macroscopic property of the system, is measured at fixed instants in time to obtain the two-time correlators $C_{ij} = \langle Q(t_i)Q(t_j)\rangle$, with $t_i < t_j$, defined as the expectation value of the product of the two outcomes, namely,

$$\begin{aligned} \langle Q(t_i)Q(t_j)\rangle &= \sum_{x_i,x_j=\pm 1} x_i x_j Pr(Q(t_i) = x_i, Q(t_j) = x_j) \\ &= \sum_{x_i,x_j} x_i x_j Pr(Q(t_j) = x_j | Q(t_i) = x_i) Pr(Q(t_i) = x_i). \end{aligned} \tag{1.21}$$

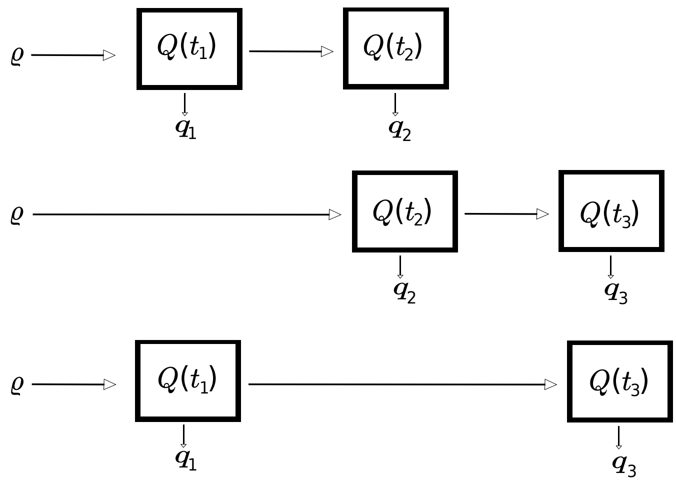


Figure 1.4: Schematic representation of the Leggett-Garg scenario. The system is prepared in the initial state ϱ and a sequence of measurements is performed at fixed instants in time, namely, t_1, t_2 and t_3, to obtain the correlators $\langle Q(t_i)Q(t_j)\rangle$.

The Leggett-Garg inequality is then defined as

$$K_3 \equiv \langle Q(t_1)Q(t_2)\rangle + \langle Q(t_2)Q(t_3)\rangle - \langle Q(t_1)Q(t_3)\rangle \leq 1. \tag{1.22}$$

As for CHSH inequality, the proof is straightforward: By **MR** we can assign a definite value, $+1$ or -1, to each $Q(t_i)$ in every run of the experiment, independent of whether $Q(t_i)$ is measured or not. By **NIM** such a value must be independent of which other measurement are performed before or after t_i. The bound $K_3 \leq 1$ is given by a maximization over all possible ± 1 assignments for $Q(t_i)$ respecting the above conditions in a way analogous to Eq.(1.6), namely

$$\max K_3 = \max_{Q(t_i)=\pm 1} \{Q(t_1)Q(t_2) + Q(t_2)Q(t_3) - Q(t_1)Q(t_3)\} = 1 \tag{1.23}$$

1.4.2 Continuous variables

It is interesting to notice that, for the case of a bounded continuous variable measurement, which we can always normalize as $Q(t_i) \in [-1, 1]$, the same bound applies. Obviously, analogous arguments apply also the the case of LHV and NCHV bounds. Since in Chapt. 5 we will consider the Leggett-Garg inequality in the limit of an infinite number of outcomes, it is interesting here to show in detail that even in the continuous limit the classical bound does not change.

The correlator C_{ij} can be defined as

$$\langle Q(t_i)Q(t_j)\rangle = \int_{x_i,x_j\in[-1,1]} x_i x_j \rho(x_i, x_j) dx_i dx_j, \tag{1.24}$$

where $x_i, x_j \in [-1, 1]$ represent the continuous-variable outcomes for $Q(t_i), Q(t_j)$ and $\rho(x_i, x_j)$ is the corresponding joint probability distribution.

The difference is that the maximum must be calculated over a the interval $[-1, 1]$, but the bound remains the same, namely,

$$\max_{|Q(t_i)|\leq 1} \{Q(t_1)Q(t_2) + Q(t_2)Q(t_3) - Q(t_1)Q(t_3)\} = 1. \tag{1.25}$$

The maximum is taken on the three-dimensional cube $\{|Q(t_i)| \leq 1 \mid i = 1, 2, 3\}$, it can be proven that such a maximum is obtained at the vertices of the cube as follows. The function $f = xy+yz-xz$ we want to maximize is a harmonic function (i.e., $\nabla^2 f = 0$), so its maximum on a compact set, the cube $|x|, |y|, |z| \leq 1$, is achieved on the boundaries of the set. One can then check the maximum on each face of the cube, which corresponds to fixing one coordinate, let us say x, to ± 1. We have, therefore a new function $\tilde{f}(y, z)$ on a square. Again, $\tilde{f}$ is harmonic, so we just have to check the boundaries, so we either fix $y = \pm 1$ or $z = \pm 1$. We then have an harmonic function on a segment, which achieves its maximum on the boundary points.

Eq. (1.25) fixes the value for deterministic assignments, expectation values are given by

$$\begin{aligned} K_3 = \langle Q(t_1)Q(t_2)\rangle + \langle Q(t_2)Q(t_3)\rangle - \langle Q(t_1)Q(t_3)\rangle = \\ \int f(x, y, z)\rho(x, y, z)dxdydz \leq \\ \max |f(x, y, z)| \int \rho(x, y, z)dxdydz = 1, \end{aligned} \tag{1.26}$$

where ρ is a classical probability distribution for x, y, z, i.e., $Q(t_1), Q(t_2), Q(t_3)$, and f is defined as above.

1.4.3 Quantum violations

For dichotomic measurements, temporal correlations appearing in LG inequality can be computed as follows. Given a ± 1-valued measurement Q, with spectral decomposition $Q = \Pi_+ - \Pi_-$, and assuming Lüders rule for the state update after the measurement [57, 58], namely, $\rho \to \Pi_\pm \rho \Pi_\pm$, up to a normalization factor and depending on the outcome ± 1, the value of the temporal correlation can be written as

$$\langle Q(t_i)Q(t_j)\rangle = \sum_{a,b=\pm} q_a q_b \mathrm{tr}(\Pi_b U_{ji} \Pi_a U_{i0} \rho_0 U_{i0}^\dagger \Pi_a U_{ji}^\dagger), \tag{1.27}$$

where q_a represent the outcome, ± 1, associated with Π_a, ρ_0 is the initial state of the system and $U_{ji} = U(t_j - t_i) = e^{-iH(t_j - t_i)}$ is the unitary time-evolution operator for some Hamiltonian H.

Already for a two level system undergoing coherent oscillations between two states, associated with values $+1$ and -1 for the property Q, one can reach the value

$$\langle Q(t_1)Q(t_2)\rangle + \langle Q(t_2)Q(t_3)\rangle - \langle Q(t_1)Q(t_3)\rangle = \frac{3}{2}, \tag{1.28}$$

thus violating the bound (1.22).

Leggett and Garg original proposal was to test their inequality on rf-SQUID flux qubit [8]. The first experimental verification was performed 25 years later by Palacios-Laloy *et al.* [59] on a similar system, a superconducting qubit of the trasmon type, but with continuous weak measurements instead of projective ones. Many experiments followed, on a wide range of different systems such as photons [60, 61, 62, 63], defect center in diamonds [64, 65], and nuclear magnetic resonance [66, 67].

1.5 Correlations polytopes

In the previous sections we have seen the derivation of specific inequalities giving necessary conditions for the existence of classical hidden variable theories. Actually, inequalities (1.5), (1.10) and (1.22) are part of a set of necessary and sufficient conditions, each one for the specific scenario, for the existence of the corresponding HV theories.

A first result in this direction was proven by Fine [6], who showed that the CHSH inequality, together with its variations given by all possible outcome relabelling, i.e., $A_i \to -A_i$, $B_j \to -B_j$, provide necessary and sufficient conditions for the existence of LHV model for such a scenario.

Such ideas were subsequently generalized by Pitowsky [7, 14] into a systematic approach to the characterization of sets of classical correlations, the *correlation polytope* approach. Notice that, even though the correlation polytope approach was originally discussed for Bell scenario, it can be easily adapted to Kochen-Specker and Leggett-Garg scenarios.

The main idea at the basis of Pitowsky's approach is that classical probability assignments are defined as convex combinations of deterministic assignments. Thus, representing deterministic assignments for a set of events as vectors, the corresponding set of classical probabilities will be a *convex polytope*, i.e., a set generated by convex combinations of a finite set of vectors. By Weyl-Minkowski theorem (see, e.g., [14]), each convex polytope has a double description: One as the convex hull of its vertices u_ε, i.e., the $\mathcal{V}$-representation, and one as a (finite) intersection of half-spaces which generates it, each one given by a linear inequality, i.e., the $\mathcal{H}$-representation.

1.5.1 Definition

We shall provide rigorous definitions and discuss some simple examples to clarify the concepts involved. We use the definition of correlation polytope given in [68], which is the natural generalization of Pitowsky's notion [14] to higher order correlations.

Given a set of propositions $\mathcal{G} = (A_1, \ldots, A_n)$ and a family $\mathcal{I}$ of subsets of $\mathcal{G}$, $\mathcal{I} \subset 2^{\mathcal{G}}$, we define the sets S_k, $k = 2, \ldots, m$ with $m \leq n$, as the sets of logical conjunctions

$$S_k = \{A_{i_1} \wedge \cdots \wedge A_{i_k} \mid i_j \neq i_{j'},\ \{A_{i_1}, \ldots, A_{i_k}\} \in \mathcal{I}\ \}. \tag{1.29}$$

The lines of the truth table associated to the above set of propositions and logical conjunctions between them, namely the 2^n vectors of $\mathbb{R}^{|\mathcal{G}|+|S_2|+\ldots+|S_m|}$,

$$u_\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n, \ldots, \varepsilon_i\varepsilon_j, \ldots, \varepsilon_{i_1}\varepsilon_{i_2}\cdots\varepsilon_{i_m}, \ldots), \tag{1.30}$$

where $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in \{0,1\}^n$, are called the vertices of the correlation polytope. Their convex hull, i.e., the set of points generated by their convex combinations, is called the correlation polytope associated with $\mathcal{I}$ and denoted as $COR(\mathcal{I})$.

It is convenient to introduce the following notation which makes apparent the correspondence between coordinates and joint probabilities. The coordinates of a point $p \in \mathbb{R}^{|\mathcal{G}|+|S_2|+\ldots+|S_m|}$ will be denoted as

$$p = (p_1, \ldots, p_n, \ldots p_{ij}, \ldots, p_{i_1\ldots i_m}, \ldots). \tag{1.31}$$

In general, the set $\mathcal{I}$ will be different from $2^{\mathcal{G}}$, since the vector (1.31) must contain only the joint probabilities that are actually measured, namely, for joint measurements of different subsystems in Bell scenarios, for compatible measurement in noncontextuality scenarios, and for sequential measurements in Leggett-Garg scenarios.

Linear inequalities now arise as a consequence of Weyl-Minkowski theorem [14]: Each convex polytope, i.e., the convex hull of a finite set of vertices, is also a convex polyhedron, i.e., a bounded set described by a finite set of linear inequalities, and vice-versa. From each set of inequalities, a minimal set of non *redundant*, i.e., not implied by the others, can be extracted. Geometrically, such a minimal set is given by the inequalities tangent to the *facets* of the polytope, also known as *tight inequalities*. We recall that a face F of a polytope P is a subset $F \subset P$ such that for any $x \in F$, every decomposition $x = \alpha y + (1-\alpha)z$ for $y, z \in P$, implies $y, z \in F$. *Facets* are defined as the $(d-1)$-dimensional faces of a d-dimensional polytope.

As a consequence, an inequality is tight if and only if it is valid, i.e., satisfied by all points of the polytope, and it is saturated by a set of vertices generating a $(d-1)$-dimensional affine subspace. A schematic representation of the above notions is given in Fig. (1.5).

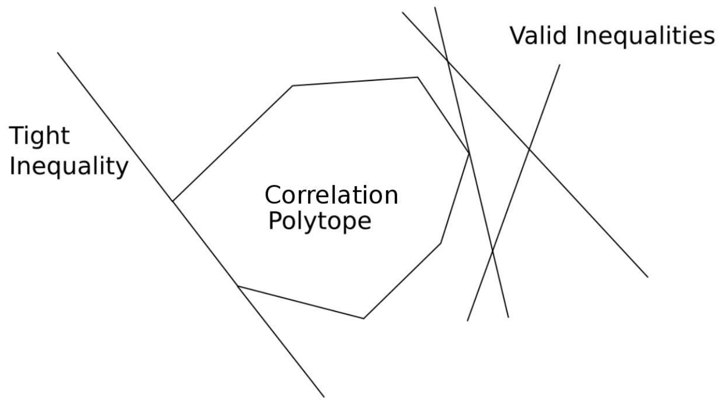


Figure 1.5: Schematic representation of tight and non-tight inequalities

Pitowsky's main result [7] can be stated as follow

Theorem (Pitowsky 1986). *Given a measurement scenario described by the set $\mathcal{I}$, the vector of measurable probabilities p belongs to the corresponding correlation polytope if and only if there exists a classical probability space representation for p.*

The basic idea is that every probability assignment, represented here as a vector, is a convex combination of deterministic assignments, the vertices of the polytope. Deterministic assignments will therefore be the points of the probability space, i.e., the atomic events, and the coefficients of convex combination will give the relative frequencies at which such events can happen.

1.5.2 Examples

It is instructive to consider a simple example: The correlation polytope for two propositions, A_1, A_2, and their logical conjunction $A_1 \wedge A_2$. The four vertices of the polytope correspond to the rows of the following truth table for $(A_1, A_2, A_1 \wedge A_2)$

A_1	A_2	$A_1 \wedge A_2$
0	0	0
0	1	0
1	0	0
1	1	1

(1.32)

As can be easily deduced from Fig. (1.6), vertices in table (1.32) form the tetrahedron with the following facet inequalities

$$\begin{aligned} p_1 - p_{12} &\geq 0, \\ p_2 - p_{12} &\geq 0, \\ p_{12} &\geq 0, \\ 1 - p_1 - p_2 + p_{12} &\geq 0, \end{aligned} \tag{1.33}$$

with the coordinate labelling as in Eq. (1.31).

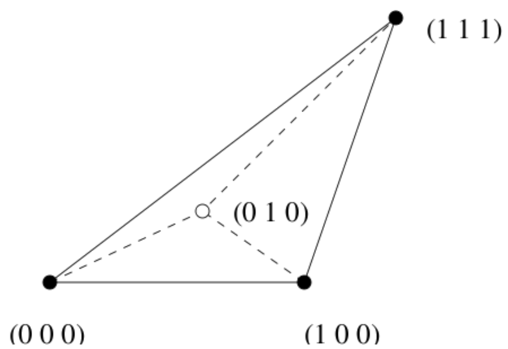


Figure 1.6: Correlation polytope for the truth table (1.32). The four vertices correspond to the rows of Table 1.32 and coordinates are labelled consequently. Inequalities (1.33) correspond to the four faces of the tetrahedron: The plane $p_{12} = 0$ is the plane tangent to vertices $(0,0,0),(1,0,0)$ and $(0,1,0)$, the plane $p_1 - p_{12} = 0$ is the one tangent to the vertices $(0,0,0),(0,1,0)$, and $(1,1,1)$, and so on.

The next example, which is directly related to HV models, is the CHSH scenario of Sect. 1.2. To make the notation consistent with the above definitions, let us denote Alice's measurements as A_1 and A_2, with outcome $0,1$, and Bob's measurements as A_3 and A_4, again with outcome $0,1$. The set of propositions is, therefore, $\mathcal{G} = (A_1, A_2, A_3, A_4)$. Moreover, the measurements associated with A_i and A_j, for $i = 1,2$ and $j = 3,4$, can be performed jointly and, consequently, it makes sense to consider the following set of logical conjunctions $S_2 = \{A_1 \wedge A_3, A_1 \wedge A_4, A_2 \wedge A_3, A_2 \wedge A_4\}$. The associated polytope is described by $2^4 = 16$ vertices in $\mathbb{R}^8$, namely

$$u_\varepsilon = (\varepsilon_1,\ \varepsilon_2,\ \varepsilon_3,\ \varepsilon_4,\ \varepsilon_1\varepsilon_3,\ \varepsilon_1\varepsilon_4,\ \varepsilon_2\varepsilon_3,\ \varepsilon_2\varepsilon_4), \qquad \varepsilon_i \in \{0,1\}, \tag{1.34}$$

where ε_i represents a classical $\{0,1\}$-valued assignment to proposition A_i and $\varepsilon_i\varepsilon_j$ the classical assignment for the logical conjunction $A_i \wedge A_j$.

The convex hull of vertices (1.34) gives a set of linear inequalities constraining the coordinates of a generic point in $\mathbb{R}^8$, which we denote by

$$p = (p_1, p_2, p_3, p_4, p_{13}, p_{14}, p_{23}, p_{24}). \tag{1.35}$$

The interpretation of such a geometrical object is the following: Given a vector belonging to the polytope, each component represents the joint probability for the corresponding subset of propositions, e.g., p_{13} represents the joint probability $Prob(A_1 \wedge A_3)$, namely $Prob(A_1 = 1, A_3 = 1)$. The whole polytope gives possible ranges for such joint probabilities if the underlying probabilistic structure is assumed to be classical, i.e., given by a probability space or, equivalently (since the number of proposition is finite) by a normalized measure on a finite Boolean algebra.

The CHSH polytope is then given by the following inequalities [14]

$$0 \leq p_{ij} \leq p_i, \; 0 \leq p_{ij} \leq p_j, \; i=1,2, \; j=3,4 \tag{1.36}$$
$$p_i + p_j - p_{ij} \leq 1, \; i=1,2, \; j=3,4, \tag{1.37}$$
$$-1 \leq p_{13} + p_{14} + p_{24} - p_{23} - p_1 - p_4 \leq 0, \tag{1.38}$$
$$-1 \leq p_{23} + p_{24} + p_{14} - p_{13} - p_2 - p_4 \leq 0, \tag{1.39}$$
$$-1 \leq p_{14} + p_{13} + p_{23} - p_{24} - p_1 - p_3 \leq 0, \tag{1.40}$$
$$-1 \leq p_{24} + p_{23} + p_{13} - p_{14} - p_2 - p_3 \leq 0. \tag{1.41}$$

The inequalities (1.38)-(1.41) are different variants, obtained via a permutation of the measurement settings, of the Clauser-Horne inequality [69]. They are obtained from the CHSH inequality, and corresponding variants, via a relabelling of the outcomes, i.e., $\{-1,+1\} \to \{0,1\}$.

Via the correlation polytope method, the problem of finding necessary and sufficient conditions for the existence of a HV model for a given measurement scenario amounts to the geometric problem known as the *hull problem.* Algorithms solving the hull problem are known, and different implementation are available (e.g., `cdd` [70], `lrs` [71], and `porta` [72]). However, the running time of such algorithms grows exponentially in the number of vertices of the polytope and the number of vertices grows exponentially with the number of measurement settings. The problem, therefore, becomes rapidly computationally intractable, even in the bipartite case the membership problem (deciding whether p belongs to $COR(\mathcal{I})$) is NP-complete [73].

1.6 Tsirelson bound

Bell, noncontextuality, and Leggett-Garg inequalities bound the possible correlations in, respectively, LHV, NCHV and MRHV theories. A similar approach can be pursued for quantum theory by asking whether there exist similar bounds for quantum correlations. The answer was given by Tsirelson [74], who developed a general framework for treating such a problem in Bell scenarios and proved an inequality bounding the quantum correlations in the CHSH scenario.

1.6.1 Original argument

The following simple proof of Tsirelson bound has been presented by Landau [75]. The Bell operator of Eq. (1.5) can be written as

$$\mathcal{B} = A_0B_0 + A_0B1 + A_1B_0 - A_1B_1 = A_0(B_0 + B_1) + A_1(B_0 - B_1), \tag{1.42}$$

then taking the square and using that $A_i^2 = B_j^2 = \mathbb{1}$ we obtain

$$\mathcal{B}^2 = 4\mathbb{1} - [A_0, A_1][B_0, B_1]. \tag{1.43}$$

Since we are dealing with $\pm$-valued observables, $\|A_i\| = \|B_j\| = 1$, which implies $\|[A_0, A_1]\| \leq 2$ and $\|[B_0, B_1]\| \leq 2$, and $\|[A_0, A_1][B_0, B_1]\| \leq 4$. The quantum value for the CHSH scenario can, thus, be bounded by

$$\langle \mathcal{B} \rangle^2 \leq \langle \mathcal{B}^2 \rangle \leq 8 \quad \Longrightarrow \quad \langle \mathcal{B} \rangle \leq 2\sqrt{2}. \tag{1.44}$$

Tsirelson's original result was derived by proving the correspondence between quantum correlations for observables of norm bounded by 1 and scalar product of vectors in a Euclidean vector space. It is instructive to recall this result since it will play a fundamental role in the redefinition of the problem as a semidefinite program.

Theorem (Tsirelson 1980)**.** *The following conditions for real numbers c_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, m$ are equivalent.*

(1) *There exist a C^*-algebra with identity $\mathcal{A}$ and self-adjoint elements $A_1, \ldots, A_n$, $B_1, \ldots, B_n$, and a state f on $\mathcal{A}$ such that*

$$[A_i, B_j] = 0, \quad \|A_i\| \leq 1, \ \|B_j\| \leq 1, \quad f(A_i B_j) = c_{ij} \tag{1.45}$$

(2) *There are unit vectors $x_1, \ldots, x_n$, $y_1, \ldots, y_m$ in a $(n+m)$-dimensional Euclidean vector space such that*

$$x_i \cdot y_j = c_{ij}, \tag{1.46}$$

where $\cdot$ denotes the Euclidean scalar product.

1.6.2 Semidefinite programming approach

Wehner's approach

Such a correspondence allows us to reformulate the problem as a maximization problem over the set of positive semidefinite matrices in the following way [76]. For the CHSH scenario, we construct the Gram matrix G of scalar products of vectors x_i, y_j, which eventually we want to associate with the corresponding quantum correlations,

$$G = \begin{pmatrix} x_1 \cdot x_1 & x_1 \cdot x_2 & x_1 \cdot y_1 & x_1 \cdot y_2 \\ x_2 \cdot x_1 & x_2 \cdot x_2 & x_2 \cdot y_1 & x_2 \cdot y_2 \\ y_1 \cdot x_1 & y_1 \cdot x_2 & y_1 \cdot y_1 & y_1 \cdot y_2 \\ y_2 \cdot x_1 & y_2 \cdot x_2 & y_2 \cdot y_1 & y_2 \cdot y_2 \end{pmatrix} . \tag{1.47}$$

We recall that given k vectors $v_1, \ldots, v_k$, their Gram matrix is defined as the $k \times k$ matrix whose entry ij is given by the scalar product $v_i \cdot v_j$. By construction the matrix is symmetric and positive semidefinite, i.e., $vGv \geq 0$ for all vectors v, which we denote as $G \succeq 0$. In fact, G can be written as $A^t A$, where the columns of A are given by the vectors v_i. Conversely, every positive semidefinite matrix can be written as a Gram matrix, just by taking as corresponding vectors the columns of the decomposition $A^t A$, where A can be, e.g., the square root of the original matrix.

The quantum bound for the CHSH scenario can be, therefore, computed via the maximization

$$\begin{aligned} \text{maximize:} \quad & \sum_{ij} \lambda_{ij} G_{ij}, \\ \text{subjected to:} \quad & G = G^T \succeq 0 \text{ and for all } i,\ G_{ii} = 1. \end{aligned} \tag{1.48}$$

for a suitable choice of the coefficients λ_{ij} that singles out the expression

$$x_1 \cdot y_1 + x_1 \cdot y_2 + x_2 \cdot y_1 - x_2 \cdot y_2 \ . \tag{1.49}$$

The above maximization problem is known as *semidefinite program* (SDP) [77], a class of optimization problems that are known to be efficiently solvable numerically, with the optimality of the solution certifiable up to an arbitrary precision. A similar approach has also been investigate by other authors [78] and finally a general solution, based on SDP and valid for arbitrary number of measurement settings and outcomes in Bell and noncontextuality scenarios, has been provided by Navascués, Pironio and Acín (NPA) [79, 80], as we shall discuss below.

Navascués-Pironio-Acín method

NPA method allows one not only to calculate the quantum violation of a given Bell inequality, but also to test the membership of a set of correlation with respect to the quantum set, in an analogous way as the correlation polytope membership test.

NPA method, for the simple case of a two-party Bell scenario, can be briefly described as follows. Let us denote by $\Pi(r|s)$ the projector operator associated with the result r for the measurement of the setting s. The set of indices is partitioned in two disjoint sets, A and B, representing, respectively, Alice's and Bob's settings. Such projectors satisfy

$$\begin{aligned} \Pi(r|s) &= \Pi(r|s)^\dagger, && \text{(Hermiticity)}, \\ \Pi(r|s)\Pi(r'|s) &= \delta_{rr'}\Pi(r|s), && \text{(Orthogonality)}, \\ \sum_r \Pi(r|s) &= \mathbb{1}, && \text{(Completeness)}, \\ [\Pi(r|s), \Pi(r'|s')] &= 0 \text{ if } s \in A \text{ and } s' \in B, && \text{(Commutativity)}. \end{aligned} \tag{1.50}$$

We can define a sequence of measurement results and settings as a vector $(\mathbf{r}|\mathbf{s}) = (r_1, \ldots, r_k|s_1, \ldots, s_k)$ and define the corresponding sequence of projectors as $\Pi(\mathbf{r}|\mathbf{s}) = \Pi(r_1|s_1)\Pi(r_2|s_2)\ldots\Pi(r_k|s_k)$, where, by definition, $\Pi((\mathbf{r}|\mathbf{s}) = \mathbb{1}$ for the sequence of length zero. For a given set of sequences $S = \{(\mathbf{r}|\mathbf{s})\}$, we introduce the $|S| \times |S|$ matrix of moments

$$M_{\mathbf{r}|\mathbf{s};\mathbf{r}'|\mathbf{s}'} = \mathrm{tr}(\rho\Pi(\mathbf{r}|\mathbf{s})^\dagger\Pi(\mathbf{r}'|\mathbf{s}')). \tag{1.51}$$

As a consequence of equations (1.50), the entries of M satisfy a set of linear constraints. Moreover, M is positive semidefinite. In fact, given a vector v, $v^\dagger M v = \mathrm{tr}(\rho C^\dagger C) \geq 0$, where $C = \sum_{\mathbf{r}|\mathbf{s}\in S} v_{\mathbf{r}|\mathbf{s}}\Pi(\mathbf{r}|\mathbf{s})$.

The maximization problem can be therefore written as

$$\begin{aligned} \text{maximize:} \quad & \sum_{ij} \lambda_{ij} M_{ij}, \\ \text{subjected to:} \quad & M \succeq 0 \text{ and linear constraints implied by (1.50)}, \end{aligned} \tag{1.52}$$

Notice that, in general, M is a Hermitian complex matrix, but if the coefficients λ_{ij} are real, than for any solution M of the problem (1.52), the complex conjugate M^* is also a solution, and, consequently, the combination $(M + M^*)/2$. Without a loss of generality, we can therefore consider only real matrices M.

As opposed to the problem (1.48) and, more generally, the analysis of quantum bounds for expressions containing only pair correlations observables, where Tsirelson theorem gives an explicit correspondence between vector of the Gram matrix, i.e., the solution of the SDP, and the quantum observables to be measured, in the more general framework of NPA method it is not always possible to reconstruct the quantum observables attaining the bound computed via SDP. In general, the problem (1.52) gives only an upper bound to the maximum quantum violation of a given Bell or noncontextuality inequality. However, NPA proved that by extending the matrix M to include longer sequences $S = \{(\mathbf{r}|\mathbf{s})\}$ one gets stricter upper bounds and in the limit $|S| \to \infty$ the SDP bound coincide with the quantum bound. A more detailed discussion of this mechanism can be found in Chapt. 4.

Cabello-Severini-Winter method

An alternative SDP-based approach to the computation of quantum bounds has been proposed by Cabello, Severini, and Winter [81, 82]. The starting point is the observation that correlations in

Bell and noncontextuality inequality can be written as a positive linear combination of probabilities of events. For instance, the KCBS inequality 5.3 can be equivalently written as

$$S_{KCBS} = \sum_{i=0}^{4} P(-1,+1|i,i+1) \overset{\text{NCHV}}{\leq} 2, \tag{1.53}$$

where $P(-1,+1|i,i+1) \equiv Prob(A_i = -1, A_{i+1} = 1)$ and the sum is taken modulo 4. In QM, such events are represented projectors, e.g., $P(-1,+1|i,i+1) = \text{tr}(\rho\Pi_i^+\Pi_{i+1}^-)$, and mutually exclusive events correspond to orthogonal projectors, e.g., $\Pi_i^+\Pi_i^- = 0$.

The authors noticed the similarity between Eq.(1.53) and the definition of the Lovász number of a graph $G = (V,E)$ [83, 84], namely

$$\vartheta(G) = \max_{v_i,\psi} \sum_{i\in V} |\langle\psi|v_i\rangle|^2, \tag{1.54}$$

where the sum is take over all vectors $|\psi\rangle$ and over all vectors $|v_i\rangle$ forming an orthogonal representation (OR) of G, namely a set of vectors in $\mathbb{R}^d$, such that $\langle v_i|v_j\rangle = 0$ whenever $i,j \in V$ are non adjacent vertices, i.e., $\{ij\} \notin E$.

The maximum of the expression S_{KCBS} in QM can be in fact written as

$$\max_{\rho,\Pi_i^{\pm}} \sum_i \text{tr}(\rho\Pi_i^+\Pi_{i+1}^-) = \max_{v_i,\psi} \sum_{i\in V} |\langle\psi|v_i\rangle|^2, \tag{1.55}$$

where each vertex of the graph $G = (V,E)$ correspond to a projector appearing on the l.h.s. of (1.55) and two vertices are adjacent if the corresponding projectors are non orthogonal (add note on the confusion of OR for the graph and the complement). Notice that the use of a pure state $|\psi\rangle$ instead of ρ is no restriction since, by a convexity argument, the maximum of S_{KCBS} is always obtained with a pure state, and the same for the use of one-dimensional projectors $|v_i\rangle\langle v_i|$, since $\langle\psi|\Pi|\psi\rangle = |\langle\psi|v_i\rangle|^2$ where $|v_i\rangle = \Pi|\psi\rangle/\sqrt{\langle\psi|\Pi|\psi\rangle}$.

Initially introduced as an upper bound on the Shannon capacity of a graph [83], the Lovász number is a well studied object in graph theory. It can be efficiently computed via semidefinite programming.

1.7 Linear and semidefinite programming

To conclude the introduction we discuss two classes of convex optimization problems that can be efficiently solved via numerical methods, namely, linear programming and semidefinite programming. More details can be found in Refs.[77, 85]

Definition of a semidefinite program

In general terms, a *semidefinite program* can be defined as a minimization of a linear function of the variable $x \in \mathbb{R}^m$ subjected to a matrix inequality

$$\begin{aligned} \text{minimize:} \quad & c\cdot x, \\ \text{subjected to:} \quad & F(x) \equiv x_0 + \sum_{i=1}^{m} x_i F_i \succeq 0, \end{aligned} \tag{1.56}$$

where the problem data are the vector $c \in \mathbb{R}^m$ and $m+1$ symmetric $n \times n$ real matrices $F_0, F_1, \ldots, F_m$, $\cdot$ denotes the Euclidean scalar product as above, and the symbol $\succeq$ denotes the fact that F is positive semidefinite, i.e., $vF(x)v \geq 0$ for all $v \in \mathbb{R}^n$. We say that x is *feasible* if $F(x) \succeq 0$.

Notice that the optimization is performed over a convex set since given $F(x) \succeq 0$ and $F(y) \succeq 0$, for all λ such that $0 \leq \lambda \leq 1$,

$$F(\lambda x + (1-\lambda)y) = \lambda F(x) + (1-\lambda)F(y) \succeq 0. \tag{1.57}$$

Semidefinite programs are therefore a subclass of the more general convex optimization problems [85].

Definition of a linear program

A *linear program* is defined as a minimization of a linear function over a convex polyhedral set, more precisely,

$$\begin{aligned} \text{minimize:} \quad & c \cdot x, \\ \text{subjected to:} \quad & Ax + b \geq 0, \end{aligned} \tag{1.58}$$

where A is a $n \times m$ real matrix, b a n-dimensional vector and the inequality sign $\geq$ is intended componentwise.

It can be easily verified that the linear program (1.58) can be written as the semidefinite program (1.56) by defining

$$F_0 = diag(b), \quad F_i = diag(a_i), \text{ for } i = 1, \ldots, m, \tag{1.59}$$

where $diag(b)$ denotes the diagonal $n \times n$ matrix with diagonal entries the entries of b, and a_i are the columns of the matrix A. Semidefinite programming can, in fact, be regarded as a generalization of linear programming obtained by replacing the componentwise inequalities by matrix inequalities, i.e., positive-semidefiniteness conditions.

Duality and numerical computation

Given a semidefinite program of the form (1.56), one can define the *dual problem* as

$$\begin{aligned} \text{maximize:} \quad & -\text{tr}(ZF_0), && (1.60) \\ \text{subjected to:} \quad & \text{tr}(ZF_i) = c_i, \text{ for all } i = 1, \ldots m \\ & Z \succeq 0, && (1.61) \end{aligned}$$

where the maximization is over the variable Z, a $n \times n$ real symmetric matrix. The original problem (1.56) is called the *primal problem*. We say that a real and symmetric matrix Z is *dual feasible* if $\text{tr}(ZF_i) = c_i$, for all $i = 1, \ldots m$, and $Z \succeq 0$.

One can prove that the problem (1.60) can be rewritten as (1.56), and, thus, every the dual of a semidefinite program is still a semidefinite program.

The fundamental property of the dual semidefinite program is that it gives bounds to the optimal value of the primal problem, and viceversa. In fact, suppose x is primal feasible and Z is dual feasible, then

$$c \cdot x + \text{tr}(ZF_0) = \sum_{i=1}^{m} \text{tr}(ZF_i x_i) + \text{tr}(ZF_0) = \text{tr}(ZF(x)) \geq 0, \tag{1.62}$$

where we used that $\mathrm{tr}(AB) \geq 0$ for $A = A^t \succeq 0$ and $B = B^t \succeq 0$. Eq. (1.62) implies

$$-\mathrm{tr}(ZF_0) \leq c \cdot x, \tag{1.63}$$

namely, that the dual objective value of every feasible dual point Z gives a lower bound on the primal objective value of any primal feasible point x.

Let us denote the optimal value of the primal problem (1.56) as

$$p^* = \inf\{c \cdot x \mid F(x) \succeq 0\ \} \tag{1.64}$$

and the optimal of the dual problem (1.60) as

$$d^* = \sup\{-\mathrm{tr}(ZF_0) \mid Z \succeq 0,\ \mathrm{tr}(ZF_i) = c_i,\ \text{for all } i = 1, \dots m\ \}. \tag{1.65}$$

From Eq. (1.63) it follows that for any primal feasible vector x and dual feasible matrix Z

$$-\mathrm{tr}(ZF_0) \leq d^* \leq p^* \leq c \cdot x, \tag{1.66}$$

We can now state the main theorem on duality for SDP [77]

Theorem. *We have $p^* = d^*$ if either of the following condition holds.*

(i) The primal problem is strictly feasible, i.e., there exists x with $F(x) \geq 0$.

(ii) The dual problem is strictly feasible, i.e., there exists Z with $Z \succeq 0$ and $\mathrm{tr}(ZF_i) = c_i$ for all $i = 1, \dots, m$.

Moreover, if both conditions hold, p^* and d^* are, respectively, the minimum and the maximum of the sets appearing in Eqs. (1.64),(1.65).

The above results implies that a numerical solution of a SDP gives the interval in which the exact solution lies. In fact, what the numerical algorithms do is to computed a feasible point for the primal problem, let us denote it as p_{num}, and a feasible point for the dual problem, let us denote it as d_{num}, thus giving

$$d_{num} \leq d^* \leq p^* \leq p_{num}. \tag{1.67}$$

The exact solution can then be, in principle, approximate up to an arbitrary precision.

Chapter 2

Noncontextuality inequalities from variable elimination

In this chapter we present a general method for deriving complete sets of Bell and noncontextuality inequalities alternative to the direct solution of the hull problem for correlation polytopes, discussed in the previous chapter, which provides some computational advantages and allows for a complete characterization of the correlation polytope even for scenarios with an arbitrary number of settings. The main results of this chapter have been published in Refs. [86, 87].

Our method is based on the application of Fourier-Motzkin (FM) method of variable elimination for systems of linear inequalities to conditions derived in Ref. [36] as consistency conditions for putting together partial extensions of quantum probabilities in order to obtain a classical probability description. More precisely, such conditions are expressed in terms of a systems of linear inequalities where also correlations between incompatible observables appear as variables: A classical probability space representation exists for a given set of QM predictions if and only if the corresponding system of linear inequalities admits a solution; Bell, or noncontextuality, inequalities are obtained by eliminating the variables associated with correlations between incompatible observables.

Our approach can be seen as a generalization of Fine's derivation of the Clauser-Horne-Shimony-Holt (CHSH) polytope [6]. It provides a generalization of the result obtained by Śliwa [88] and Collins and Gisin [89], namely, the appearance of only a finite number of families of Bell inequalities in measurement scenarios where one experimenter is allowed to choose between an arbitrary number of different measurements, and it allows for a complete characterization of a specific n-setting noncontextuality scenario.

In the following, we shall present the general method and discuss it more in details by means of simple examples. We then proceed to analyze some Bell and noncontextuality scenarios, providing in some cases analytical results, and showing the computational advantage in others.

2.1 Extension of measures and consistency conditions

First we need a result on extension of probability measures which will be the basis of our approach to the computation of $\mathcal{H}$-representation for correlation polytopes. More details can be found in Ref.[36].

Proposition 1. *Consider any set of probabilities p_i on a set of yes/no (i.e., $\{0,1\}$-valued) observables A_i and correlations p_{ij} on a subset of pairs A_i, A_j, defining a probability on each pair A_i, A_j, with $p_i = \langle A_i \rangle$, $p_j = \langle A_j \rangle$, $p_{ij} = \langle A_i A_j \rangle$; describing observables A_i as vertices and the above pairs* *as edges in a graph, any set of predictions associated with a tree graph, i.e., a graph without closed loops, admits a classical representation.*

We call such a graph representation a *compatibility graph.* Some examples of compatibility graphs are given in Figs. 2.1, 2.2, 2.3, 2.4, and examples of tree graphs are given by Fig. 2.1 (b),(c),(d).

Prop. 1 can then be generalized as follows [36].

Proposition 2. *the same holds with yes/no observables A_i substituted by free Boolean algebras $\mathcal{A}_i$, p_i by probabilities on $\mathcal{A}_i$, p_{ij} by probabilities on the Boolean algebra freely generated by the union of the sets of generators of $\mathcal{A}_i$ and $\mathcal{A}_j$.*

We recall that a Boolean algebra is freely generated by n generators $B_1, \ldots, B_n$ if such generators are as much unconstrained as possible, i.e., they satisfy no conditions except those necessary conditions defining a Boolean algebra (e.g., distributive law). Since all Boolean algebras freely generated by $n < \infty$ generators are isomorphic, for the sake of simplicity, we can think of the the algebra of subsets 2^X of the set $X = \{0,1\}^n$, with set theoretic operations $(\cap, \cup, ^c)$; then the subsets $B_i = \{(x_1, \ldots, x_n) \in X \mid x_i = 1\}$, for $i = 1, \ldots, n$, can be taken as free generators. In terms of propositions and truth assignments, the free Boolean algebra assumption amounts to the assumption that each possible $\{0,1\}$-valued assignment to propositions is admissible. For more details see [50]. Notice that this is precisely the way how we define the 2^n vertices of the correlation polytope in Eq.(1.30).

The above results suggest a general method for the computation of $\mathcal{H}$-representation for correlation polytopes, based on exploiting the automatically existing classical representations for subsets of observables with compatibility relations described by tree graphs, as in Prop. 1 and 2 above. Conditions for classical representability arise as consistency (i.e., coincidence on intersections) conditions for putting together partial extensions associated with subgraphs, giving rise to a description of the initial compatibility graph as a tree graph on such extended nodes. Such consistency conditions are expressed in terms of the existence of a solution for a set of linear inequalities. One of the main application of Fourier-Motzkin algorithm is precisely deciding whether a system of inequalities has a solution. It is thus sufficient to solve the hull problem for a smaller polytope, i.e., compute the single partial extensions, then constructing a higher dimensional polytope from that solution, i.e., impose the consistency conditions, and finally apply FM algorithm, i.e., derive the conditions for the existence of a solution.

We recall that the Fourier-Motzkin method consists in eliminating a variable from a system of linear inequalities by summing, after a proper normalization, all inequalities where it appears with plus sign with all inequalities where it appears with minus sign. From a geometric point of view, since the system of linear inequalities can represent a polytope (more generally, a cone), the above operation amounts to a projection onto the coordinates associated to the remaining variables; for more details see [90].

Although our method is general, different strategies are possible corresponding to different partitions of the initial graph into subgraphs. We shall introduce the details of our method by means of some simple examples. The first example is the derivation of the CHSH polytope. It is interesting to notice that it is analogous to that presented by Fine [6]; our method can be seen as a generalization of his idea to an arbitrary number of observables. We shall then discuss more complex Bell scenarios and some related computational results, and finally the complete characterization of the correlation polytope for the n-cycle noncontextuality scenario.

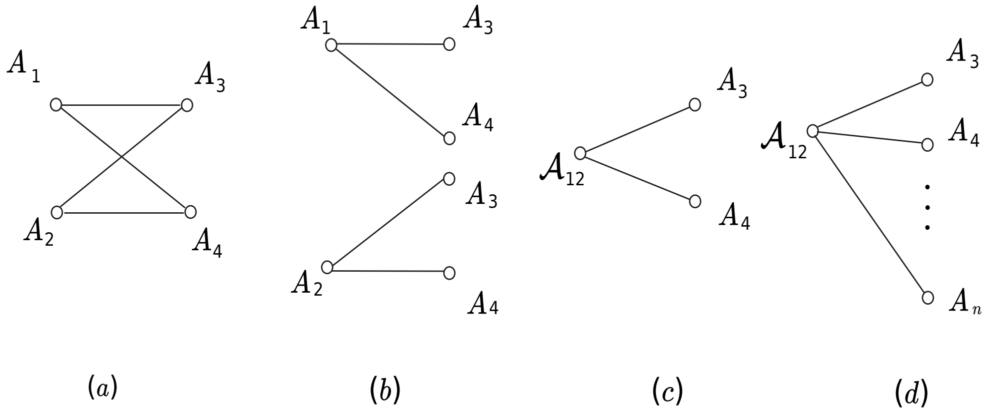


Figure 2.1: (a) Graph of the compatibility relations between the observables in the CHSH scenario. (b) Partition in two subset of three observables with intersection $\{A_1, A_2\}$. (c) Tree graph obtained by extending the probability measure on the algebra generated by A_1, A_2. (d) Asymmetric case with additional observables on Bob's side.

2.2 Bell inequalities

2.2.1 CHSH polytope from Bell-Wigner polytope

The CHSH polytope is generated by the following set of vertices

$$u_\varepsilon = (\varepsilon_1,\ \varepsilon_2,\ \varepsilon_3,\ \varepsilon_4,\ \varepsilon_1\varepsilon_3,\ \varepsilon_1\varepsilon_4,\ \varepsilon_2\varepsilon_3,\ \varepsilon_2\varepsilon_4), \qquad \varepsilon_i \in \{0,1\}. \tag{2.1}$$

As previously discussed, it is associated with a bipartite measurement scenario in which Alice can choose between two measurements, associated with propositions A_1 and A_2, and Bob can choose between two measurements, associated with propositions A_3 and A_4.

As already recognized by Fine [6] and discussed also in [36, 91], the existence of a classical description for the four observables is equivalent to the existence of classical descriptions for the two subsystems, $\{A_1, A_2, A_3\}$ and $\{A_1, A_2, A_4\}$, coinciding on $\{A_1, A_2\}$. In fact, the two classical descriptions would give rise to an extension of the probability assignment to the four observables satisfying the hypothesis of Prop. 2 (see Fig. 2.1 (a),(b),(c)).

The constraints on the subsystem $\{A_1, A_2\}$ imposed by the third observable, A_3 or A_4, are described by the Bell-Wigner polytope [14], i.e., the correlation polytope associated with three proposition and their pairwise logical conjunctions, which is given by the following inequalities:

$$0 \le p_{ij} \le p_i,\ \ 0 \le p_{ij} \le p_j,\ \ ij = 12, 1s, 2s, \tag{2.2}$$
$$p_i + p_j - p_{ij} \le 1,\ \ ij = 12, 1s, 2s, \tag{2.3}$$
$$1 - p_1 - p_2 - p_s + p_{12} + p_{1s} + p_{2s} \ge 0, \tag{2.4}$$
$$p_1 - p_{12} - p_{1s} + p_{2s} \ge 0, \tag{2.5}$$
$$p_2 - p_{12} - p_{2s} + p_{1s} \ge 0, \tag{2.6}$$
$$p_s - p_{1s} - p_{2s} + p_{12} \ge 0, \tag{2.7}$$

for fixed $s = 3$ or 4.

Classical representability for QM prediction in the CHSH scenario amounts, therefore, to the existence of a common solution for the two systems of linear inequalities, namely, it amounts to the existence of a value for p_{12} consistent with the constraints imposed by classical descriptions for the two subsystems of three observables.

As a consequence of general properties of Fourier-Motzkin method (see [90]), a system of inequalities admits a solution if and only if the projected system, i.e., the system obtained by eliminating one or more variables, admits a solution. It follows that the above set of measurements admits a classical description if and only if measured correlations, i.e., correlations between compatible observables, satisfy the system of inequalities obtained by eliminating p_{12}.

In order to eliminate p_{12}, we just apply Fourier-Motzkin method: We sum inequalities where p_{12} appears with opposite sign and keep inequalities where it does not appear. A posteriori, we realize that only redundant inequalities arise from (2.2) and (2.3) and the combination of inequalities with the same index s.

The only interesting, i.e., non redundant, inequalities are those obtained from the combination of (2.4)–(2.7) for different s, namely

$$-1 \leq p_{13} + p_{14} + p_{24} - p_{23} - p_1 - p_4 \leq 0, \tag{2.8}$$

$$-1 \leq p_{23} + p_{24} + p_{14} - p_{13} - p_2 - p_4 \leq 0, \tag{2.9}$$

$$-1 \leq p_{14} + p_{13} + p_{23} - p_{24} - p_1 - p_3 \leq 0, \tag{2.10}$$

$$-1 \leq p_{24} + p_{23} + p_{13} - p_{14} - p_2 - p_3 \leq 0, \tag{2.11}$$

that, together with the Eqs. (2.2), (2.3) in which p_{12} does not appear, give the $\mathcal{H}$-representation of the CHSH polytope.

2.2.2 Bipartite $(2, n)$ scenario

An analogous argument applies to the scenario in which Alice can choose between two measurements and Bob can choose among $n > 2$ measurements, associated with propositions $A_3, \ldots, A_{n+2}$: The initial system of inequalities is still given by (2.2)–(2.7), but with s taking values in $\{3, 4, \ldots, n+2\}$ (see Fig. 2.1 (d)).

Since only one variable (i.e., p_{12}) has to be eliminated, at most two inequalities with different index s can be combined to give a valid inequality. Therefore, the final set of inequalities is given by (2.8)–(2.11) with the pair $3, 4$ substituted by any pair i, j with $i, j \in \{3, \ldots, n+2\}$ and $i < j$. This is precisely the result obtained in Refs. [89, 88].

2.2.3 Two parties, three settings

Now consider the bipartite scenario in which Alice can choose among three measurements, associated with propositions A_1, A_2 and A_3, and Bob can choose among three measurements, associated with propositions A_4, A_5 and A_6.

Analogously to the previous case, see Figure 2.2, the existence of a classical description for the six observables is equivalent to the existence of classical descriptions for the three subsystems, $\{A_1, A_2, A_3, A_4\}$, $\{A_1, A_2, A_3, A_5\}$ and $\{A_1, A_2, A_3, A_6\}$, coinciding on $\{A_1, A_2, A_3\}$. A probability on $\{A_1, A_2, A_3\}$ is completely defined once the probabilities $p_1, p_2, p_3, p_{12}, p_{13}, p_{23}, p_{123}$ are given.

It is therefore sufficient to calculate the correlation polytope associated with probabilities p_1, p_2, p_3, p_s, $p_{1s}, p_{2s}, p_{3s}, p_{12}, p_{13}, p_{23}, p_{123}$, then consider the system given by all the above inequalities for $s = 4, 5, 6$ and eliminate the variables $p_{12}, p_{13}, p_{23}, p_{123}$.

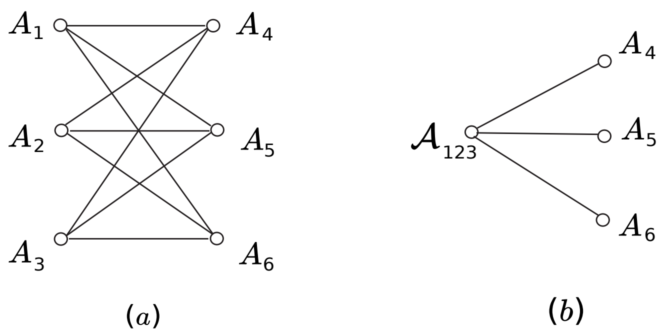


Figure 2.2: (a) Graph of the compatibility relations between the observables in the bipartite (3,3) scenario. (b) Tree graph obtained by extending the measure on the algebra generated by A_1, A_2, A_3.

2.2.4 Bipartite $(3, n)$ scenario

Again, by adding observables only on Bob's side, one just obtains more copies of the initial system of inequalities, but with different indices s. The situation is analogous to that depicted in Fig. 2.1 (d), but with $\mathcal{A}_{12}$ substituted by $\mathcal{A}_{123}$

In particular, since four variables have to be eliminated, at most $2^4 = 16$ inequalities with different indices s can be combined. As a result, all families of valid inequalities for the general case $(3, n)$ already arise in the case in which Alice performs 3 measurement and Bob 16.

2.2.5 Multipartite $(m, \ldots, m, n)$ scenario

The above argument can be extended to the case of p parties in which the first $p - 1$ can choose among m measurements, while the last one can choose among $n > m$ measurements: All families of inequalities can be obtained by studying the case in which the last experimenter performs 2^k measurements, where k is the number of variables to be eliminated.

2.2.6 Computational results

We have presented an alternative method for the computation of half-space representation for correlation polytopes based on algebraic conditions and variable elimination. Besides the explicit calculations presented above, a reasonable question is: Does our method provide any computational advantage with respect to existing methods? In order to show the advantages of the tree graph method, we have computed the $\mathcal{H}$-representation for some simple polytopes, with (i) our tree graph method using existing software implementing the Fourier-Motzkin algorithm; specifically, we used `porta` [72], and (ii) using standard software for solving the hull problem; specifically, we used `cdd`.

For simple cases like the $(2, 2)$ (i.e., the CHSH) and $(3, 3)$ scenarios, the computation is equally fast with both methods. However, remarkably, our tree graph method is noticeably faster to compute asymmetric scenarios:

For the $(3, 4)$ scenario, the tree graph method implemented with `porta` completed the calculation in ≈ 11 minutes, while `cdd` needed ≈ 20 minutes. The 11 minutes include the time (seconds) required to calculate the initial polytope (see Sec. 2.2.3).

For the $(3, 5)$ scenario, the tree graph method implemented with `porta` completed the calculation in ≈ 72 minutes, while `cdd` was still running after a week and we had to stop it.

All computations were performed on the same machine with an Intel Xeon CPU running at 3.20 GHz.

2.3 Noncontextuality inequalities: The n-cycle scenario

The n-cycle contextuality scenario is given by n observables $X_0, \ldots, X_{n-1}$ and the set of maximal contexts.

$$\mathcal{C}_n = \{\{X_0, X_1\}, \ldots, \{X_{n-2}, X_{n-1}\}, \{X_{n-1}, X_0\}\}. \tag{2.12}$$

This scenario is the natural generalization of CHSH and KCBS scenarios, the most fundamental scenarios for, respectively, nonlocality and contextuality. From the point of view of contextuality scenarios, the n-cycle scenario for odd n has been investigated by Liang, Spekkens, and Wiseman [92], who derived the inequality (2.17) for odd n and discuss the maximal quantum violation. As a bipartite Bell scenario, i.e., for even n, it has been investigate by Braunstein and Caves [93], who derived the inequality (2.17) for even n, later the maximal quantum violation was discussed [76].

In this section we give a complete characterization of the n-cycle polytope for any n, proving that the only facets inequalities of the polytope are given by Eqs.(2.13a)-(2.13d) together with Eq.(2.17).

A compatibility graph representation for the n-cycle scenario is given in Fig. 2.3. Here, it is convenient to use ± 1-valued variables instead of 0/1-valued. The corresponding correlations, i.e., expectation values of the product of the outcomes, are related to probabilities by the affine invertible transformation defined as

$$4p(++|X_iX_{i+1}) = 1 + \langle X_i \rangle + \langle X_{i+1} \rangle + \langle X_iX_{i+1} \rangle, \tag{2.13a}$$
$$4p(+-|X_iX_{i+1}) = 1 + \langle X_i \rangle - \langle X_{i+1} \rangle - \langle X_iX_{i+1} \rangle, \tag{2.13b}$$
$$4p(-+|X_iX_{i+1}) = 1 - \langle X_i \rangle + \langle X_{i+1} \rangle - \langle X_iX_{i+1} \rangle, \tag{2.13c}$$
$$4p(--|X_iX_{i+1}) = 1 - \langle X_i \rangle - \langle X_{i+1} \rangle + \langle X_iX_{i+1} \rangle. \tag{2.13d}$$

As a consequence, the analysis of the correlation polytopes (i.e., number of vertices and facets, tightness of a given inequality) defined in terms of ± 1-valued or 0/1-valued variables are equivalent.

The n-cycle correlation polytope is then defined by the following 2^n vertices

$$(x_0, \ldots, x_{n-1}, x_0x_1, \ldots, x_{n-1}x_0), \quad x_i \in \{-1, 1\}. \tag{2.14}$$

In order to differentiate this with the probability case, we use the following notation for a general vector

$$(\langle X_0 \rangle, \ldots, \langle X_{n-1} \rangle, \langle X_0X_1 \rangle, \ldots, \langle X_{n-1}X_0 \rangle). \tag{2.15}$$

A set of tight inequalities is already given by the positivity conditions on the terms in Eq.(2.13), namely

$$1 + \langle X_i \rangle + \langle X_{i+1} \rangle + \langle X_iX_{i+1} \rangle \geq 0 \tag{2.16a}$$
$$1 + \langle X_i \rangle - \langle X_{i+1} \rangle - \langle X_iX_{i+1} \rangle \geq 0 \tag{2.16b}$$
$$1 - \langle X_i \rangle + \langle X_{i+1} \rangle - \langle X_iX_{i+1} \rangle \geq 0 \tag{2.16c}$$
$$1 - \langle X_i \rangle - \langle X_{i+1} \rangle + \langle X_iX_{i+1} \rangle \geq 0. \tag{2.16d}$$

one can easily proved that they are saturated by a set of affinely independent vertices (2.14). Such conditions are trivially satisfied in QM since they involve only pairs of jointly measurable observables.

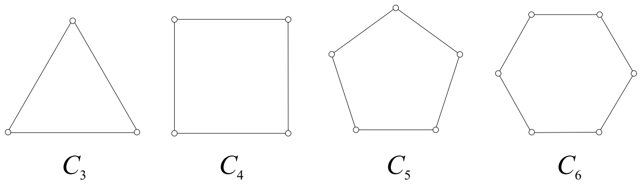


Figure 2.3: Graphs associated to the compatibility relations among the observables X_i for $n = 3, 4, 5, 6$. C_4 corresponds to CHSH case with the labelling of nodes A_1, B_1, A_2, B_2, in the usual notation for Alice and Bob observables, and C_5 corresponds to KCBS case with the labelling $X_0, \ldots, X_4$.

The remaining facets of the n-cycle polytope are given by the 2^{n-1} inequalities

$$\Omega = \sum_{i=0}^{n-1} \gamma_i \langle X_i X_{i+1} \rangle \overset{\text{NCHV}}{\leq} n-2, \tag{2.17}$$

where $\gamma_i \in \{-1, 1\}$ such that the number of $\gamma_i = -1$ is odd.

Our proof is based on the fact that the existence of a classical probability model for the observables $\{X_0, \ldots, X_{n-1}\}$ is equivalent to the existence of classical models for $\{X_0, \ldots, X_{n-2}\}$ and $\{X_0, X_{n-1}, X_{n-2}\}$, coinciding on their intersection $\{X_0, X_{n-2}\}$, see Fig. 2.4 (a),(b),(c). More precisely, if the two subsets of observables in Fig.2.4 (b) $(n-1)$-cycle and 3-cycle, admit a classical representation, *i.e.*all the corresponding inequalities are satisfied, then the set of probabilities can be extended, following Prop. 1,2, as in Fig. 2.4 (c), *i.e.*two classical model for $\{0, 1, \ldots, n-3, n-2\}$ and for $\{0, n-1, n-2\}$ coinciding on their intersection $\{0, n-2\}$. By Prop. 2, such a set already admits a classical representation.

Such a consistency condition for the intersection is written in terms of the "unmeasurable correlation" $\langle X_0 X_{n-2} \rangle$, *i.e.*, a correlation between observables that are not in a context and therefore cannot be jointly measured, but have nevertheless a well-defined correlation in every classical model [91]. The final set of inequalities must not contain the variable $\langle X_0 X_{n-2} \rangle$, which must be removed by applying Fourier-Motzkin (FM) elimination.

We can now proceed by induction on n. The case $n = 3$ is known. For the inductive step, following the above argument, we calculate the n-cycle inequalities by combining the $(n-1)$-cycle inequalities for the subset $\{X_0, \ldots, X_{n-2}\}$ with the 3-cycle inequalities for $\{X_0, X_{n-1}, X_{n-2}\}$. We apply FM elimination on the variable $\langle X_0 X_{n-2} \rangle$ from the whole set of inequalities. All inequalities in (2.17) are obtained by combining one inequality, of the same form, for the $(n-1)$-cycle with one for the 3-cycle, and are in the right number. In fact, in half of the $(n-1)$-cycle inequalities, $\langle X_0 X_{n-2} \rangle$ appears with the $+$ sign, and in half with the $-$ sign, and the same for the 3-cycle. The number of possible combination is, therefore, given by $2^{n-3} \cdot 2 + 2^{n-3} \cdot 2 = 2^{n-1}$.

Combining two inequalities for the $(n-1)$-cycle, or two for the 3-cycle gives a redundant inequality, as happens for combination of positivity conditions (2.16) with inequalities of the form (2.17), the latter being obtainable as a sum of $n-1$ (or 3) positivity conditions. There are no other inequalities.

Tightness can be proved by showing that inequalities (2.17) correspond to facets of the $2n$-dimensional correlation polytope, *i.e.*, they are saturated by $2n$ noncontextual vertices which generate an affine subspace of dimension $2n-1$. First, focus on the inequality of the odd n-cycle for which all $\gamma_i = 1$. It is saturated by $2n$ vertices which can be written as $(\pm v_i, w_i)$, for $i = 0, \ldots, n-1$, where v_i is a n-dimensional vector given by a cyclic permutation of the components of $v_{n-1} = (+1, -1, +1, \ldots, +1)$ and w_i's components are given by the corresponding products, of v_i's components, namely, one component equal to 1 and $n-1$ components equal to -1. Up to a

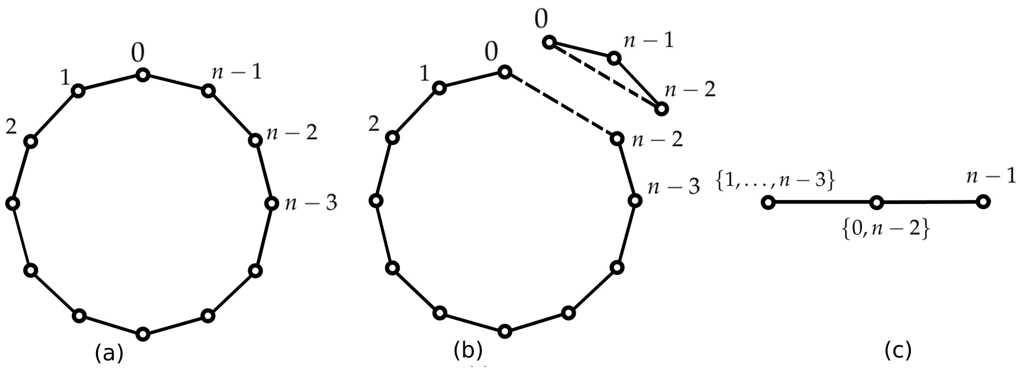


Figure 2.4: (a) n-cycle scenario. (b) subsets of observables that can be associated with the $(n-1)$-cycle and 3-cycle scenario by considering the "unmeasurable correlation" $\langle X_0 X_{n-2}\rangle$ (dashed line). (c) Extended classical model that can be obtained if the two subset admits a classical representation coinciding on their intersection. Such a model is automatically classical as it can be depicted as a tree graph.

reordering of the vectors, it holds $v_i + v_{i+1} = 2e_{i+1}$, with addition modulo n and $e_0, \ldots, e_{n-1}$ the canonical basis of $\mathbb{R}^n$, and $w_i + (1, 1, \ldots, 1) = 2e_i$. As a consequence, $\{(\pm v_i, w_i)\}_{i=1,\ldots,2n}$ is a basis of $\mathbb{R}^{2n}$, and, therefore, such vectors generate an affine subspace of dimension $2n - 1$. Since all the other vertices and inequalities are obtained from this one via the mapping $X_i \mapsto -X_i$, we are done for the odd case. For the even case, the proof is slightly different: Take the inequality which has $\gamma_0 = -1$ and all the other $\gamma_i = 1$. Again, it is saturated by $2n$ vertices which can be written as $(\pm v_i, w_i)$, for $i = 0, \ldots, n-1$, with $v_0 = \sum_{k=0}^{n-1}(-1)^k e_k$, $w_0 = -\sum_{k=0}^{n-1} e_k$ and

$$v_i = e_0 + e_1 + \sum_{2\le k\le i} (-1)^k e_k + \sum_{i<k\le n-1} (-1)^{k+1} e_k, \tag{2.18}$$

$$w_i = e_0 + e_i - \sum_{k\neq 0,i} e_k, \qquad for\ i = 1, \ldots, n-1. \tag{2.19}$$

The above vectors satisfy $v_{n-1} - v_0 = 2e_0$, $v_1 - v_0 = 2e_1$, and $v_i - v_{i+1} = (-1)^{i+1} e_{i+1}$, for $i = 1, \ldots, n-2$, and $w_i + w_j - 2w_0 = 4e_0$, $w_i - w_0 = 2(e_0 + e_i)$ for $i, j = 1, \ldots, n-1$. Again, this implies that the vertices are affinely independent, and all the other vertices and inequalities are generated via the relabelling $X_i \mapsto -X_i$

2.4 Discussion

We developed an alternative method for the computation of the correlation polytope associated with Bell and noncontextuality scenarios based on results on automatic extension of probability measures and Fourier-Motzkin algorithm of variable elimination for systems of linear inequalities.

We applied our method to different Bell and noncontextuality scenarios. For the Bell scenario, we derive some result on the minimal computation needed for a completely characterization of asymmetric scenarios where one experimenter has an arbitrary number of settings, thus generalizing a result by Śliwa [88] and Collins-Gisin [89]. Moreover, we performed some explicit computations to show the advantages with respect to usual algorithms solving directly the hull problem.

For noncontextuality scenario, we are able to give a complete characterization, i.e., the complete set of tight inequalities, for the n-cycle scenario, which is the natural generalization of the two most fundamental scenarios, respectively, for nonlocality and contextuality, namely, the CHSH and

KCBS scenario. To complete the discussion on the n-cycle scenario, in Chapt. 4, we will compute the quantum bound by means of Cabello-Severini-Winter method [81, 82].

Chapter 3

Optimal tests for state-independent contextuality

In the previous chapter, we discussed the problem of finding tight inequalities for a given measurement scenario and the computational difficulties associated with it. The design of tests of quantum versus HV theories usually involves as a first step the derivation of tight inequalities, i.e., the optimal inequalities corresponding to the boundaries of the classical set of probabilities, and subsequently the search for a quantum state and observables giving the maximal violation.

Despite the existence of alternative approaches, still in many relevant cases such inequalities cannot be computed and the optimization over the set of states and observables cannot be performed. In particular for SIC scenarios, even the simplest case, i.e., Yu and Oh's [56] scenario, is too complex to allow for such an approach.

In this chapter, we approach the problem from a different perspective. Our idea comes from the observation that for SIC scenarios the set of measurements is known, thus there is no need for an optimization over the set of observables, and the violation is the same for any state, thus there is no need for an optimization over the set of quantum states. The result is an optimization method for noncontextuality inequalities based on linear programming that provides inequalities with the maximal quantum violation. Such a problem can be solved efficiently with standard numerical techniques and with the optimality of the solution guaranteed.

We shall apply our method to some of the most important SIC scenarios showing the advantages of our improved inequalities. Once such inequalities are computed, one can easily checked for their tightness, i.e., whether they correspond to facets of the associated polytope. We find that in all the cases we analyzed, our optimal inequalities, in the sense of a maximal quantum violation, are also tight with respect to the associated polytope. We also discuss a possible generalization of the method that search for inequalities violated by any state, but with a different degree of violation. We shall also discuss the proper way of performing such texts experimentally. The results of this chapter have been published in Ref. [94], a brief discussion of the meaning of experimental test for contextuality have been presented in [95].

3.1 Optimization method

Given the increased complexity of the measurement scenarios in the case of SIC, we need a more compact notation with respect to the simple one used in the previous chapter. In addition, in order to have a direct comparison between the original SIC inequality proposed by previous authors and our improved version, we shall consider ± 1-valued observables instead of 0/1-valued, as we did for the n-cycle scenario in Sect. 2.3. Given some dichotomic quantum observables $A_1, A_2, \dots, A_n$, we denote with $\underline{c}$ a context, i.e., a set of indices, such that A_k and A_ℓ are compatible whenever $k, \ell \in \underline{c}$, i.e., $[A_k, A_\ell] = 0$. Notice that valid contexts are also those defined by a single observable, i.e., with $|\underline{c}| = 1$.

As we shall see below, it may be interesting to consider only a certain admissible subset $\mathcal{C}$ of the set of all possible contexts $\{\underline{c}\}$. The contextuality scenario will be defined by the observables $A_1, \ldots, A_n$, together with the list of admissible contexts $\mathcal{C}$.

The set of all (contextual as well as noncontextual) correlations for such a scenario can be represented by the standard construction in terms of vectors $\vec{v} = (v_{\underline{c}} \mid \underline{c} \in \mathcal{C})$, where $v_{\underline{c}}$ is the expectation value of the product of the values of the observables indexed by $\underline{c}$.

In a NCHV model, each observable has a fixed assignment $\vec{a} \equiv (a_1, \ldots, a_n) \in \{-1, 1\}^n$ for the observables $A_1, \ldots, A_n$, and accordingly each entry in $\vec{v}$ is exactly the product of the assigned values, i.e., $v_{\underline{c}} = \prod_{k\in\underline{c}} a_k$. The most general noncontextual HV model predicts fixed assignments $\vec{a}^{(i)}$ with probabilities p_i, and hence the set of correlations that can be explained by a noncontextual HV models is characterized by the convex hull of the models with fixed assignments, thus forming the noncontextuality polytope.

Then, a noncontextuality inequality is an affine bound on the noncontextuality polytope, i.e., a real vector $\vec{\lambda}$ such that $\eta \geq \vec{\lambda} \cdot \vec{v}$ for all correlation vectors v that originate from a noncontextual model:

$$\eta \geq \sum_{\underline{c}\in\mathcal{C}} \lambda_{\underline{c}} \prod_{k\in\underline{c}} a_k, \tag{3.1}$$

for any assignment $\vec{a} \equiv (a_1, \ldots, a_n) \in \{-1, 1\}^n$.

In quantum mechanics, in contrast, the measurement of the entry $v_{\underline{c}}$ corresponds to the expectation value $\langle \prod_{k\in\underline{c}} A_k \rangle_\rho$, where ρ specifies the quantum state. Thus the value of $\vec{\lambda} \cdot \vec{v}$ predicted by quantum mechanics is given by $\langle T(\vec{\lambda}) \rangle_\rho$, with

$$T(\vec{\lambda}) = \sum_{\underline{c}\in\mathcal{C}} \lambda_{\underline{c}} \prod_{k\in\underline{c}} A_k. \tag{3.2}$$

If the expectation value exceeds the noncontextual limit η, then the inequality demonstrates contextual behaviour, yielding the quantum violation

$$\mathcal{V} = \frac{\max_\rho \langle T(\vec{\lambda}) \rangle_\rho}{\eta} - 1. \tag{3.3}$$

For a given contextuality scenario, we can define optimal inequalities in the sense of a maximal value for the parameter $\mathcal{V}$. We shall discuss the relation between such a notion of optimality with the usual notion of tightness.

We recall that, in general, this optimization is difficult to perform and it is not always clear that an optimal inequality also yields the most significant violation [96]. However, if we require a state independent violation of the inequality, without loss of generality, $T(\vec{\lambda}) = \mathbb{1}$ and hence the optimization over the quantum state ϱ vanishes. Then, the coefficient vector $\vec{\lambda}$ and the noncontextual bound η are optimal if η is minimal under the constraint $T(\vec{\lambda}) = \mathbb{1}$ and if the inequalities in Eq. (3.1) are satisfied. That is, we ask for a solution $(\eta^*, \vec{\lambda}^*)$ of the optimization problem

$$\begin{aligned} \text{minimize:} \quad & \eta, \\ \text{subject to:} \quad & T(\vec{\lambda}) = \mathbb{1} \text{ and} \\ & \eta \geq \sum_{\underline{c}\in\mathcal{C}} \lambda_{\underline{c}} \prod_{k\in\underline{c}} a_k, \text{for all } \vec{a}. \end{aligned} \tag{3.4}$$

This optimization problem is a linear program and such programs can be solved efficiently by standard numerical techniques and optimality is then guaranteed. We implemented this optimization using CVXOPT [97] for Python, which allows us to study inequalities with up to $n = 21$ observables and $|\mathcal{C}| = 131$ contexts. Note that this program also solves the feasibility problem, whether a contextuality scenario exhibits SIC at all. This is the case, if and only if the program finds a solution with $\eta < 1$ and thus $\mathcal{V} > 0$.

The optimal coefficients $\vec{\lambda}^*$ are, in general, not unique but rather form a polytope defined by the system of inequalities

$$\eta^* \geq \sum_{\underline{c}\in\mathcal{C}} \lambda_{\underline{c}} \prod_{k\in\underline{c}} a_k, \tag{3.5}$$

for any $\vec{a} \equiv (a_1, \ldots, a_n) \in \{-1, 1\}^n$ and the maximal value η^*. This leaves the possibility to find optimal inequalities with further special properties. There are at least two important properties that one may ask for. Firstly, from an experimental point of view, it would be desirable to have some of the coefficients $\lambda_{\underline{c}} = 0$, since then the context $\underline{c}$ does not need to be measured. In general, it will depend on the experimental setup, which coefficients $\lambda_{\underline{c}} = 0$ yield the greatest advantage. For the sequential measurement schemes it is natural to choose the longest measurement sequences. Secondly, there might be *tight* inequalities among the optimal solutions: An inequality is tight, if the affine hyperplane given by the solutions of $\eta = \vec{\lambda}\cdot\vec{x}$ is tangent to a facet of the noncontextuality polytope. This property can be readily checked using Pitowsky's construction [14]: Denote by d the affine dimension of the noncontextuality polytope and choose those assignments $\vec{a}$, for which Eq. (3.1) is saturated. Then, the inequality is tangent to a facet if and only if the affine space spanned by the vertices $\vec{v} \equiv (\prod_{k\in\underline{c}} a_k \mid \underline{c} \in \mathcal{C})$ is $(d-1)$-dimensional.

Furthermore, we mention that the condition of state independence might be loosened to only require that the quantum violation is *at least* $\mathcal{V}$ for all quantum states. This corresponds to replacing the condition $T(\vec{\lambda}) = \mathbb{1}$ by the condition that $T(\vec{\lambda}) - \mathbb{1}$ is positive semidefinite, namely

$$\begin{aligned} \text{minimize:} \quad & \eta, \\ \text{subject to:} \quad & T(\vec{\lambda}) - \mathbb{1} \geq 0 \text{ and} \\ & \eta \geq \sum_{\underline{c}\in\mathcal{C}} \lambda_{\underline{c}} \prod_{k\in\underline{c}} a_k, \text{for all } \vec{a}. \end{aligned} \tag{3.6}$$

Then, the linear program in Eq. (3.4) becomes the semidefinite program (3.6), which still can be solved by standard numerical methods, e.g., CVXOPT [97] for Python, with optimality of the solution guaranteed. However, for the examples that we consider in the following, the semidefinite and the linear program yield the same results, namely, that the every state gives rise to the same violation of the inequality.

3.2 Applications

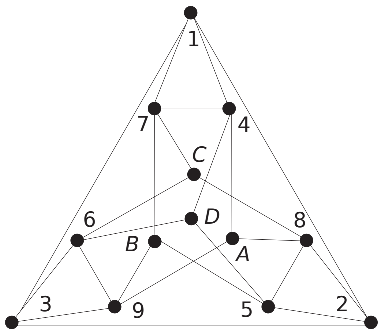
3.2.1 Yu and Oh

We now apply our method to the SIC scenario for a qutrit system introduced by Yu and Oh [56]. Qutrit systems are of fundamental interest, since no smaller quantum system can exhibit a contextual behavior [4, 5]. It has been shown that this scenario is the simplest possible SIC scenario for a qutrit [98, 99].

For a qutrit system, we consider dichotomic observables of the form

$$A_i = \mathbb{1} - 2|v_i\rangle\langle v_i|, \tag{3.7}$$

and in the Yu-Oh scenario, there are 13 observables defined by the 13 unit vectors $|v_i\rangle$ provided in Fig. 3.1.



$$
\begin{array}{lll}
v_1 = (1,0,0) & v_5 = (1,0,-1) & v_A = (-1,1,1) \\
v_2 = (0,1,0) & v_6 = (1,-1,0) & v_B = (1,-1,1) \\
v_3 = (0,0,1) & v_7 = (0,1,1) & v_C = (1,1,-1) \\
v_4 = (0,1,-1) & v_8 = (1,0,1) & v_D = (1,1,1) \\
 & v_9 = (1,1,0) &
\end{array}
$$

Figure 3.1: Graph of the compatibility relations between the observables for the Yu-Oh scenario. Nodes represent vectors $|v_i\rangle$, or the observables A_i defined in (3.7), and edges represent orthogonality, or compatibility, relations.

$\underline{c}$	YO	opt_2	opt_3	$\underline{c}$	YO	opt_2	opt_3	$\underline{c}$	YO	opt_2	opt_3
1	2	2	1	A–D	2	1	2	3, 9	−1	−2	−1
2	2	3	1	1,2	−1	−1	−2	4, 7	−1	0	−1
3	2	3	1	1,3	−1	−1	−2	5, 8	−1	−2	−1
4	2	1	1	1,4	−1	−1	−1	6, 9	−1	−2	−1
5	2	2	1	1,7	−1	−1	−1	$*$,A–D	−1	−1	−2
6	2	2	1	2,3	−1	−2	−2	1,2,3	–	–	0
7	2	1	1	2,5	−1	−2	−1	1,4,7	–	–	−3
8	2	2	1	2,8	−1	−2	−1	2,5,8	–	–	−3
9	2	2	1	3,6	−1	−2	−1	3,6,9	–	–	−3

Table 3.1: Coefficients $\lambda_{\underline{c}}$ of inequalities for the Yu-Oh scenario. The column $\underline{c}$ labels the different contexts, YO the coefficients in the inequality of Ref. [56], opt_2 an optimal tight inequality with contexts of maximal size 2, opt_3 an optimal tight inequality with contexts of all sizes. For compactness, the coefficients in the column YO have been multiplied by 50/3, for the column opt_2 by 52/3 and for the column opt_3 by 83/3. The row labelled "A–D" shows the coefficients for the contexts $\{A\}$, $\{B\}$, $\{C\}$, $\{D\}$ and the row labelled "$*$,A–D" shows the coefficients for $\{4, A\}$, $\{8, A\}$, $\{9, A\}$, $\{5, B\}$, $\{7, B\}$, $\{9, B\}$, $\{6, C\}$, $\{7, C\}$, $\{8, C\}$, $\{4, D\}$, $\{5, D\}$, $\{6, D\}$.

In the according graph, each operator is represented by node $i \in V$ of the graph $G = (V, E)$ and an edge $(i, j) \in E$ indicates that $|v_i\rangle$ and $|v_j\rangle$ are orthogonal, $|v_j\rangle\langle v_i| = 0$, so that A_i and A_j are compatible. The original inequality takes into account all contexts of size one and two, $\mathcal{C}_{\text{YO}} = \{\{1\}, \ldots, \{D\}\} \cup E$ and the coefficients were chosen to $\lambda_{\underline{c}} = -3/50$ if $\underline{c} \in E$ and $\lambda_{\underline{c}} = 6/50$ else. This yields an inequality with a state-independent quantum violation of $\mathcal{V} = 1/24 \approx 4.2\%$.

With the linear program we find that the maximal violation for the contexts $\mathcal{C}_{\text{YO}}$ is $\mathcal{V} = 1/12 \approx 8.3\%$ and thus twice that of the inequality in Ref [56]. Interestingly, among the optimal coefficients $\vec{\lambda}^*$ there is a solution which is tight and for which the coefficients $\lambda_{4,7}$ vanishes, cf. Table 3.1, column "opt$_2$" for the list coefficients. We find that up to symmetries, $\lambda_{4,7}$ is the only context that can be omitted while still preserving optimality.

The maximal contexts in the Yu-Oh scenario are of size three, and hence it is possible to include also the corresponding terms in the inequality, i.e., we extend the contexts $\mathcal{C}_{\text{YO}}$ by the contexts $\{1,2,3\}$, $\{1,4,7\}$, $\{2,5,8\}$, and $\{3,6,9\}$. Since this increases the number of parameters in the inequality, there is a chance that this case allows an even higher violation. In fact, the maximal violation is $\mathcal{V} = 8/75 \approx 10.7\%$. Again, it is possible to find tight inequalities with vanishing coefficients, and in particular the context $\{1,2,3\}$ can be omitted; the list of coefficients is given in Table 3.1, column "opt$_3$".

Using Pitowsky's approach, checking the tightness of an inequality amounts to a problem of affine independence that can be easily transformed in a linear independence problem and solved efficiently, e.g., by Gaussian elimination. We verified that the inequalities opt$_2$ and opt$_3$ are tight with respect to the corresponding correlation polytope, wheareas YO is not tight.

3.2.2 Extended Peres-Mermin set

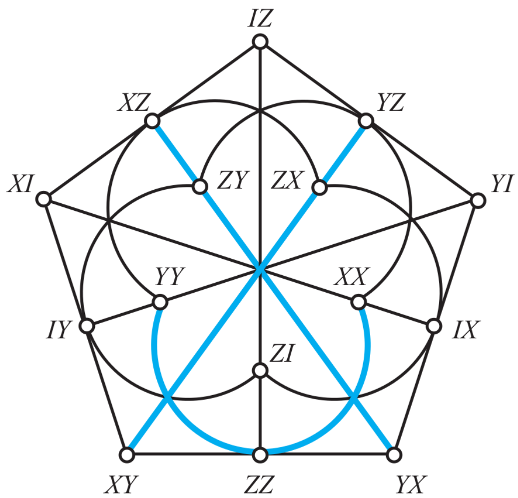


Figure 3.2: The set of observables represented by two-fold tensor products of Pauli matrices except for the identity has 15 ob- servables (represented by points in the graph; the notation is $XI = \sigma_x \otimes \mathbb{1}$) which can be grouped in 15 sets of compatible observables (represented by lines; collinear points correspond to compatible observables). Each point is on three lines and each line is incident with three points. Black (cyan) lines represent sets in which the product of all the matrices is $\mathbb{1}$ ($-\mathbb{1}$). It is impossible to assign noncontextual results 1 or +1 to all the 15 observables in agreement with the predictions of QM: The product of the results of a black (cyan) line must be +1 (1)

The *extended Peres-Mermin set* uses as observables all 15 products of Pauli operators on a two-qubit system. Observables and compatibility relations are depicted in Fig. 3.2.2. The original SIC inequality proposed for this scenario [100] is given by

$$\sum_{\underline{c} \in \mathcal{C}^+} \langle \prod_{k \in \underline{c}} A_k \rangle - \sum_{\underline{c} \in \mathcal{C}^-} \langle \prod_{k \in \underline{c}} A_k \rangle \leq 9, \tag{3.8}$$

where $\mathcal{C}^+$ is the set of contexts denoted by black lines in Fig.3.2.2, i.e., contexts such that the product of the observables gives $+\mathbb{1}$, and $\mathcal{C}^-$ is the set of contexts denoted by cyan lines, i.e., such that the product of the observables gives $-\mathbb{1}$.

In the language used above, the optimal violation is $\mathcal{V} = 2/3$, where only contexts of size three need to be measured and, with the proper normalization, $\lambda_{\underline{c}} = 1/15$, except $\lambda_{xx,yy,zz} = \lambda_{xz,yx,zy} = \lambda_{xy,yz,zx} = -1/15$. We verified that (3.8) is optimal and among the optimal solutions no simpler inequality exists. Also in this case the inequality is corresponds to a facet of the associated polytope.

3.2.3 Cabello, Estebaranz, and García-Alcaine's 18-vector proof

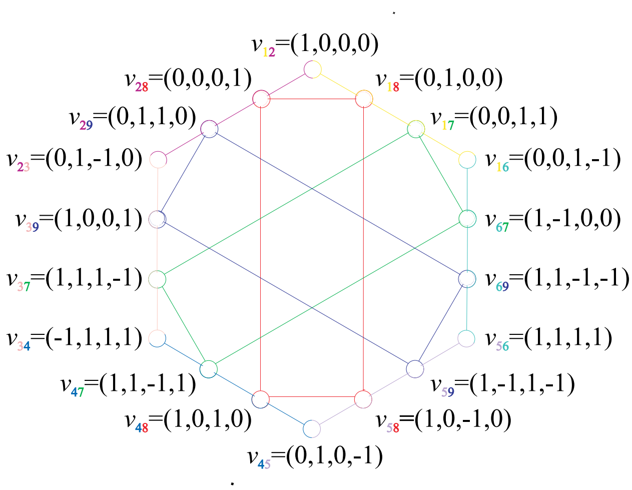


Figure 3.3: Graphical representation for the contexts in the 18-vector KS set of Ref. [47]. Each side of the regular hexagon and each square represent a context, vectors in each context are mutually orthogonal.

The KS set provided by Cabello *et al.* [47], uses a four-level system and 18 observables derived as in Eq. (3.7) for the 18 (unnormalized) vectors depicted in Fig.3.3. For contexts up to size 2 the maximal violation is $\mathcal{V} = 1/17 \approx 5.9\%$ and is given by the inequality (cf. [101]),

$$\sum_i \langle P_i \rangle - \frac{1}{2} \sum_{ij} \langle P_i P_j \rangle \leq 4, \tag{3.9}$$

where $P_i = |v_i\rangle\langle v_i|$ for the (normalized) vectors v_i of Fig. 3.3, and the sum is over all possible contexts of size 2.

Including all contexts the maximal violation is $\mathcal{V} = 2/7 \approx 28.6\%$ and is given by the inequality (cf. [51])

$$\sum_{maximal\ \underline{c}} \langle \prod_{i \in \underline{c}} A_i \rangle \leq 7, \tag{3.10}$$

where $A_i = 2P_i - \mathbb{1}$, and the sum is over the 18 maximal contexts of four observables depicted in Fig.3.3.

Both inequalities (3.9),(3.10) can be proven to be tight.

The situation where only contexts up to size 3 are admissible has not yet been studied and we find numerically a maximal violation of $\mathcal{V} \approx 14.3\%$.

3.3 Experimental tests

The main motivation for the introduction of the notions of compatible measurements and contexts is given by the property of commuting observables. As we discussed in Chapt. 1, compatible observable not only can be measured jointly, but also in sequence in any order and repeatedly multiple times always giving, in the ideal case, the same result as the initial measurement.

For an experimental test of noncontextuality, two possible measurement schemes naturally arise, namely

a) the joint measurement,

b) the sequential measurement.

The possible advantages and disadvantages of both methods from the point of view of experimental tests have been discussed in Ref. [102], we briefly recall here the main argument of the authors.

In the joint measurement scenario, each single device corresponds to an entire context, and the outcomes must specify the outcome of each of the observables that is jointly measured. The only way to argue that such single-observable outcomes are independent of the device used is to repeat multiple times and in different order the joint measurement and the single-observable measurement. Precisely because such a single-observable device is able to reproduce the result of the corresponding joint measurements, i.e., those involving such an observable, there is good reason to assume that the outcome is independent of the joint measurement performed, i.e., of the context.

In the sequential measurement scenario, the experimenter starts already with single-observable measurements. Such measurement device can be tested directly in a similar way as above, i.e., by repeating the measurements many times and in different order and verify whether the results are unchanged.

The physical motivation for unchanged results in the sequential measurement scenario is much stronger since the devices used are identical in different contexts. On the other hand, the joint measurement device might correspond to a completely different experimental setup even for the unchanged setting of a context, so it is more difficult to maintain that the outcome for the unchanged setting is unchanged.

Moreover, in the sequential measurement scheme it is possible to discuss the influence from one measurement to the subsequent, and it also possible to perform a test (albeit limited) of such influence, and to include it in the analysis as correction terms to ideal NCHV bound, as discussed in Ref. [102].

Independently of the chosen measurement scheme, it is clear from the above discussion that it is a fundamental requirement for any experimentalist to convincingly argue that the devices used in the experiment produce a context-independent outcome.

An indicative example of this approach is given by the experiment of Kirchmair *et al.* [53], where the Peres-Mermin inequality is tested via sequential measurements. Here, each observable is implemented by an experimental procedure, namely, a sequence of non-local unitaries and flourescence detection, which is always the same in every measurement context. Analogously, the state preparation procedure is always the same independently of the sequence of measurements to be performed afterwards.

Another example, where such context-independent experimental procedures are not carefully followed, is given by the test of Yu and Oh's inequality by Zu *et al.* [103]. Here, observables are implemented by different experimental procedures, depending on the measurement context, and even the initial state is changed in a context-dependent way. In particular, the measurement of each of the observables A_A, A_B, A_C, A_D of Fig. 3.1 is performed always in a different way. See also the discussion in [95].

Another test of the Yu and Oh's inequality, performed with trapped ions and following the same procedure as Ref. [53], has then been performed by Zhang *et al.* [104]. Notice, however, that the first preprint submitted to the arXiv [105] presented the same problems as the experiment in Ref. [103].

3.4 Discussion

Among the most striking aspects where contextuality is more general than nonlocality is that the former can be found to be independent of the quantum state. For this state-independent scenario, we showed that the search for the optimal inequality reduces to a linear program, which can be solved numerically with optimality guaranteed. We studied several cases of this optimization and find that in all those instances one can construct noncontextuality inequalities with a state independent violation that are, in addition, tight. This is in particular the case for the most fundamental scenario of state independent contextuality [56] and we presented two essentially different inequalities—one involves at most contexts of size two, the other of size three. We hence lifted the Yu-Oh scenario to the same fundamental status as the CHSH Bell inequality [26, 69], which is the simplest scenario for nonlocality.

Chapter 4

Quantum bounds for temporal correlations

The assumptions of realism and locality lead to bounds on the correlations between observable quantities —the Bell inequalities, and these bounds are violated in quantum mechanics. Interestingly, this quantum violation is limited for many Bell inequalities and does not reach the maximal possible value. For instance, the CHSH inequality bounds the correlation [26]

$$\mathcal{B} = \langle A_1 \otimes B_1 \rangle + \langle A_1 \otimes B_2 \rangle + \langle A_2 \otimes B_1 \rangle - \langle A_2 \otimes B_2 \rangle, \tag{4.1}$$

where A_i and B_j are measurements on two different particles. On the one hand, local realistic models obey the CHSH inequality $\mathcal{B} \leq 2$, which is violated in quantum mechanics. On the other hand, the maximal quantum value is upper bounded by Tsirelson's bound [74] $\mathcal{B} \leq 2\sqrt{2}$. Whereas this bound holds within quantum mechanics, it has turned out that hypothetical theories that reach the algebraic maximum $\mathcal{B} = 4$ without allowing faster-than-light communication are possible [37]. This raises the question of whether the bounded quantum value can be derived on physical grounds from fundamental principles. Partial results are available, and principles have been suggested that bound the correlations [41, 43, 106, 107, 108] (see also the discussion in Ref. [109]).

The question of how and why quantum correlations are fundamentally limited has been discussed mainly in the scenario of bipartite and multipartite measurements. What happens, however, if we shift the attention from spatially separated measurements to temporally ordered measurements? There is no need to measure on distinct systems as in Eq. (4.1), but rather, we may perform sequential measurements on the same system. Then, an elementary property of quantum mechanics becomes important: The measurement changes the state of the system. In fact, this allows us to temporally "transmit" a certain amount of information [23], and one would expect that the correlations in the temporal case can be larger than in the spatial situation.

We stress that sequential measurements also have been considered in the analysis of quantum contextuality (the Kochen-Specker theorem [5]) and macrorealism (Leggett-Garg inequalities [8]). The research in this fields has triggered experiments involving sequential measurements. For demonstrating such a contradiction between classical and quantum physics, e.g., the KCBS correlation, appearing in Eq. (1.10),

$$\begin{aligned} \mathcal{S}_5 = & \langle A_1 A_2 \rangle_{\text{seq}} + \langle A_2 A_3 \rangle_{\text{seq}} + \langle A_3 A_4 \rangle_{\text{seq}} + \langle A_4 A_5 \rangle_{\text{seq}} \\ & - \langle A_5 A_1 \rangle_{\text{seq}} \end{aligned} \tag{4.2}$$

has been considered [34, 110]. Here, $\langle A_i A_j \rangle_{\text{seq}}$ denotes a sequential expectation value that is the average of the product of the value of the observables A_i and A_j when first A_i is measured, and afterwards A_j. One can show that for macrorealistic theories as well as for noncontextual models the bound $\mathcal{S}_5 \leq 3$ holds, but in quantum mechanics, this can be violated.

Here however, we are rather interested in the fundamental bounds on the temporal quantum correlations, with no assumption about the compatibility of the observables. Special cases of this problem have been discussed before: For Leggett-Garg inequalities, maximal values for two-level systems have been derived [110, 111], and temporal inequalities similar to the CHSH inequality have been discussed [23, 112].

In this chapter, we provide a method that allows us to compute the maximal achievable quantum value for an arbitrary inequality and thus we solve the problem of bounding temporal quantum correlations. First, we will discuss a simple method, which can be used for expressions as in Eq. (4.2), where only sequences of two measurements are considered. Then, we introduce a general method which can be used for arbitrary sequential measurements, resulting in a complete characterization of the possible quantum values. Interestingly, our methods characterize temporal correlations exactly, whereas for the case of spatially separated measurements only converging approximations are known.

For the convenience of the reader, the technical details are collected in Sect. 4.5. The results of this chapter have been published in [24].

4.1 Sequential projective measurements

When determining the maximal value for sequential measurements as in Eq. (4.2) we consider projective measurements, as these are the standard textbook examples of quantum measurements. The underlying formalism has been established by von Neumann [19] and Lüders [57]. According to Lüders' rule, an observable A with possible results ± 1 is described by two projectors Π_+ and Π_- such that $A = \Pi_+ - \Pi_-$. If the observable A is measured, the quantum state is projected onto the space of the observed result, namely,

$$\varrho \mapsto \Pi_\pm \varrho \Pi_\pm / \mathrm{tr}(\varrho \Pi_\pm). \tag{4.3}$$

For the moment, we restrict to rule 4.3. We shall see in the next chapter, the consequences of using a more general state-update rule, e.g., von Neumann's original proposal [19], and how such different state-update rules for projective measurements can be exploited to construct dimension witnesses.

Applying the above scheme to the case of sequential measurements, one finds that the sequential mean value can be written as

$$\langle A_i A_j \rangle_{\mathrm{seq}} = \mathrm{tr}(\varrho \Pi^i_+ \Pi^j_+ \Pi^i_+) + \mathrm{tr}(\varrho \Pi^i_- \Pi^j_- \Pi^i_-) - \mathrm{tr}(\varrho \Pi^i_- \Pi^j_+ \Pi^i_-) - \mathrm{tr}(\varrho \Pi^i_+ \Pi^j_- \Pi^i_+) \tag{4.4}$$

With the substitution $\Pi^{i,j}_\pm = (\mathbb{1} \pm A_{i,j})/2$, it becomes

$$\langle A_i A_j \rangle_{\mathrm{seq}} = \frac{1}{2}[\mathrm{tr}(\varrho A_i A_j) + \mathrm{tr}(\varrho A_j A_i)]. \tag{4.5}$$

It is interesting to notice that for pairs of ± 1-valued observables such a mean value does not depend on the order of the measurement [23].

4.2 The simplified method

We first show how the maximal quantum mechanical value for an expression such as $\mathcal{S}_5$ in Eq. (4.2) can be determined. First, we consider a set $\mathcal{A} = \{A_i\}$ of ± 1-valued observables and a general expression $C = \sum_{ij} \lambda_{ij} \langle A_i A_j \rangle_{\text{seq}}$. The correlations given in Eq. (4.2) are just a special case of this scenario. Then, we consider the matrix built up by the sequential mean values $X_{ij} = \langle A_i A_j \rangle_{\text{seq}}$. This matrix has the following properties: (i) it is real and symmetric, $X = X^T$, (ii) the diagonal elements equal one, $X_{ii} = 1$, and (iii) the matrix has no negative eigenvalue (or $v^T X v \geq 0$ for any vector v), denoted as $X \succeq 0$ (see Sect. 4.5.2). A similar construction for the matrix X, together with the optimization problem below, has been considered before in relation with Bell inequalities [76]. However, our method involves a different notion of correlations, namely that given by Eq. (4.5).

The main idea is now to optimize the expression $C = \sum_{ij} \lambda_{ij} X_{ij}$ over all matrices with the properties (i)–(iii) above. Hence, we consider the optimization problem

$$\begin{aligned} \text{maximize:} \quad & \sum_{ij} \lambda_{ij} X_{ij}, \\ \text{subjected to:} \quad & X = X^T \succeq 0 \text{ and for all } i,\ X_{ii} = 1. \end{aligned} \tag{4.6}$$

Since all matrices X that can originate from a sequence of quantum measurements will be of this form, one performs the optimization over a potentially larger set. Thus, the solution of this optimization is, in principle, just an upper bound on the maximal quantum value of $\mathcal{S}_5$. Note that the optimization itself can be done efficiently and is assured to reach the global optimum since it represents a so-called semidefinite program [77]. In the case of $\mathcal{S}_5$, this optimization can even be solved analytically and gives

$$\mathcal{S}_5 \leq \frac{5}{4}\left(1 + \sqrt{5}\right) \approx 4.04. \tag{4.7}$$

It turns out that appropriately chosen measurements on a qubit already reach this value [110]. Hence, this upper bound is tight. More generally, one can prove (see Sects.4.5.2 and 4.5.4) that each matrix X with the above properties has a sequential quantum representation. Finally, note that if the observables in each sequence are required to commute, then the maximal quantum value for $\mathcal{S}_5$ is known to be $\Omega_{QM} = 4\sqrt{5} - 5 \approx 3.94$ [92, 87].

4.3 The general method

The above method can only be used for correlations terms of sequences of at most two ± 1-valued observables. In the following, we discuss the conditions allowing a given probability distribution to be realized as sequences of measurements on a single quantum system in the general setting. We label as $\mathbf{r} = (r_1, r_2, \ldots, r_n)$ the results of an n-length sequence obtained by using the setting $\mathbf{s} = (s_1, s_2, \ldots, s_n)$. The ordering is such that r_1, s_1 label the result and the setting for the first measurement etc. The outcomes of any such sequence are sampled from the sequential conditional probability distribution

$$P(\mathbf{r}|\mathbf{s}) \equiv P_{\text{seq}}(r_1, r_2, \ldots, r_n | s_1, s_2, \ldots, s_n). \tag{4.8}$$

In the case of projective quantum measurements, each individual result r of any setting s is associated with a projector Π_r^s, which altogether satisfy two requirements: For each setting the operators must sum up to the identity, i.e., $\sum_r \Pi_r^s = \mathbb{1}$, and they satisfy the orthogonality relations $\Pi_r^s \Pi_{r'}^s = \delta_{rr'} \Pi_r^s$, where $\delta_{rr'}$ is the Kronecker symbol. Finally, after the measurement with the setting s and result r, the quantum state is transformed according to Lüders rule $\varrho \mapsto \Pi_r^s \varrho \Pi_r^s / P(r|s)$.

In the following, we say that the a conditional probability distribution $P(\mathbf{r}|\mathbf{s})$ has a sequential projective quantum representation if there exists a suitable set of such operators Π_r^s and an appropriate initial state ϱ such that

$$P(\mathbf{r}|\mathbf{s}) = \text{tr}[\Pi(\mathbf{r}|\mathbf{s}) \Pi(\mathbf{r}|\mathbf{s})^\dagger \varrho], \tag{4.9}$$

with the shorthand $\Pi(\mathbf{r}|\mathbf{s}) = \Pi_{r_1}^{s_1}\Pi_{r_2}^{s_2}\cdots\Pi_{r_n}^{s_n}$.

Whether a given distribution $P(\mathbf{r}|\mathbf{s})$ indeed has such a representation can be answered via a so-called matrix of moments, which often appears in moment problems [76, 79, 80, 113]. This matrix, denoted as M in the following, contains the expectation value of the products of the above-used operators $\Pi(\mathbf{r}|\mathbf{s})$ at the respective position in the matrix. In order to identify this position we use as a label the abstract operator sequence $\mathbf{r}|\mathbf{s}$ for both row and column index. In this way the matrix is defined as

$$M_{\mathbf{r}|\mathbf{s};\mathbf{r}'|\mathbf{s}'} = \langle \Pi(\mathbf{r}|\mathbf{s})\Pi(\mathbf{r}'|\mathbf{s}')^\dagger|.\rangle \tag{4.10}$$

Whenever this matrix is indeed given by a sequential projective quantum representation, the matrix M satisfies two conditions: (a) linear relations of the form $M_{\mathbf{r}|\mathbf{s};\mathbf{k}|\mathbf{l}} = M_{\mathbf{r}'|\mathbf{s}';\mathbf{k}'|\mathbf{l}'}$ if the underlying operators are equal as a consequence of the properties of normalization and orthogonality of projectors, (b) $M \succeq 0$ since $v^\dagger M v \geq 0$ holds for any vector v, because such a product can be written, by defining $C = \sum_{\mathbf{r}|\mathbf{s}} v_{\mathbf{r}|\mathbf{s}}\Pi(\mathbf{r}|\mathbf{s})$, as the expectation value $\langle CC^\dagger\rangle_\varrho \geq 0$, which is non-negative for any operator C. Finally, note that certain entries of this matrix are the given probability distribution, for instance, at the diagonal $M_{\mathbf{r}|\mathbf{s};\mathbf{r}|\mathbf{s}} = P(\mathbf{r}|\mathbf{s})$. The main point, however, is the converse statement: Given a moment matrix with properties (a) and (b) above, the associated probability distribution $P(\mathbf{r}|\mathbf{s})$ always has a sequential projective quantum representation (see Sect.4.5.4).

Hence, the search for quantum bounds represents again a semidefinite program. The fact that this characterization is sufficient is in stark contrast with the analogue technique in the spatial Bell-type scenario [80], where one needs to use moment matrices of an increasing size n to generate better superset characterizations which only become sufficient in the limit $n \to \infty$. However, indirectly, the sufficiency of our method has already been proven in this context [80] (see 4.5.4 for details).

4.4 Applications

To demonstrate the effectiveness of our approach, we discuss four examples. First, we consider the original Leggett-Garg inequality

$$\begin{aligned}\mathcal{S} = \langle M(t_1)M(t_2)\rangle_{\text{seq}} + \langle M(t_2)M(t_3)\rangle_{\text{seq}} \\ -\langle M(t_1)M(t_3)\rangle_{\text{seq}} \leq 1.\end{aligned} \tag{4.11}$$

This bound holds for macrorealistic models, and it has been shown that in quantum mechanics values up to $\mathcal{S} = 3/2$ can be observed [8, 110, 111]. Our methods allow us not only to prove that this value is optimal for any dimension and any measurement, provided the state-update rule is the one of Eq. (4.3), but also to, for instance, determine all values in the three-dimensional space of temporal correlations $\langle M(t_i)M(t_j)\rangle$, which can originate from quantum mechanics. The detailed description is given in Fig. 4.1, and the calculations are given in the Sect. 4.5.1. We shall see in the next chapter, that more general state-update rules allow for a higher value.

Second, we consider the case of sequential measurements for the N-cycle scenario of Chapter 2, namely, the expression

$$\mathcal{S}_N = \sum_{i=0}^{N-2} \langle A_i A_{i+1}\rangle_{\text{seq}} - \langle A_{N-1}A_0\rangle_{\text{seq}}. \tag{4.12}$$

For this case, everything can be solved analytically (see Sect.4.5.1 for the general sequential bound, and Sect. 4.5.4 for the quantum bound for compatible measurements) leading to the bound

$$\mathcal{S}_N \leq N\cos\left(\frac{\pi}{N}\right), \tag{4.13}$$

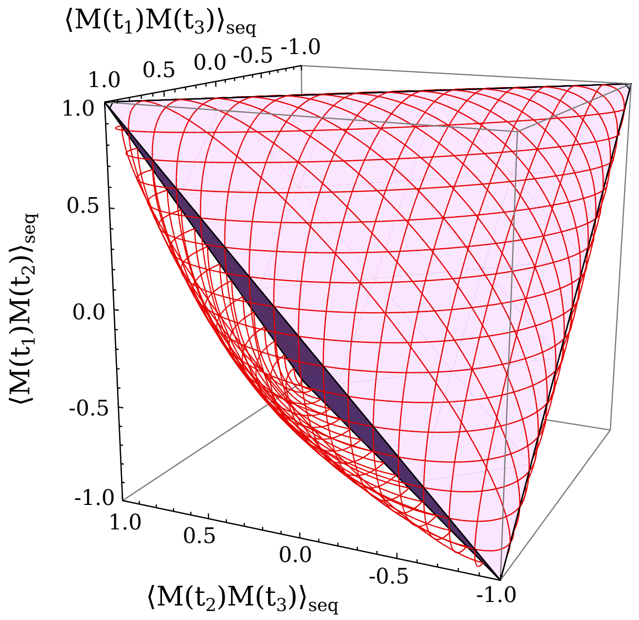


Figure 4.1: Complete characterization of the possible quantum values for the simplest Leggett-Garg scenario, with dichtomic measurements and Lüders rule (4.3). In this case, three different times are considered, resulting in three possible correlations $\langle M(t_1)M(t_2)\rangle_{\text{seq}}$, $\langle M(t_1)M(t_3)\rangle_{\text{seq}}$, and $\langle M(t_2)M(t_3)\rangle_{\text{seq}}$. In this three-dimensional space, the possible classical values form a tetrahedron, characterized by Eq. (4.11) and variants thereof. The possible quantum mechanical values form a strictly larger set with curved boundaries.

which can be reached by suitably chosen measurements. This value has already occurred in the literature [110], but only qubits have been considered. Our proof shows that it is valid in arbitrary dimension. Note that the fact that the maximal value is obtained on a qubit system is not trivial, although the measurements are dichotomic. For Kochen-Specker inequalities with dichotomic measurements examples are known, where the maximum value cannot be attained in a two-dimensional system [115] and also for Bell inequalities this has been observed [116, 117].

As a third application, we consider the noncontextuality scenario recently discovered by S. Yu and C. H. Oh [56]. There, thirteen measurements on a three-dimensional system are considered, and a noncontextuality inequality is constructed, which is violated by any quantum state. It has been shown that this scenario is the simplest situation where state-independent contextuality can be observed [98], so it is of fundamental importance. We can directly apply our method to the original inequality by Yu and Oh, as well as recent improvements [94] and compute the corresponding Tsirelson-like bounds. We recall that our results are not directly related to the phenomenon of quantum contextuality, since no compatibiliy of the measurements is assumed, but they show the effectiveness of our method even on complex scenarios, namely, inequalities containing 37 or 41 terms, that involve sequential measurements. Our results are summarized in Table 4.1.

Another class of inequalities is given by the guess-your-neighbor's-input inequalities [118], which if viewed as multipartite inequalities, show no quantum violation but a violation with the use of postquantum no-signalling resources. We calculate the sequential bound for the case of measurement sequences of length three, instead of measurement on three parties. We consider

$$P(000|000) + P(110|011) + P(011|101) \\ +P(101|110) \leq \Omega_{C,Q} \leq \Omega_S \leq \Omega_{NS}, \tag{4.14}$$

Ineq.	NCHV bound	State-independent quantum value	Algebraic maximum	**Sequential bound**
Yu-Oh	16	$50/3 \approx 16.67$	50	17.794
Opt2	16	$52/3 \approx 17.33$	52	20.287
Opt3	25	$83/3 \approx 27.67$	65	32.791

Table 4.1: Bounds on the quantum correlations for the Kochen-Specker inequalities in the most basic scenario. Three inequalities were investigated: First, the original inequality proposed in Ref. [56] and the optimal inequalities from Ref. [94] with measurement sequences of length two (Opt2) and length three (Opt3). For each inequality, the following numbers are given: The maximum value for noncontextual hidden variable (NCHV) models, the state-independent quantum violation in three-dimensional systems (obtained in Refs. [56, 94]), the algebraic maximum and the maximal value that can be attained in quantum mechanics for the sequential measurements. The latter bound is higher than the state independent quantum value, since the observables do not have to obey the compatibility relations occurring in the Kochen-Specker theorem. Notice that the sequential bound is obtained as a maximization over the set of possible observables and states, thus it is in general state-dependent. Interestingly, in all cases the maximal quantum values are significantly below the algebraic maximum.

with the notation $P(r_1, r_2, r_3|s_1, s_2, s_3)$ as before, and possible results and settings $r_i \in \{0, 1\}$ and $s_i \in \{0, 1\}$. We find that

$$\Omega_S \approx 1.0225, \tag{4.15}$$

while it is known that $\Omega_{C,Q} = 1$ and $\Omega_{NS} = \frac{4}{3}$, where the indices C, Q, S, NS label, respectively, the classical, quantum, sequential and no-signalling bounds. So, in this case, the bound for sequential measurements is higher than the bound for spatially separated measurements. This also highlights the greater generality of our method in comparison with the results of Ref. [23]: There, only temporal inequalities with sequences of length two have been considered, where in addition the measurements can be split in two separate groups. In this case it turned out that the bounds were always reached with commuting observables. Our examples show that this is usually not the case, when longer measurement sequences are considered.

4.5 Details of the calculations

4.5.1 Discussion of the simplest Leggett-Garg scenario

In this part we provide some further details about how to determine the set of possible quantum values for the simplest non-trivial Leggett-Garg scenario as shown in Fig. 4.1. Here it is assumed that one can measure an observable M at three different time instances t_1, t_2, t_3, which gives rise to three different observables $A_i = M(t_i)$ with $i = 1, 2, 3$.

However, rather than being interested in determining the full sequential probability $P(\mathbf{r}|\mathbf{s})$ for all possible combinations we are here only interested in some limited information, namely only for the correlation space. This means that from a general distribution we only want to reproduce the correlations terms $\langle A_i A_j|\rangle$seq with $1 \leq i < j \leq 3$ each defined by

$$\langle A_i A_j \rangle_{\text{seq}} = P(r_i = r_j|i, j) - P(r_i \neq r_j|i, j). \tag{4.16}$$

Thus we want to characterize the set

$$\langle S\rangle_{\text{qm}} = \{q_{ij} \in \mathbb{R}^3 : q_{ij} = \langle A_i A_j\rangle_{\text{seq}}, \\ \langle A_i A_j\rangle_{\text{seq}} \text{ has projective quantum rep.}\}. \tag{4.17}$$

For this we refer to problem given by (4.6), with

$$X = \begin{bmatrix} 1 & \langle A_1 A_2\rangle_{\text{seq}} & \langle A_1 A_3\rangle_{\text{seq}} \\ \langle A_1 A_2\rangle_{\text{seq}} & 1 & \langle A_2 A_3\rangle_{\text{seq}} \\ \langle A_1 A_3\rangle_{\text{seq}} & \langle A_2 A_3\rangle_{\text{seq}} & 1 \end{bmatrix}. \tag{4.18}$$

Any matrix of this form has a sequential projective quantum representation if and only if X is positive semidefinite. However a matrix satisfies $X \succeq 0$ if and only if the determinant of all principal minors are non-negative. This gives

$$\mathcal{S}_{\text{qm}} = \{q_{ij} \in \mathbb{R}^3 : |q_{ij}| \leq 1, \\ 1 + 2q_{12}q_{13}q_{23} \geq q_{12}^2 + q_{13}^2 + q_{23}^2\}. \tag{4.19}$$

which is the plotted region of Fig. 4.1.

We mention that via the general method one can also in principle determine the achievable probability distribution of a general scenario. However, this requires the solution of a SDP with some unknown entries, and hence an analytic solution is in general not accessible.

4.5.2 Detailed discussion of sequential bounds for the N-cycle inequalities

We first need the general form [87] for Eq. (4.12)

$$\mathcal{S}_N(\gamma) = \sum_{i=0}^{N-1} \gamma_i \langle A_i A_{i+1}\rangle_{\text{seq}}, \tag{4.20}$$

where the indices are taken modulo N and $\gamma = (\gamma_0, \ldots, \gamma_{N-1}) \in \{-1, 1\}^N$ with an odd number of -1. Since any two assignments γ and γ' can be converted into each other via some substitutions $A_i \to -A_i$, the quantum bound does not depend on the particular choice of γ. For the case of odd N, we can consider the expression

$$\mathcal{S}_N = -\sum_{i=0}^{N-1} \langle A_i A_{i+1}\rangle_{\text{seq}}, \tag{4.21}$$

with index i taken modulo N. The optimization problem in Eq. (4.6), therefore, can be expressed as

$$\begin{aligned} \text{maximize:} \quad & \frac{1}{2}\text{tr}(WX) \\ \text{subjected to:} \quad & X = X^T \succeq 0 \text{ and } X_{ii} = 1 \text{ for all } i, \end{aligned} \tag{4.22}$$

where W is the circulant symmetric matrix

$$W = -\begin{bmatrix} 0 & 1 & \ldots & 0 & 1 \\ 1 & 0 & 1 & & 0 \\ \vdots & 1 & 0 & \ddots & \vdots \\ 0 & & \ddots & \ddots & 1 \\ 1 & 0 & \ldots & 1 & 0 \end{bmatrix}. \tag{4.23}$$

The condition $X \succeq 0$, i.e., $v^T X v \geq$ for any real vector v, follows from the fact that $\langle A_i A_j \rangle_{\text{seq}} = \frac{1}{2}\text{tr}[\varrho(A_i A_j + A_j A_i)]$ and the fact that the matrix $Y = \text{tr}[\varrho(A_i A_j)]$ fulfils $v^T Y v \geq$ for any real vector v, and X is the real part of Y.

By using the vector $\lambda = (\lambda_1, \ldots, \lambda_N)$, the dual problem for the semidefinite program in Eq. (4.22) can be written as (see Ref. [77] for a general treatment and Ref. [76] for the discussion of a similar problem)

$$\begin{aligned} \text{minimize:} \quad & \text{tr}(diag(\lambda)) \\ \text{subjected to:} \quad & -\frac{1}{2}W + diag(\lambda) \succeq 0, \end{aligned} \tag{4.24}$$

where $diag(\lambda)$ denotes the diagonal matrix with entries $\lambda_1, \ldots, \lambda_N$.

Let us denote with p and d optimal values for, respectively, the primal problem in Eq. (4.22) and the dual problem in Eq. (4.24). Then $d \geq p$. We shall provide a feasible solution for the dual problem with $d = N\cos(\frac{\pi}{N})$ and a feasible solution for the primal problem with $p = d$, this will guarantee the optimality of our primal solution.

We start by finding the maximum eigenvalue for W. Since W is a circulant matrix, its eigenvalues can be written as [119]

$$\mu_j = -2\cos\left(\frac{2\pi j}{N}\right) \tag{4.25}$$

for $j = 0, \ldots, N-1$, and $\mu_{max} = 2\cos\left(\frac{\pi}{N}\right)$ the maximum eigenvalue.

For a pair of Hermitian matrices A, B, it holds $\mu_{min}(A+B) \geq \mu_{min}(A) + \mu_{min}(B)$, where μ_{min} denotes the minimum eigenvalue. Therefore, $\lambda = (\cos\left(\frac{\pi}{N}\right), \ldots, \cos\left(\frac{\pi}{N}\right))$ is a feasible solution for the dual problem and $\text{tr}[diag(\lambda)] = N\cos\left(\frac{\pi}{N}\right)$, and $p \leq N\cos\left(\frac{\pi}{N}\right)$.

Now consider the matrix $X'_{ij} = (x_i, x_j)$, with $x_1, \ldots, x_N$ unit vectors in a 2-dimensional space such that the angle between x_i and x_{i+1} is $\frac{N+1}{N}\pi$, and $(\cdot, \cdot)$ denoting the scalar product. Clearly, X' is positive semidefinite. Since $X'_{i,i+1} = -\cos\left(\frac{\pi}{N}\right)$, it follows that $p = d = N\cos\left(\frac{\pi}{N}\right)$ and the solution X' is optimal.

In order to prove that X' can be obtained as matrix of expectation values for sequential measurements, we define for a 3-dimensional unit vector $\vec{a}$ the observable $\sigma_a \equiv \vec{\sigma} \cdot \vec{a}$, where $\vec{\sigma}$ denots the vector of the Pauli matrices. Then, by Eq. (4.5), $\langle \sigma_a \sigma_b \rangle_{seq} = \vec{a} \cdot \vec{b}$, independently of the initial quantum state ϱ. In fact, explicit observables reaching this bound have already been discussed in the literature [110, 115].

For the case N even, we can consider the expression

$$\mathcal{S}_N = \sum_{i=0}^{N-2} \langle A_i A_{i+1} \rangle_{seq} - \langle A_0 A_{N-1} \rangle_{seq}, \tag{4.26}$$

and the maximization problem can be expressed as a SDP as in Eq. (4.22), with the proper choice of the matrix W. Such a SDP has been solved in Ref. [76]. The solution is analogous to the previous one: A set of observables, for a two-level system, saturating the bound, again, independently of the quantum state, is given by observables $A_i = \vec{\sigma} \cdot \vec{x_i}$, where the vectors x_i are on a plane with an angle $\frac{\pi}{N}$ separating x_i and x_{i+1}.

As opposed to the odd N case, such a bound can be also reached with commuting operators, this corresponds to the well known maximal violation of Braunstein-Caves inequalities [76].

The above results prove that the bound computed in Ref. [115] for sequential measurements on qubits, coinciding with the value explicitly obtained in Ref. [110], is valid for any dimension of the quantum system on which measurements are performed.

Finally, we stress that the construction of the above set of observables from the solution of the SDP, i.e., the matrix X or the set of vectors $\{x_i\}$ such that $X_{ij} = (x_i, x_j)$, is general. We recall that the vectors $\{x_i\}$ can be obtained, e.g., as the columns of the matrix $\sqrt{X}$ and, therefore, the dimension of the subspace spanned by them is equal to the rank of the matrix X. In the previous case, since we were dealing with vectors in dimension $d \leq 3$, we used the property of Pauli matrices

$$\{\sigma_a, \sigma_b\} \equiv \sigma_a\sigma_b + \sigma_b\sigma_a = 2(\vec{a} \cdot \vec{b})\mathbb{1}. \tag{4.27}$$

For matrices X with higher rank, the corresponding vectors $\{x_i\}$ will span a real vector space V of dimension $d > 3$. Now for general complex vector spaces V with a symmetric bilinear form $(\ ,\ )$, an analogue of Eq. (4.27), namely

$$\{A_v, A_u\} = 2(v, u)\mathbb{1}, \quad \text{for any } u, v \in V \tag{4.28}$$

can be established by a representation of associated Clifford algebra, cf. Ref. [120, 121]

As a consequence, for every positive semidefinite real matrix X with diagonal elements equal to 1, one can find a set of unit vectors $\{x_i\}$ giving $X_{ij} = (x_i, x_j)$ and a set of ±1-valued observables $\{A_i\}$, associated with $\{x_i\}$, such that

$$\langle A_iA_j\rangle_{seq} = \mathrm{tr}\left[\frac{1}{2}\varrho(A_iA_j + A_jA_i)\right] = (x_i, x_j), \tag{4.29}$$

for all quantum states ϱ. In particular, if the rank of X is d, such operators can be chosen as $2^d \times 2^d$ Hermitian matrices [122]. This shows the completeness of the simplified method.

4.5.3 Completeness of the general method

In this part we shortly comment on the completeness of the presented general method. As pointed out, this has already been proven indirectly in the context of the spatial bipartite case [80].

At first let us change slightly the notation in order to make it closer to the one used in Ref. [80]. In the following, we do not explicitly consider the matrix M, but rather a slightly smaller matrix where one erases some trivial constraints. In the following the set $\{E_i\}$ contains all projectors Π_k^s, but one of the outcomes k from each setting s is left out. We also use a single subscript to identify setting and outcome. Then the matrix

$$\chi^n_{\mathbf{uv}} = \mathrm{tr}[E(\mathbf{u})E(\mathbf{v})^\dagger\rho] \tag{4.30}$$

with $\mathbf{u} = (u_1, u_2, \ldots, u_l)$ is built from all products $E(\mathbf{u}) = E_{u_1}E_{u_2}\cdots E_{u_l}$ of the operators $\{E_i\}$ of at most length $l \leq n$, and the single extra "sequence" $\mathbf{u} = 0$ of the identity operator, $E(0) = \mathbb{1}$. Again this matrix has to satisfy linear relations parsed as $\chi^n_{\mathbf{uv}} = \chi^n_{\mathbf{u'v'}}$, if the operators fulfil $E(\mathbf{u})E(\mathbf{v})^\dagger = E(\mathbf{u}')E(\mathbf{v}')^\dagger$ as a consequence of the orthogonality properties of projectors, and that $\chi^n \succeq 0$.

That this matrix is positive semidefinite can be verified as follows: Let us first assume that there exists a sequential projective quantum representation. Consider the operator $C = \sum_{\mathbf{u}} c_{\mathbf{u}}E(\mathbf{u})^\dagger$ with arbitrary $c_{\mathbf{u}} \in \mathbb{C}$ and evaluate the expectation value of $CC^\dagger$, which provides

$$\mathrm{tr}(CC^\dagger\varrho) = \sum_{\mathbf{u},\mathbf{v}} c_{\mathbf{u}}\mathrm{tr}[E(\mathbf{u})^\dagger E(\mathbf{v})\varrho]c^*_{\mathbf{v}} \tag{4.31}$$

$$= \sum_{\mathbf{u},\mathbf{v}} c_{\mathbf{u}}\chi^n_{\mathbf{uv}}c^*_{\mathbf{v}} \geq 0. \tag{4.32}$$

The final inequality holds because $CC^\dagger \succeq 0$ and $\rho \succeq 0$ are both positive semidefinite operators. Since $c_{\mathbf{u}} \in \mathbb{C}$ are arbitrary the condition given by Eq. (4.32) means that $\chi^n \succeq 0$ is positive semidefinite.

For the reverse one needs a way to construct an explicit sequential projective quantum representation out of the matrix χ^n satisfying the above properties. For this, clearly more difficult part, we refer to Ref. [80] and just mention the solution. For the given positive semidefinite matrix χ^n one associates a set of vectors $\{|e_{\mathbf{u}}\rangle\}$ by the relation $\chi^n_{\mathbf{uv}} = \langle e_{\mathbf{u}}||e_{\mathbf{v}}\rangle$. From this set of vectors one now constructs an appropriate state and corresponding projective measurements by $\hat{\mathcal{H}} = \mathrm{span}(\{|e_{\mathbf{u}}\rangle\})$, $\hat{\rho} = |e_0\rangle\langle e_0|$, and $\hat{E}_i = \mathrm{proj}(\mathrm{span}(\{|e_{\mathbf{u}}\rangle : u_1 = i\}))$ where proj means the projector onto the given subspace. That these solution satisfies all the required constraints is shown in the proof of Theorem 8 of Ref. [80]. An analogous mathematical result, valid only for the case of dichotomic observables, has been presented also in Ref. [114].

In the spatial case considered in Ref. [80], some of these operators, additionally, have to commute since they should correspond to measurements onto different local parts. This cannot be inferred, in general, by a finite level χ^n and this is eventually the reason why in the spacial case arbitrary high order terms have to be considered. However, luckily, since in our situation the measurements of different settings may well fail to commute we can rely on a finite level n.

4.5.4 Quantum bounds for compatible measurements in the N-cycle scenario

We now complete the discussion of the N-cycle noncontextuality scenario by computing, by means of the CSW method, the quantum bound for the inequalities (4.12) in the case of sequence of compatible measurements.

Such bounds are given by

$$\Omega_{\mathrm{QM}} = \begin{cases} \frac{3n\cos\left(\frac{\pi}{N}\right)-N}{1+\cos\left(\frac{\pi}{N}\right)} & \text{for odd } N, \\ N\cos\left(\frac{\pi}{N}\right) & \text{for even } N. \end{cases} \tag{4.33}$$

Let us first discuss the case of odd N. As in the above case of general sequential measurements, without loss of generality, we can restrict our discussion to the inequalities in which $\gamma_i = -1$ for all i, namely,

$$\mathcal{S}_N = -\sum_{i=0}^{N-1} \langle A_i A_{i+1}\rangle_{\mathrm{seq}}, \tag{4.34}$$

Using that

$$\pm\langle A_i A_{i+1}\rangle = 2\Big[p(+\pm|A_i,A_{i+1}) + p(-\mp|A_i,A_{i+1})\Big] - 1, \tag{4.35}$$

we can rewrite $\mathcal{S}_N$ as $2\Sigma - N$, where Σ is a sum of probabilities.

In quantum mechanics, any sum of probabilities is bounded from above by the Lovász ϑ-function, $\vartheta(G)$, of the graph G in which nodes are the arguments of the probabilities and edges link exclusive events (e.g., $(++|A_0,A_1)$ and $(--|A_1,A_2)$) [81, 82].

If N is odd, the graph G associated to Σ is the prism graph of order N, Y_N (see Fig. 4.2). We shall prove that its ϑ-function is

$$\vartheta(Y_N) = \frac{2N\cos\left(\frac{\pi}{N}\right)}{1+\cos\left(\frac{\pi}{N}\right)}, \tag{4.36}$$

therefore, if N is odd, the quantum bound for compatible measurements Ω_{QM} is bounded from above by $2\vartheta(Y_N) - N$. The following quantum state and observables saturates this bound [92]: $|\psi\rangle = (1,0,0)$ and $A_j = 2|v_j\rangle\langle v_j|v_j - \mathbb{1}$, where $|v_j\rangle = (\cos\theta, \sin\theta\cos[j\pi(N-1)/N], \sin\theta\sin[j\pi(N-1)/N])$ and $\cos^2\theta = \cos(\pi/N)/(1+\cos(\pi/N))$.

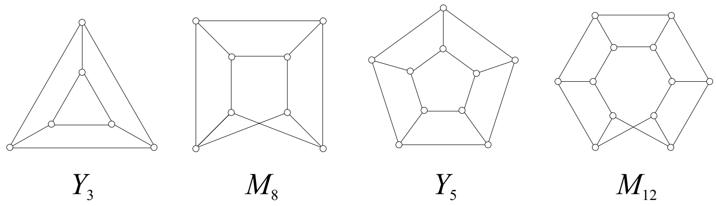


Figure 4.2: Graphs associated to the sum of probabilities Σ in the N-cycle inequalities for $N = 3, 4, 5, 6$.

Equation 4.36) can be proven as follows. An orthonormal representation (OR) for a graph $G = (V, E)$ is a set of unit vectors $\{v_i\}$ associated with vertices $V = \{i\}$ such that two vectors are orthogonal if the corresponding vertices are adjacent, $i.e.(i, j) \in E$. Lovász ϑ function is defined as the maximum, over all OR, of the norm of the operator given by sum of the unidimensional projectors associated with vectors [83, 84]. Notice that different vertices can be mapped onto the same vector, but then the corresponding projector appears in the sum once for each vertex associated with it.

For the prism graph Y_N, in general, it holds $\vartheta(Y_N) \leq 2\vartheta(C_N) = \frac{2N\cos(\frac{\pi}{N})}{1+\cos(\frac{\pi}{N})}$ since a graph consisting in two copies of C_N, let us denote it as G, can be obtained from Y_N by removing the edges connecting vertices of the outer cycle with those of the inner cycle.

Consider an OR for C_N, say $v_0, \ldots, v_{N-1}$, which gives the maximum value for the norm of the corresponding sum of projectors, i.e $\vartheta(C_N)$. Clearly, the $2N$ vectors v_i, v_i', with $v_i' = v_i$, for $i = 0, \ldots, N-1$, form a OR for G, giving $\vartheta(G) = 2\vartheta(C_N)$. To show that $\vartheta(Y_N) = \vartheta(G) = 2\vartheta(C_N)$, it is sufficient to notice that the above vectors are also an OR for Y_N. Such an OR is obtained by associating v_i with the ith vertex of the outer cycle and the vector v_{i+1}' with the ith vertex of the inner cycle. This completes the discussion for the case of odd N.

For even N, the proof can be obtained simply by noting that such inequalities are closely related to the Braunstein-Caves inequalities [93], whose Tsirelson bound was found in [76]. The following quantum state and observables saturates the bound for general sequential measurements given in Eq. (4.13) $|\psi\rangle = (0, 1/\sqrt{2}, -1/\sqrt{2}, 0)$ and $A_j = \tilde{A}_j \otimes \mathbb{1}$ for even j and $A_j = \mathbb{1} \otimes \tilde{A}_j$ for odd j, where $\tilde{A}_j = \cos(j\pi/N)\sigma_x + \sin(j\pi/N)\sigma_z$ and σ_x, σ_z are Pauli matrices.

4.6 Discussion

In this chapter, we presented a general method to compute tight bounds for quantum correlations in the sequential measurement scenario.

For interpreting our results, let us note that our scenario is more general than the scenarios considered by Leggett and Garg and Kochen and Specker. Leggett and Garg consider measurements of the same observable on a system subjected to the time evolution $\varrho(t) = U(t)\varrho(0)U^\dagger(t)$, with $U(t) = e^{-iHt}$ for some Hamiltonian H, which is mapped onto the observables in the Heisenberg picture. In our case, the observables can be connected via unitaries. In fact, since the bound is independent of the dimension, one can always extend the Hilbert space and the corresponding observables, but without increasing the rank of the density matrix, such that all observables have the same eigenvalues with the same degeneracy. However, such unitaries, in general, do not correspond to a time evolution of the form $U(t) = e^{-iHt}$ for some self-adjoint operator H, i.e., they do not form a strongly-continuous one-parameter group [16]. Compared with the Kochen-Specker scenario, our approach is more general since it does not assume that the measurements in a sequence are commuting. Nevertheless, if one wishes to connect existing noncontextuality inequalities to information processing tasks, it is important to know the maximal quantum values (also if the observables do not commute), in order to characterize the largest quantum advantage possible.

Furthermore, we emphasize that in our derivation it was assumed that the measurements are described by projective measurements and this condition is indeed important. In fact, this sheds light on the role of projective measurements: One can easily construct classical devices with a memory, which give for sequential measurements as in Eq. (4.2) the algebraic maximum $\mathcal{S}_5 = 5$. These classical devices must also have a quantum mechanical description. Our results show, however, that in this quantum mechanical description a more general dynamical evolution than the Lüders rule is required. From this perspective, our results prove that the memory that can be encrypted in quantum systems by projective measurements is bounded.

Our results lead to the question of why quantum mechanics does not allow us to reach the algebraic maximum of temporal correlations, as long as projective measurements are considered.

In the next chapter, we consider a more general measurement scheme that still involves projective measurements but a state-update rule more general than Lüders rule and closer to von Neumann's original proposal [19]. It will be shown that the algebraic maximum can actually be reached, but only in the limit of a infinite-dimensional system.

We believe that proper generalizations of information-theoretic concepts such as communication complexity [106], information causality [107], the E-principe [41], or Local Orthogonality [43] might play a role here, but we leave this question for further research. A first step in explaining quantum mechanics from information theoretical principles lies in the precise characterization of all possible temporal quantum correlations, and our work presents an operational solution to this problem.

Chapter 5

Dimension witnesses

The recent progress in the experimental control and manipulation of physical systems at the quantum level opens new possibilities (e.g., quantum communication, computation, and simulation), but, at the same time, demands the development of novel theoretical tools of analysis. There are already tools which allow us to recognize quantum entanglement and certify the usefulness of quantum states for quantum information processing tasks [123, 124]. However, on a more fundamental level, there are still several problems which have to be addressed. For example, how can one efficiently test whether measurements actually access all the desired energy levels of an ion? How to certify that the different paths of photons in an interferometer can be used to simulate a given multi-dimensional quantum system? Similar questions arise in the analysis of experiments with orbital angular momentum, where high-dimensional entanglement can be produced [125, 126], or in experiments with electron spins at nitrogen-vacancy centers in diamond, where the quantumness of the measurements should be certified [127].

The challenge is to provide lower bounds on the *dimension* of a quantum system only from the statistics of measurements performed on it. More precisely, one certifies lower bounds on the dimension of the underlying Hilbert space, where the measurement operators act on. Such bounds can be viewed as lower bounds on the complexity and the number of levels accessed by the measurement devices: If the measurement operators act non-trivially only on a small subspace, then all measurements results can be modeled by using a low-dimensional quantum system only. Note that this is not directly related to the rank of a density matrix. In fact, a pure quantum state acting on a one-dimensional subspace only can still give rise to measurement results, which can only be explained assuming a higher-dimensional Hilbert space.

The problem of estimating the Hilbert space dimension has been considered in different scenarios, and slightly different notions of dimension were involved. Brunner and coworkers introduced the concept of quantum "dimension witnesses" by providing lower bounds on the dimension of composite systems from the violation of Bell inequalities [128, 129]. The nonlocal properties of the correlations produced are clearly the resource used for this task. As a consequence, even if the experimenter is able to access and manipulate many levels of her systems locally, but she is not able to entangle those levels, the above test fails to certify such a dimension. Such a task can therefore be interpreted as a test of the type of entanglement and correlations produced, namely, how many levels or degrees of freedom the experimenter is able to entangle.

In a complementary scenario, several different states of a single particle are prepared and different measurements are carried out [130, 131, 132]. This approach has also recently been implemented using photons [133, 134]. In this situation, the dimension of the system can be interpreted as the dimension of the set of states the experimenter is able to prepare.

As a third possibility, also the continuous time evolution can be used to bound the dimension of a quantum system [135]. In this case, the relevant notion of dimension is that of the set of states generated by the dynamical evolution of the system.

In this chapter we focus on sequential measurements on a single system, a type of measurements used in tests of quantum versus classical theories, e.g., contextuality and Leggett-Garg inequalities, and we show how they can be used to provide lower bounds on the dimension of quantum systems. We recall that quantum contextuality is a genuine quantum effect leading to the Kochen-Specker theorem, which states that quantum mechanics is in contradiction to non-contextual hidden variable (NCHV) models [4, 5, 21, 39, 92]. In fact, already in the first formulation of the theorem the dimension plays a central role [39].

In Sect. 5.1, we derive bounds for the several important noncontextuality inequalities for different dimensions and scenarios. More precisely, we analyze the KCBS inequality both for the case of compatible and incompatible measurements, in Sects. 5.1.1 and 5.1.2. In Sects. 5.1.3, 5.1.4, we apply the same analysis to the Peres-Mermin inequality. Finally, we discuss the case of imperfect measurements and interpret some experimental data from a contextuality experiment in our framework, in Sects. 5.1.5, 5.1.6. The possibility of an application of the same analysis to different noncontextuality inequalities is then discussed in Sect. 5.1.7.

The experimental violation of these bounds automatically provides a lower bound on the dimension of the system, showing that noncontextuality inequalities can indeed be used as dimension witnesses. Remarkably, contextuality can be used as a resource for bounding the dimension of quantum systems in a state-independent way.

In the second part of the chapter, namely, Sect. 5.2, we focus on a single inequality, the original Leggett-Garg inequality, but with a more general measurement scheme. Different bounds can be derived, which depend on the dimension of the underlying quantum system. Analogously to the previous case, the experimental violation of such bounds in sequential measurement scenario provides a lower bound on the dimension of the system.

This illustrates clearly the difference with the existing schemes: Dimension witnesses derived according to Refs. [131, 132] certify the minimum classical or quantum dimension spanned by a set of preparations. They distinguish between classical and quantum dimension d, but, in general, not between quantum dimension d and classical dimension $d + 1$. They require at least $d + 1$ preparations to certify a dimension d. On the other side, dimension witnesses based on Bell's theorem or contextuality certify the minimum quantum dimension accessed by the measurement devices acting on a system prepared in the a single state. Contrary to the Bell scenario [128, 129], in our approach the initial state and its nonlocal properties play no role and the result of our test can directly be interpreted as the minimal number of levels accessed and manipulated by the measurement apparatus.

Technical details are presented separately in Sect. 5.3, the uninterested reader can skip this part. The results of this chapter have been published in Refs. [115, 25].

5.1 Noncontextuality and dimension witnesses

In this section, we shall discuss the application of noncontextuality inequalities as dimension witnesses. We shall analyze the usual scenario for contextuality tests, and a more general scenario involving imperfect measurements.

5.1.1 The KCBS inequality

We first turn to the state-dependent case. The simplest system showing quantum contextuality is a quantum system of dimension three [39]. The simplest noncontextuality inequality in three dimensions is the KCBS inequality [34]. For that, one considers

$$\langle \chi_{\text{KCBS}} \rangle = \langle AB \rangle + \langle BC \rangle + \langle CD \rangle + \langle DE \rangle + \langle EA \rangle, \tag{5.1}$$

where A, B, C, D, and E are measurements with outcomes -1 and 1, and the measurements in the same mean value $\langle \ldots \rangle$ are compatible [1], i.e., are represented in quantum mechanics by commuting operators. The mean value itself is defined via a sequential measurement: For determining $\langle AB \rangle$, one first measures A and then B on the same system, multiplies the two results, and finally averages over many repetitions of the experiment.

The KCBS inequality states that

$$\langle \chi_{\text{KCBS}} \rangle \overset{\text{NCHV}}{\geq} -3, \tag{5.2}$$

where the notation "$\overset{\text{NCHV}}{\geq} -3$" indicates that -3 is the minimum value for any NCHV theory. As we have seen, a value of $\langle \chi_{\text{KCBS}} \rangle = 5 - 4\sqrt{5} \approx -3.94$ can be reached on a three-dimensional quantum system, if the observables and the initial state are appropriately chosen. This quantum violation of the NCHV bound does not increase in higher-dimensional systems [92, 87], and the violation of the KCBS inequality has been observed in recent experiments with photons [136, 137].

Given the fact that quantum contextuality requires a three-dimensional Hilbert space, it is natural to ask whether a violation of Eq. (5.2) implies already that the system is not two-dimensional. The following observation shows that this is the case:

Observation 1. Consider the KCBS inequality where the measurements act on a two-dimensional quantum system and are commuting, i.e., $[A, B] = [B, C] = [C, D] = [D, E] = [E, A] = 0$. Then, the classical bound holds:

$$\langle \chi_{\text{KCBS}} \rangle \overset{\text{2D,com.}}{\geq} -3. \tag{5.3}$$

Proof of Observation 1. First, if two observables A and B are compatible, then $|\langle A \rangle \pm \langle AB \rangle| \leq 1 \pm \langle B \rangle$. This follows from the fact that A and B have common eigenspaces and the relation holds separately on each eigenspace. Second, in two dimensions, if $[A, B] = 0 = [B, C]$, then either $B = \pm\mathbb{1}$ or $[A, C] = 0$. The reason is that, if B is not the identity, then it has two one-dimensional eigenspaces. These are shared with A and C, so A and C must be simultaneously diagonalizable.

Considering the KCBS operator χ_{KCBS}, the claim is trivial if $A, \ldots, E$ are all compatible, because then the relation holds separately on each eigenspace. It is only possible that not all of them commute if there are two groups in the sequence $\{A, B, C, D, E\}$ of operators separated by identity operators. Without loss of generality, we assume that the groups of commuting operators are $\{E, A\}$ and $\{C\}$ so that $B = b\mathbb{1} = \pm\mathbb{1}$ and $D = d\mathbb{1} = \pm\mathbb{1}$. This gives

$$\begin{aligned} \langle \chi_{\text{KCBS}} \rangle &= b\langle A \rangle + b\langle C \rangle + d\langle C \rangle + d\big(\langle E \rangle + d\langle EA \rangle\big) \\ &\geq b\langle A \rangle + b\langle C \rangle + d\langle C \rangle - 1 - d\langle A \rangle \\ &= (b - d)\langle A \rangle + (b + d)\langle C \rangle - 1 \geq -3 \end{aligned} \tag{5.4}$$

and proves the claim. In this argumentation, setting observables proportional to the identity does not change the threshold, but in general it is important to consider this case, as this often results in higher values. □

It should be added that Observation 1 can also be proved using a different strategy: Given two observables on a two-dimensional system, one can directly see that if they commute, then either one of them is proportional to the identity, or their product is proportional to the identity. In both cases, one has a classical assignment for some terms in the KCBS inequality and then one can check by exhaustive search that the classical bound holds. Details are given in Sect. 5.3.1.

Furthermore, Observation 1 can be extended to generalizations of the KCBS inequality with more than five observables, i.e., the N-cycle scenario [87]: For that, one considers

$$\langle \chi_N \rangle = \sum_{i=1}^{N-1} \langle A_i A_{i+1} \rangle + s \langle A_N A_1 \rangle, \tag{5.5}$$

where $s = +1$ if N is odd and $s = -1$ if N is even. For this expression, the classical bound for NCHV theories is given by $\langle \chi_N \rangle \geq -(N-2)$. In fact, the experiment in Ref. [136] can also be viewed as measurement of $\langle \chi_6 \rangle$.

The discussion of the possible mean values $\langle \chi_N \rangle$ in quantum mechanics differs for even and odd N. If N is odd, the maximal possible quantum mechanical value is $\langle \chi_N \rangle = \Omega_N \equiv -[3N \cos(\pi/N) - N]/[1+\cos(\pi/N)]$ and this value can already be attained in a three-dimensional system [92, 87]. The proof of Observation 1 can be generalized in this case, implying that for two-dimensional systems the classical bound $\langle \chi_N \rangle \geq -(N-2)$ holds. So, for odd N, the generalized KCBS inequalities can be used for testing the quantum dimension.

If N is even, the scenario becomes richer: First, quantum mechanics allows to obtain values of $\langle \chi_N \rangle = \Omega_N \equiv -N \cos(\pi/N)$, but this time this value requires a four-dimensional system [87]. For two-dimensional quantum systems, the classical bound $\langle \chi_N \rangle \geq -(N-2)$ holds. For three-dimensional systems, one can show that if the observables A_i in a joint context are different ($A_i \neq \pm A_{i+1}$) and not proportional to the identity, then still the classical bound holds (for details see Sect. 5.3.2). However, if two observables are the same, e.g., $A_1 = -A_2$, then $\langle A_1 A_2 \rangle = -1$ and $\langle \chi_N \rangle = -1 + \langle \chi_{N-1} \rangle$. In summary, for even N, we have the following hierarchy of bounds

$$\langle \chi_N \rangle \overset{\text{2D,com.}}{\geq} -(N-2) \overset{\text{3D,com.}}{\geq} -1 + \Omega_{N-1} \overset{\text{4D,com.}}{\geq} \Omega_N. \tag{5.6}$$

Here, the notation $\overset{\text{2D,com.}}{\geq}$ etc. means that this bound holds for commuting observables in two dimensions. All these bounds are sharp. This shows that extended KCBS inequalities are even more sensitive to the dimension than the original inequality.

5.1.2 The KCBS inequality with incompatible observables

In order to apply Observation 1 the observables must be compatible. Since this condition is not easy to guarantee in experiments [102], we should ask whether it is possible to obtain a two-dimensional bound for the KCBS inequality when the observables are not necessarily compatible. We can state:

Observation 2. If the observables $A, \ldots, E$ are dichotomic observables but not necessarily commuting, then, for any two-dimensional quantum system,

$$\langle \chi_{\text{KCBS}} \rangle \overset{\text{2D}}{\geq} -\frac{5}{4}(1+\sqrt{5}) \approx -4.04. \tag{5.7}$$

This bound is sharp and can be attained for suitably chosen measurements. A general proof, i.e., valid for any dimension, of the above bound has been given in Chapter 4. However, since in this case we are dealing only with two-dimensional system, it is interesting to present a simpler proof of this statement that will also give introduce some of the techniques and ideas that we will use in the rest of the chapter.

The strategy of proving this bound is the following: If the observables are not proportional to the identity, one can write $A = |A^+\rangle\langle A^+| - |A^-\rangle\langle A^-|$ and $B = |B^+\rangle\langle B^+| - |B^-\rangle\langle B^-|$, and express $|A^+\rangle\langle A^+|$ and $|B^+\rangle\langle B^+|$ in terms of their Bloch vectors $|\mathfrak{a}\rangle$ and $|\mathfrak{b}\rangle$. Then, one finds that

$$\langle AB \rangle = 2|\langle A^+||B^+\rangle|^2 - 1 = \langle \mathfrak{a}||\mathfrak{b}\rangle. \tag{5.8}$$

This property holds for all projective measurements on two-dimensional systems and is, together with a generalization below [see Eq. (5.15)] a key idea for deriving dimension witnesses. Note that it implies that the sequential mean value $\langle AB \rangle$ is independent of the initial quantum state and also of the temporal order of the measurements [23]. Eq. (5.8) allows us to transform the KCBS inequality into a geometric inequality for three-dimensional Bloch vectors. Additional details of the proof are given in Sect. 5.3.3.

Observation 2 shows that the bound for NCHV theories can be violated already by two-dimensional systems, if the observables are incompatible. This demonstrates that experiments, which aim at a violation of Eq. (5.2) also have to test the compatibility of the measured observables, otherwise the violation can be explained without contextuality.

It must be added that Observation 2 cannot be used to witness the quantum dimension, since we showed in the previous chapter that Eq. (5.7) holds for all dimensions. As we see below, this difficulty can be surmounted by considering NC inequalities in which quantum mechanics reaches the algebraic maximum.

5.1.3 The Peres-Mermin inequality

In order to derive the state-independent quantum dimension witnesses, let us consider the sequential mean value [51],

$$\begin{aligned}\langle \chi_{\mathrm{PM}} \rangle =& \langle ABC \rangle + \langle bca \rangle + \langle \gamma\alpha\beta \rangle + \langle A\alpha a \rangle + \langle bB\beta \rangle \\ &- \langle \gamma c C \rangle, \end{aligned} \tag{5.9}$$

where the measurements in each of the six sequences are compatible. Then, for NCHV theories the bound

$$\langle \chi_{\mathrm{PM}} \rangle \overset{\mathrm{NCHV}}{\leq} 4 \tag{5.10}$$

holds. In a four-dimensional quantum system, however, one can take the following square of observables, known as the Peres-Mermin square [48, 49]

$$\begin{aligned} A &= \sigma_z \otimes \mathbb{1}, & B &= \mathbb{1} \otimes \sigma_z, & C &= \sigma_z \otimes \sigma_z, \\ a &= \mathbb{1} \otimes \sigma_x, & b &= \sigma_x \otimes \mathbb{1}, & c &= \sigma_x \otimes \sigma_x, \\ \alpha &= \sigma_z \otimes \sigma_x, & \beta &= \sigma_x \otimes \sigma_z, & \gamma &= \sigma_y \otimes \sigma_y. \end{aligned} \tag{5.11}$$

These observables lead for any quantum state to a value of $\langle \chi_{\mathrm{PM}} \rangle = 6$, demonstrating state-independent contextuality. The quantum violation has been observed in several recent experiments [53, 54, 55]. Note that the sequences in Eq. (5.9) are defined such that each observable occurs either always in the first or always in the second or always in the third place of a measurement a sequence. This difference to the standard version does not matter at this point (since the observables in any row or column commute), but it will become important below.

The PM inequality is of special interest for our program since it is violated up to the algebraic maximum with four-dimensional quantum systems and the violation is state-independent. Therefore, this inequality is a good candidate for dimension witnesses without assumptions on the measurements. First, we can state:

Observation 3. If the measurements in the PM inequality are dichotomic observables on a two-dimensional quantum system and if the measurements in each mean value are commuting, then one cannot violate the classical bound,

$$\langle \chi_{\mathrm{PM}} \rangle \overset{\mathrm{2D,\,com.}}{\leq} 4. \tag{5.12}$$

If one considers the same situation on a three-dimensional system, then the violation is bounded by

$$\langle \chi_{\mathrm{PM}} \rangle \overset{\text{3D, com.}}{\leq} 4(\sqrt{5}-1) \approx 4.94. \tag{5.13}$$

These bounds are sharp.

The idea for proving this statement is the following: If one considers the three commuting observables in each mean value and assumes that they act on a three-dimensional system, then three cases are possible: (a) one of the three observables is proportional to the identity, or (b) the product of two observables is proportional to the identity, or (c) the product of all three observables is proportional to the identity. One can directly show that if case (c) occurs in some mean value, then the classical bound $\langle \chi_{\mathrm{PM}} \rangle \leq 4$ holds. For the cases (a) and (b), one can simplify the inequality and finds that it always reduces to a KCBS-type inequality, for which we discussed already the maximal quantum values in different dimensions [see Eq. 5.6]. Details are given in Sect.5.3.3.

5.1.4 The PM inequality with incompatible observables

Let us now discuss the PM inequality, where the observables are not necessarily compatible. Our results allow us to obtain directly a bound:

Observation 4. Consider the PM operator in Eq. (5.9), where the measurements are not necessarily commuting projective measurements on a two-dimensional system. Then we have

$$\langle \chi_{\mathrm{PM}} \rangle \overset{\text{2D}}{\leq} 3\sqrt{3} \approx 5.20. \tag{5.14}$$

Proof. One can directly calculate as in the proof of Observation 2 that for sequences of three measurements on a two-dimensional system

$$\langle ABC \rangle = \langle A \rangle \langle BC \rangle \tag{5.15}$$

holds. Here, $\langle A \rangle = \mathrm{tr}(\varrho A)$ is the usual expectation value, and $\langle BC \rangle$ is the state-independent sequential expectation value given in Eq. (5.8). With this, we can write:

$$\begin{aligned} \langle \chi_{\mathrm{PM}} \rangle =& \langle A \rangle (\langle BC \rangle + \langle \alpha a \rangle) + \langle b \rangle (\langle ca \rangle + \langle B\beta \rangle) \\ &+ \langle \gamma \rangle (\langle \alpha\beta \rangle - \langle cC \rangle). \end{aligned} \tag{5.16}$$

Clearly, this is maximal for some combination of $\langle A \rangle = \pm 1$, $\langle b \rangle = \pm 1$, and $\langle \gamma \rangle = \pm 1$. But for any of these choices, we arrive at an inequality that is discussed in Prop. 5 in Sect. 5.3.3. Note that due to Eq. (5.15) the order of the measurements matters in the definition of $\langle \chi_{\mathrm{PM}} \rangle$ in Eq. (5.9). This motivates our choice; in fact, for some other orders (e.g., $\langle \tilde{\chi}_{\mathrm{PM}} \rangle = \langle ABC \rangle + \langle bca \rangle + \langle \beta\gamma\alpha \rangle + \langle A\alpha a \rangle + \langle \beta b B \rangle - \langle \gamma c C \rangle$) Eq. (5.14) does not hold, and one can reach $\langle \tilde{\chi}_{\mathrm{PM}} \rangle = 1 + \sqrt{9 + 6\sqrt{3}} \approx 5.404$. □

The question arises whether a high violation of the PM inequality also implies that the system cannot be three-dimensional and whether a similar bound as Eq. (5.14) can be derived. While the computation of a bound is not straightforward, a simple argument shows already that measurements on a three-dimensional systems cannot reach the algebraic maximum $\langle \chi_{\mathrm{PM}} \rangle = 6$ for any quantum state: Reaching the algebraic maximum implies that $\langle ABC \rangle = 1$. This implies that the value of C is predetermined by the values of A and B and the value A of determines the product BC. As this holds for any quantum state, it directly follows that A, B, C (and all the other observables in the PM square) are diagonal in the same basis and commute, so the bound in Observation 3 holds. From continuity arguments it follows that there must be a finite gap between the maximal value of $\langle \chi_{\mathrm{PM}} \rangle$ in three dimensions and the algebraic maximum.

5.1.5 Imperfect measurements

In actual experimental implementations the measurements may not be perfectly projective. It is therefore important to discuss the robustness of our method against imperfections.

Notice that, since we are considering sequential measurements, another possibility for maximal violation of the above inequalities is the use of a classical device with memory, able to keep track of the measurement performed and adjust the outcomes of the subsequent measurements accordingly in order to obtain perfect correlations or anti-correlations. However, as proved in Ref. [23] and also discussed in Ref. [24], such a classical device cannot be simulated in quantum mechanics via projective measurements, more general positive operator valued measures (POVMs) are necessary.

We therefore limit our analysis to some physically motivated noise models. A noisy projective measurement A may be modelled by a POVM with two effects of the type $E^+ = (1-p)\mathbb{1}/2 + p|A^+\rangle\langle A^+|$ and $E^- = (1-p)\mathbb{1}/2 + p|A^-\rangle\langle A^-|$. Then, the probabilities of the POVM can be interpreted as coming from the following procedure: With a probability of p one performs the projective measurement and with a probability of $(1-p)$ one assigns a random outcome. For this measurement model, one can show that Observation 4 is still valid. Details and a more general POVM are discussed in Sect. 5.3.5. We add that the proof strongly depends on the chosen measurement order in $\langle \chi_{\mathrm{PM}} \rangle$ and that in any case assumptions about the measurement are made, so the dimension witnesses are not completely independent of the measurement device.

The above discussion shows that it is extremely important to test the extent to which the measurements are projective and whether they are compatible. This can be achieved by performing additional tests. For instance, one can measure observable A several times in a sequence $\langle AAA \rangle$ to test whether the measurement is indeed projective. In addition, one may measure the sequence $\langle ABA \rangle$ and compare the results of the two measurements of A, to test whether A and B are compatible. For noncontextuality inequalities it is known how this information can be used to derive correction terms for the thresholds [102], and similar methods can also be applied here.

5.1.6 Experimental results

To stress the experimental relevance of our findings, let us discuss a recent ion-trap experiment [53]. There, the PM inequality has been measured with the aim to demonstrate state-independent contextuality. For our purpose, it is important that in this experiment also all permutations of the terms in the PM inequality have been measured. This allows also to evaluate our $\langle \chi_{\mathrm{PM}} \rangle$ with the order given in Eq. (5.9). Experimentally, a value $\langle \chi_{\mathrm{PM}} \rangle = 5.36 \pm 0.05$ has been found. In view of Observation 3, this shows that the data cannot be explained by commuting projective measurements on a three-dimensional system. Furthermore, Observation 4 and the discussion above prove that, even if the measurements are noisy and noncommuting, the data cannot come from a two-dimensional quantum system.

5.1.7 Generalizations

Generalizations of our results to other inequalities are straightforward: Consider a general noncontextuality inequality invoking measurement sequences of length two and three. For estimating the maximal value for two-dimensional systems (as in Observations 2 and 4) one transforms all sequential measurements via Eqs. (5.8) and (5.15) into expressions with three-dimensional Bloch-vectors, which can be estimated. Also noise robustness for the discussed noise model can be proven, as this follows also from the properties of the Bloch vectors (cf. Prop. 10 in the Sect.5.3.5). In addition, if a statement as in Observation 3 is desired, one can use the same ideas as the ones presented here, since they rely on general properties of commuting observables in three-dimensional space. Consequently, our methods allow to transform most of the known state-independent NC inequalities (for instance, the ones presented in Refs. [51, 56, 94, 100]) into witnesses for the quantum dimension.

5.2 Sequential measurements and Leggett-Garg inequality

In this section, by considering a broader class of projective measurements than hitherto considered, we show that the maximum quantum violation of the Leggett-Garg inequality can exceed the usual bound of 3/2. More precisely, we shall show that such a bound strongly depends on the number of levels N that can be accessed by the measurement apparatus via projective measurements. We provide exact bounds for small N that exceed the known bound for the Leggett-Garg inequality, and we show that in the limit $N \to \infty$ the Leggett-Garg inequality can be violated up to its algebraic maximum.

We discuss the application of the Leggett-Garg inequality as dimension witness as well as the implication of our results for the tests of macrorealist versus quantum theory.

Let us now be more concrete and recall the simplest Leggett-Garg inequality which, for dichotomic observable $Q = \pm 1$, reads

$$K_3 \equiv C_{21} + C_{32} - C_{31} \leq 1, \tag{5.17}$$

where $C_{\beta\alpha} = \langle Q(t_\beta)Q(t_\alpha)\rangle$ is the correlation function of variable Q at the two times $t_\beta \geq t_\alpha$. For a two-level system, the maximum quantum value of K_3 is $K_3^{\max} = \frac{3}{2}$ [8], which we shall refer to as the *Lüders bound*, $K_3^{\text{Lüders}} = \frac{3}{2}$, for reasons to become clear shortly. As we showed in the previous chapter, for measurements given by just two projectors, Π_+ and Π_-, onto eigenspaces associated with results $Q = +1$ and $Q = -1$, the maximum quantum value of K_3 is the same as for the qubit, *irrespective of system size* [24]. This has been reflected in several studies: The experiment of [65] on a three-level system obtained a maximum value less than $\frac{3}{2}$; on the theory side, multi-level quantum systems such as a large spin [138], optoelectromechanical systems [139] and photosynthetic complexes [140] have also been observed to obey $K_3 \leq K_3^{\text{Lüders}}$. From this, one might conclude that nothing new is to be gained from considering higher dimensional systems. Were this the case, the bound for the qubit would apply in all generality and $K_3^{\text{Lüders}}$ could be identified with the relevant temporal Tsirelson bound. However, as we will show, with a more general projective measurement scheme, violations of (5.17) for multi-level systems can exceed the qubit value.

Other than in an invasive scenario (where the algebraic maximum is trivially achieved, e.g., a classical device with memory or its quantum realization via POVMs [23]), the only hint that a violation of (5.17) greater than $K_3^{\text{Lüders}}$ is possible has come in the recent work by Dakić *et al.* [141]. There, however, the excess violation was claimed to stem from correlations beyond quantum theory. In contrast, our excess violations are found within the standard framework of quantum theory and projective measurements. This we achieve by considering measurements on an N-level system that can project the state in one of M different subspaces, $2 \leq M \leq N$, with outcomes that are nevertheless associated with either $Q = +1$ or $Q = -1$. From a macroscopic-realist point-of-view, this leaves (5.17) unchanged. From a quantum perspective, however, the choice of M determines the state-update rule under projective measurement: For $M = 2$ the projection is onto one of two subspaces, corresponding to Lüders rule for dichotomic measurements [57]; whereas $M = N$ is the case of a complete degeneracy-breaking measurement, as initially proposed by von Neumann [19] (see also Ref. [58] for a discussion). These additional possibilities for state reduction are ultimately responsible for the increased violations.

In the present section, we use the example of a large spin precessing in a magnetic field to demonstrate that violations $K_3 > \frac{3}{2}$ are possible and that the algebraic bound $K_3 = 3$ can be reached. We then discuss the exact upper bounds for small $M \leq 5$, and how they may be obtained with few-dimensional systems with $N \leq 9$.

Our results emphasize even more the value of sequential measurement scheme as dimension witnesses. Similarly to noncontextuality inequalities, also the Leggett-Garg inequality, combined with such a more general class of projective measurements, provides lower bounds on the dimension of quantum systems.

In addition, our results reveal a stark contrast between spatial and temporal correlations. In fact, we discuss how a similar modification to the spatial Bell scenario does not lead to an increase in the Tsirelson bound for the corresponding Bell inequality [74].

On the basis of the discussion of the previous chapter, and given the formal symmetry between Bell and Leggett-Garg inequalities [142, 143] and the general trend towards unification between temporal and spatial correlations [144, 145, 146, 147], one would expect that the Tsirelson bound for the Leggett-Garg inequality holds analogously to the spatial case. Surprisingly, we prove that this is not the case.

Moreover, we discuss how our results can be used in the discrimination of Lüders and von Neumann state-update rules [148], i.e., which one, if any, correctly represents the measurement scenario.

5.2.1 The measurement scheme

We consider measurements of a property Q, which can take values ± 1, on a N-level quantum system, with each level associated with a definite value of Q. From a macrorealist point-of-view, the fact that different levels are associated with the same value of Q is irrelevant: They may be considered as microscopically distinct states that have the same macroscopic property Q. Macrorealism and non-invasive measurability imply that at each instant of time, the system has a definite value of Q, which is independent of measurements previously performed on the system and, therefore, that the bound for Eq. (5.17) in macrorealist theories remains the same.

From a quantum mechanical perspective, the fact that the system has more than two levels, allows for many possible state-update rules. According to Lüders' rule [57], the state is updated as $\rho \mapsto \Pi_\pm \rho \Pi_\pm$, up to normalization, depending on the outcome of the measurement. On the opposite side, von Neumann's original proposal [19] is a state-update $\rho \mapsto \sum_k \Pi_\pm^{(k)} \rho \Pi_\pm^{(k)}$, where $\Pi_\pm^{(k)}$ are one-dimensional projectors. Both state-update rules are plausible, and the choice of the correct one depends on the particulars of the interaction between the system and the measurement apparatus (see Ref.[58] for a discussion).

More generally, we consider all possible intermediate cases, namely, state-update rules given by M different projectors, with $2 \leq M \leq N$, associated with either $+1$ or -1 outcome. The correlation functions are therefore given by

$$C_{\beta\alpha} = \sum_{l,m} q_l q_m \mathrm{tr}(\Pi_m U_{\beta\alpha} \Pi_l U_{\alpha 0} \rho_0 U_{\alpha 0}^\dagger \Pi_l U_{\beta\alpha}^\dagger), \tag{5.18}$$

where q_l represent the outcome ± 1 associated with Π_l, ρ_0 is the initial state of the system and $U_{\beta\alpha} = U(t_\beta - t_\alpha) = e^{-iH(t_\beta - t_\alpha)}$ is the unitary time-evolution operator for some Hamiltonian H.

5.2.2 A simple example

Consider a quantum-mechanical spin of length j in a magnetic field oriented in the x-direction. We write its Hamiltonian ($\hbar = 1$) as

$$H = \Omega J_x, \tag{5.19}$$

with Ω the level spacing and J_x the x-component of the angular momentum operator. Let us choose to measure the spin in the z direction such that the measurement projectors are $\Pi^j_m = |m;j\rangle\langle m;j|m;j$ with $|m;j\rangle$ eigenstates of the J_z operator. In this example, we only consider the von Neumann limit, $M = N = 2j+1$, and choose the measurement values to be $q^j_m = 1 - 2\delta_{m,-j}$, such that the lowest energy state is associated with the value -1, and the rest with $+1$.

Calculating the correlation functions $C_{\beta\alpha}$ for this setup, several differences with the qubit case are immediately apparent. Most importantly, the correlation functions here depend on both times, not just their difference. As corollary, the correlation functions depend on the initial state. A further difference is that, for the projectively-measured correlation functions discussed here, the order of the measurements $t_\beta > t_\alpha$ is important. This is not the case for $M = 2$, where we have seen that, for arbitrary N, the projectively-measured correlation functions are equal to the expectation value of the symmetrised product $\frac{1}{2}\{Q_j, Q_i\} = \frac{1}{2}(Q_jQ_i + Q_iQ_j)$, where the operators Q have spectral decomposition $Q = \Pi_+ - \Pi_-$, with $\Pi_\pm$ the projectors associated with the eigenvalues ± 1.

We initialise the system so that at time $t = 0$ it is in state $|\psi(t=0)\rangle = |-j;-j\rangle$ and set the measurement times as $\Omega t_1 = \pi$, $t_2 - t_1 = t_3 - t_2 = \tau$. For $N = 2$ we obtain the familiar qubit result. For $N = 3$, the LGI parameter reads:

$$K_3 = \frac{1}{16} + 2\cos(\Omega\tau) - \frac{5}{4}\cos(2\Omega\tau) + \frac{3}{16}\cos(4\Omega\tau), \tag{5.20}$$

which exhibits the key property in which we are interested — as 5.1 shows, this quantity shows a maximum of $K_3^{\max} = 1.7565$, clearly in excess of the Lüders bound.

5.2.3 Asymptotic limit

Fig. 5.1 further shows that the maximum value of K_3 for this model increases as a function of system size, N. In the limit $N \to \infty$, the maximum possible violation is $K_3^{\max} = 3$, as we now show. With measurement times $\Omega\tau = \frac{1}{2}\pi$, the correlation functions read (see Sect. 5.3.6 for details)

$$\begin{aligned} C_{31} &= -1; \qquad C_{21} = 1 - 2^{1-2j}; \\ C_{32} &= 1 - 2\frac{1}{2^{2j}} + 4\frac{1}{2^{4j}} - 2\frac{(4j)!}{4^{2j}[(2j)!]^2}. \end{aligned} \tag{5.21}$$

The corresponding value of K_3 as a function of N is shown in Fig. 5.1. For finite N, this choice of measurement time does not give the maximum violation. However, this result serves to bound $K_3^{\max}$ from below and, for large j, the asymptotic behaviour is

$$K_3 \to 3 - \sqrt{\frac{2}{\pi j}}. \tag{5.22}$$

Thus, at least in the limit that the dimension of the system becomes infinite, the K_3 LGI can be violated by quantum mechanics all the way up to the algebraic bound.

5.2.4 Maximum violations

While the precessing spin model reveals violations greater than the qubit case can occur, the violations for this system are *not* the maximum possible violations at a given N and M. Again, this is in contrast with the $M = 2$ case where the Rabi oscillation of the qubit provides the maximum violation.

To investigate the true maximum violations as a function of N and M, we combine two different methods. The maximum value for a given M can be obtained by means of the maximization method

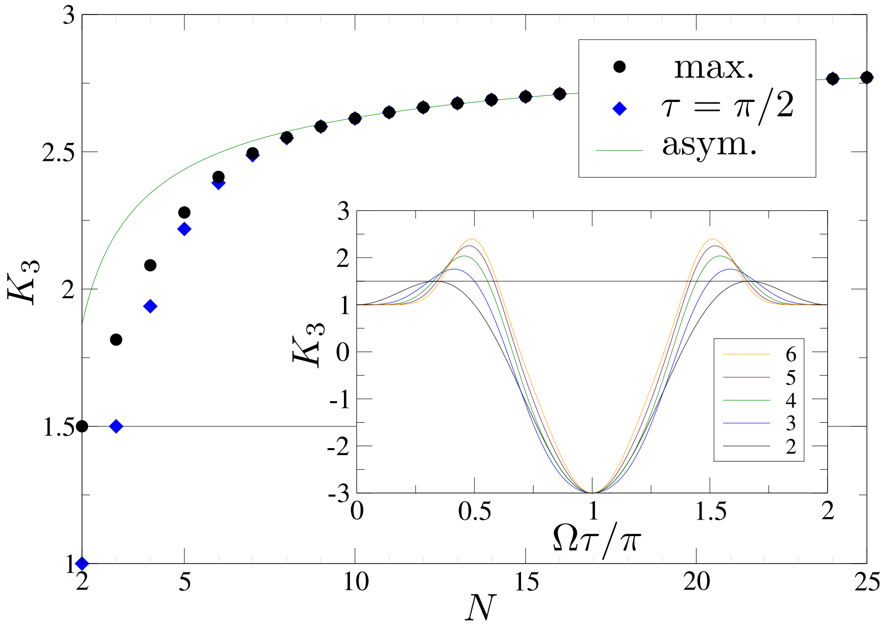


Figure 5.1: The Leggett-Garg quantity K_3 for a spin of length $j = (N-1)/2$ precessing in magnetic field with measurement times $\Omega t_1 = \pi$, $t_2 - t_1 = t_3 - t_2 = \tau$. The measurement is made with $M = N$ projectors (von Neumann scheme) in the the z-direction. **Inset:** K_3 as a function of measurement time τ for various values of N. For $N = 2$, the maximum is familiar qubit or Lüders bound $K_3^{\max} = \frac{3}{2}$ (solid line). For $N = 3$, however, the maximum value is 1.7565, and this increases with increasing N. **Main panel:** The black circles show the maximum value $K_3^{\max}$ as a function of system size $N = 2j+1$ for the spin precession model with measurement times as above. The blue diamonds show the value of K_3 with τ fixed $\Omega\tau = \pi/2$ and the solid line shows the asymptotic behaviour $K_3^{\max} \sim 3 - \sqrt{2/\pi j}$. In the limit $N \to \infty$, $K_3^{\max}$ tends to the algebraic bound of 3.

for temporal correlations presented in the previous chapter and based on semidefinite programming. This method provides an upper bound valid for any N, which is attained for any $N \geq N_{min}$. However, the exact value for N_{min} cannot be extracted from the solution, even though the method provides a state and a set of observables attaining the maximum quantum value.

We also pursue a complementary approach in which, for explicit values of N and M, we numerically maximise K_3 over time-evolution operators $U_{\beta\alpha}$ treated as general $N \times N$ unitary matrices. The results from these calculations are summarized in Tab. 5.1 and Fig. 5.2. We observe that the $M = 3$ and $M = 4$ bounds from semidefinite programming are saturated at relatively small system sizes, $N = 5$ and $N = 8$ respectively.

5.2.5 Temporal versus spatial correlations

Leggett-Garg inequalities are often referred to as 'Bell inequalities in time'; in addition, it is known the Lüders bound for the n-term generalization, for even n, of the original Leggett-Garg inequality (5.17) coincides with the Tsirelson bound [24] for the corresponding Bell inequalities [93, 76], and noncontextuality inequalities [87]. It is therefore a natural question whether the above general measurement scheme can provide excess quantum violation of Bell inequalities. The answer, however, is negative as can be easily deduced directly from the Tsirelson's proof of the quantum bound [74] or by noticing that the commutativity of the measurements, even when performed sequentially as in

SDP			MAX										
M	$K_3^{\max}$		M	N	$K_3^{\max}$		M	N	$K_3^{\max}$		M	N	$K_3^{\max}$
2	$\frac{3}{2}$		3	3	2.1547		4	4	2.3693		5	5	2.5166
3	2.211507		3	4	2.1736		4	5	2.3877		5	6	2.5312
4	2.454629		3	5	2.2115		4	6	2.4181		5	7	2.5459
5	2.579333		3	6	2.2115		4	7	2.4315		5	8	2.5506
6	2.656005		3	7	2.2115		4	8	2.4545		5	9	2.5545

Table 5.1: The maximum value of the LGI parameter K_3 as a function of system size N and number of projectors M. The leftmost results are from the semi-definite programming (SDP) approach, whilst the rest are from direct maximisation (MAX) with fixed N and M. Here the value assignments $q_m = 1 - 2\delta_{m,-j}$ were used. In general, the bound changes for different assignments, but except for the case $M = 6$, the above choice was found to give the maximum violation.

contextuality t
rule is used.

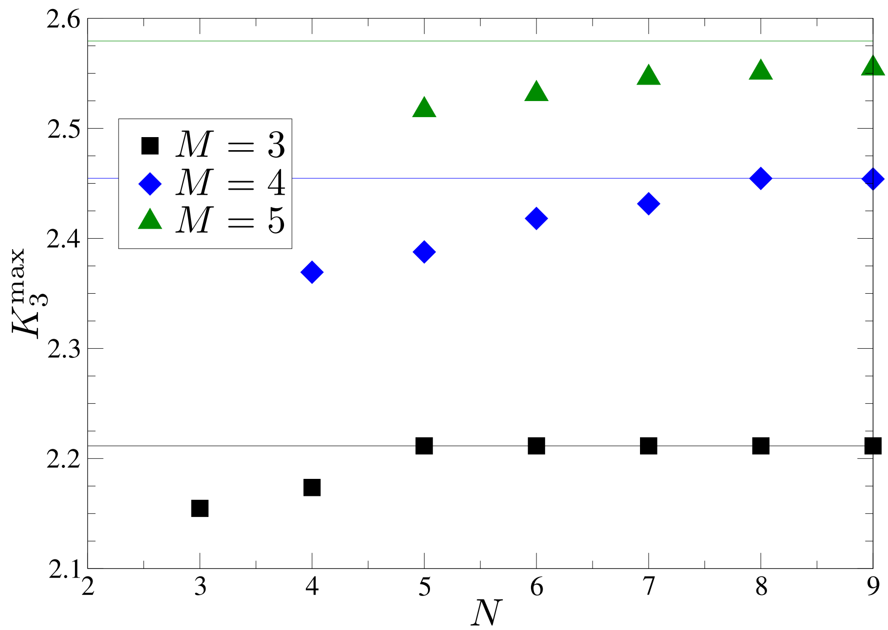


Figure 5.2: A plot of the data in Tab.5.1. The maximum values for each M (from SDP) are shown as straight lines.

5.3 Details of the calculations

5.3.1 Alternative proof of Observation 1

For an alternative proof of Observation 1, we need the following Proposition:

Proposition 3. *If two dichotomic measurements on a two-dimensional quantum system commute* $[A_i, A_{i+1}] = 0$, *then either*

(a) *one of the observables is proportional to the identity,* $A_i = \pm\mathbb{1}$ *or* $A_{i+1} = \pm\mathbb{1}$ *or,*

(b) *the product of the two observables is proportional to the identity,* $A_i A_{i+1} = \pm\mathbb{1}$.

Proof. This fact can easily be checked: The observables A_i and A_{i+1} are diagonal in the same basis and the entries on the diagonal can only be ± 1. Then, only the two cases outlined above are possible. □

Alternative proof of Observation 1. With the help of Prop. 3 one can consider each term of the KCBS inequality and make there six possible replacements. For instance, the term $\langle AB\rangle$ may be replaced by $\langle AB\rangle \mapsto \pm\langle B\rangle$ (if one sets $A \mapsto \pm\mathbb{1}$) or $\langle AB\rangle \mapsto \pm\langle A\rangle$ (if one sets $B \mapsto \pm\mathbb{1}$) or $\langle AB\rangle \mapsto \pm 1$. This results in a finite set of $6^5 = 7776$ possible replacements. Some of them are contradictory and can be disregarded, e.g., if one sets $B \mapsto \mathbb{1}$ from the term $\langle AB\rangle$ and $C \mapsto \mathbb{1}$ from the term $\langle CD\rangle$, then one cannot set $\langle BC\rangle \mapsto -1$ anymore. For the remaining replacements, one can directly check with a computer that the $\langle\chi_{\text{KCBS}}\rangle$ reduces to the classical bound. □

5.3.2 Detailed discussion of the N-cycle inequalities

First, we prove the following statement:

Proposition 4. *Consider the generalized KCBS operator*

$$\langle\chi_N\rangle = \sum_{i=1}^{N-1} \langle A_i A_{i+1}\rangle - \langle A_N A_1\rangle \tag{5.23}$$

for N even, where the A_i are dichotomic observables on a three-dimensional system, which are not proportional to the identity. Furthermore, the commuting pairs should not be equal, that is $A_i \neq A_{i+1}$. Then, the bound

$$\langle\chi_N\rangle \geq -(N-2) \tag{5.24}$$

holds.

Proof. From the conditions, it follows that the observables have to be of the form $A_i = \pm(\mathbb{1} - 2|a_i\rangle\langle a_i|)$ with $\langle a_i||a_{i+1}\rangle = 0$. This implies that the sequential measurements can be rephrased via $A_i A_{i+1} = \pm(\mathbb{1} - 2|a_i\rangle\langle a_i| - 2|a_{i+1}\rangle\langle a_{i+1}|)$. Let us first assume that the signs in front of the A_i are alternating, that is, $A_i = +(\mathbb{1} - 2|a_i\rangle\langle a_i|)$ for odd i and $A_i = -(\mathbb{1} - 2|a_i\rangle\langle a_i|)$ for even i. Then, a direct calculation leads to

$$\langle\chi_N\rangle = -(N-2) + 4\langle\sum_{k=2}^{N-1} |a_k\rangle\langle a_k|\rangle. \tag{5.25}$$

From this, $\langle\chi_N\rangle \geq -(N-2)$ follows, since the operator in the sum is positive semidefinite.

A general distribution of signs for the A_i results in a certain distribution of signs for the $A_i A_{i+1}$. If I denotes the set of index pairs $(k, k+1)$, where $A_k A_{k+1} = +(\mathbb{1} - 2|a_k\rangle\langle a_k| - 2|a_{k+1}\rangle\langle a_{k+1}|)$, then I has always an odd number of elements. We can then write:

$$\langle\chi_N\rangle = -(N-2) + 2(|I|-1) + 4\langle\sum_{k=1}^{N} \alpha_k |a_k\rangle\langle a_k|\rangle \tag{5.26}$$

where $\alpha_k = 1$ if both $(k, k+1) \notin I$ and $(k-1, k) \notin I$, $\alpha_k = 0$ if either $(k, k+1) \in I, (k-1, k) \notin I$ or $(k, k+1) \notin I, (k-1, k) \in I$, and $\alpha_k = -1$ if both $(k, k+1) \in I$ and $(k-1, k) \in I$.

It remains to show that the last two terms are non-negative. The main idea to prove this is to use the fact that an operator like $X = \mathbb{1} - |a_i\rangle\langle a_i| - |a_{i+1}\rangle\langle a_{i+1}|$ is positive semidefinite, since $|a_i\rangle$ and $|a_{i+1}\rangle$ are orthogonal.

More explicitly, let us first consider the case where the index pairs in I are connected and distinguish different cases for the number of elements in I. If $|I| = 1$, there are no k with $\alpha_k = -1$, so $2(|I|-1) + 4\langle\sum_{k=1}^N \alpha_k |a_k\rangle\langle a_k|\rangle \geq 0$. If $|I| = 2$, then $I = \{(i-1,i),(i,i+1)\}$ and there is a single $\alpha_i = -1$. In this case, one has $2|I| + 4\langle\sum_{k=1}^N \alpha_k |a_k\rangle\langle a_k|\rangle \geq 0$. This is not yet the desired bound, but it will be useful later.

If $|I| = 3$, then $I = \{(i-1,i),(i,i+1),(i+1,i+2)\}$ and we have $\alpha_i = \alpha_{i+1} = -1$. But now, the fact that $X = \mathbb{1} - |a_i\rangle\langle a_i| - |a_{i+1}\rangle\langle a_{i+1}| \geq 0$ directly implies that $2(|I|-1) + 4\langle\sum_{k=1}^N \alpha_k |a_k\rangle\langle a_k|\rangle \geq 0$. If $|I| = 4$ there are three $\alpha_k = -1$ and we can use $X \geq 0$ two times, showing that again $2|I| + 4\langle\sum_{k=1}^N \alpha_k |a_k\rangle\langle a_k|\rangle \geq 0$. All this can be iterated, resulting in two different bounds, for $|I|$ odd and $|I|$ even.

To complete the proof, we have to consider a general I which does not necessarily form a single block. One can then consider the different blocks and, since $|I|$ is odd, at least one of the blocks contains an odd number of index pairs. Then, summing up the bound for the single blocks leads to $2(|I|-1) + 4\langle\sum_{k=1}^N \alpha_k |a_k\rangle\langle a_k|\rangle \geq 0$. □

Finally, in order to justify Eq. (5.6) for the three-dimensional case, we have to discuss what happens if one of the observables is proportional to the identity. However, then the mean value $\langle\chi_N\rangle$ reduces to inequalities which will be discussed later (see Prop. 7 in Sect. 5.3.4).

5.3.3 Detailed proof of Observation 2

For computing the minimal value in two-dimensional systems, we need the following proposition. Note that the resulting value has been reported before [110], so the main task is to prove rigorously that this is indeed optimal.

Proposition 5. *Let $|\mathfrak{a}_i\rangle \in \mathbb{R}^3$ be normalized real three-dimensional vectors and define*

$$\chi_N = \sum_{i=1}^N \langle\mathfrak{a}_i||\mathfrak{a}_{i+1}\rangle \text{ for } N \text{ odd}, \tag{5.27a}$$

$$\chi_N = -\langle\mathfrak{a}_1||\mathfrak{a}_2\rangle + \sum_{i=2}^N \langle\mathfrak{a}_i||\mathfrak{a}_{i+1}\rangle \text{ for } N \text{ even}. \tag{5.27b}$$

Then we have

$$\chi_N \geq -N\cos(\frac{\pi}{N}). \tag{5.28}$$

Proof. We write $|\mathfrak{a}_i\rangle = \{\cos(\alpha_i), \sin(\alpha_i)\cos(\beta_i), \sin(\alpha_i)\sin(\beta_i)\}$ and then we have

$$\chi_N = \sum_{i=1}^N [\pm]\Big[\cos(\alpha_i)\cos(\alpha_{i+1}) + \cos(\beta_i - \beta_{i+1})\sin(\alpha_i)\sin(\alpha_{i+1})\Big], \tag{5.29}$$

where the symbol $[\pm]$ denotes the possibly changing sign of the term with $i = 1$. Let us first explain why the minimum of this expression can be obtained by setting all the $\beta_i = 0$. Without losing generality, we can assume that $|\mathfrak{a}_1\rangle$ points in the x-direction, i.e., $\alpha_1 = 0$ and $\sin(\alpha_1) = 0$. Then, only $N-2$ terms of the type $\sin(\alpha_i)\sin(\alpha_{i+1})$ remain and all of them have a positive prefactor. For given values of β_i we can choose the signs of $\alpha_2, \ldots, \alpha_{N-1}$ such that all these terms are negative, while the other parts of the expression are not affected by this. Then, it is clearly optimal to choose $\beta_2 = \beta_3 = \ldots = \beta_N = 0$. This means that all the vectors lie in the x-y-plane.

Having set all $\beta_i = 0$, the expression is simplified to $\chi_N = \sum_{i=1}^N [\pm]\cos(\alpha_i - \alpha_{i+1})$. We use the notation $\delta_i = \alpha_i - \alpha_{i+1}$ and minimize $\sum_{i=1}^N [\pm]\cos(\delta_i)$ under the constraint $\sum_{i=1}^N \delta_i = 0$. Using Lagrange multipliers, it follows that $[\pm]\sin(\delta_i) = \lambda$ for all i.

If N is odd, this means that we can express any δ_i as $\delta_i = \pi/2 \pm \vartheta + 2\pi k_i$ with $\vartheta \geq 0$. From $\cos(\pi/2 + \vartheta + 2\pi k_i) = -\cos(\pi/2 - \vartheta + 2\pi k_i)$, it follows that the sign in front of the ϑ should be identical for all δ_i, otherwise, the expression is not minimized. Let us first consider the case that all signs a positive. From the condition $\sum_{i=1}^N \delta_i = 0$, it follows that $N(\pi/2) + N\vartheta + 2\pi K = 0$, with $K = \sum_{i=1}^N k_i$. Since we wish to minimize χ_N, the angles δ_i should be as close as possible to π, which means that $|\vartheta - \pi/2|$ should be minimal. This leads to the result that one has to choose $K = -(N \pm 1)/2$. Computing the corresponding ϑ leads to $\vartheta = \pi/2 \pm \pi/N$, which results in Eq. (5.28). If the signs in front of all ϑ are negative, one can make a similar argument, but this time has to minimize $|\delta_i + \pi|$ or $|\vartheta - 3\pi/2|$. This leads to the same solutions.

If N is even, one has for $i = 2, \ldots, N$ again $\delta_i = \pi/2 \pm \vartheta + 2\pi k_i$ and the first δ_1 can be written as $\delta_1 = -\pi/2 \pm \vartheta + 2\pi k_1$. One can directly see that if the signs in front of ϑ is positive (negative) for all $i = 2, \ldots, N$ it has to be positive (negative) also for $i = 1$. A direct calculation as before leads to $\vartheta = \pi/2 \pm \pi/N$ and, again, to the same bound of Eq. (5.28). □

Proof of Observation 2. Let us first assume that none of the observables is proportional to the identity, and consider a single sequential measurement $\langle AB \rangle$ of two dichotomic noncommuting observables $A = |A^+\rangle\langle A^+| - |A^-\rangle\langle A^-| = P_+^A - P_-^A$ and $B = |B^+\rangle\langle B^+| - |B^-\rangle\langle B^-| = P_+^B - P_-^B$. We can also express $|A^+\rangle\langle A^+|$ and $|B^+\rangle\langle B^+|$ in terms of their Bloch vectors $|\mathfrak{a}\rangle$ and $|\mathfrak{b}\rangle$. Then, we have that

$$\langle AB \rangle = 2|\langle A^+|B^+\rangle|^2 - 1 = \langle \mathfrak{a}|\mathfrak{b}\rangle. \tag{5.30}$$

Note that this means that the mean value $\langle AB \rangle$ is independent of the initial quantum state. To see this relation, we write $\langle AB \rangle = \mathrm{tr}(P_+^B P_+^A \varrho P_+^A P_+^B) - \mathrm{tr}(P_-^B P_+^A \varrho P_+^A P_-^B) - \mathrm{tr}(P_+^B P_-^A \varrho P_-^A P_+^B) + \mathrm{tr}(P_-^B P_-^A \varrho P_-^A P_-^B)$. Using the fact that in a two-dimensional system $|\langle A^+|B^+\rangle|^2 = |\langle A^-|B^-\rangle|^2$ and $|\langle A^-|B^+\rangle|^2 = |\langle A^+|B^-\rangle|^2$ holds, and $\mathrm{tr}(\varrho) = 1$, this can directly be simplified to the above expression. Using the above expression, we can write $\langle \chi_{\mathrm{KCBS}} \rangle = \sum_{i=1}^5 \langle \mathfrak{a}_i|\mathfrak{a}_{i+1}\rangle$. Then, Prop. 5 proves the desired bound.

It remains to discuss the case where one or more observables in the KCBS inequality are proportional to the identity. Let us first assume that only one observable, say A_1 is proportional to the identity. Then, if the Bloch vector of ϱ is denoted by $|\mathfrak{r}\rangle$ a direct calculation shows that the KCBS operator reads

$$\langle \chi_{\mathrm{KCBS}} \rangle = \langle \mathfrak{r}|\mathfrak{a}_2\rangle + \sum_{i=2}^4 \langle \mathfrak{a}_i|\mathfrak{a}_{i+1}\rangle + \langle \mathfrak{a}_5|\mathfrak{r}\rangle, \tag{5.31}$$

and Prop. 5 proves again the claim. If two observables A_i and A_j are proportional to the identity, the same rewriting can be applied, if A_i and A_j do not occur jointly in one correlation term. This is the case if $j \neq i \pm 1$. In the other case (say, $A_1 = \mathbb{1}$ and $A_2 = -\mathbb{1}$), one has $\langle A_1 A_2 \rangle = -1$ and can rewrite

$$\langle \chi_{\mathrm{KCBS}} \rangle = -1 - \langle \mathfrak{r}|\mathfrak{a}_2\rangle + \sum_{i=3}^4 \langle \mathfrak{a}_i|\mathfrak{a}_{i+1}\rangle + \langle \mathfrak{a}_4|\mathfrak{r}\rangle, \tag{5.32}$$

and Prop. 5 implies that $\langle \chi_{\mathrm{KCBS}} \rangle \geq -4\cos(\pi/4) - 1 = -2\sqrt{2} - 1 > -5\cos(\pi/5) = -5(1+\sqrt{5})/4$. If more than two observables are proportional to the identity, the bound can be proven similarly. □

5.3.4 Proof of Observation 3

We need a whole sequence of Propositions:

Proposition 6. *If one has three dichotomic measurements $A_i, i = 1, 2, 3$ on a three-dimensional quantum system which commute pairwise $[A_i, A_j] = 0$, then either*

(a) one of the observables is proportional to the identity, $A_i = \pm\mathbb{1}$ for some i or,

(b) the product of two observables of the three observables is proportional to the identity, $A_i A_j = \pm\mathbb{1}$ for some pair i, j or,

(c) The product of all three observables is proportional to the identity, $A_1 A_1 A_3 = \pm\mathbb{1}$.

Note that these cases are not exclusive and that for a triple of observables several of these cases may apply at the same time.

Proof. This can be proven in the same way as Prop. 3, since all A_i are diagonal in the same basis. □

Proposition 7. *For sequences of dichotomic measurements the following inequalities hold:*

$$\eta_N \equiv \langle A_1 \rangle + \sum_{i=1}^{N-1} \langle A_i A_{i+1} \rangle - \langle A_N \rangle \leq N - 1. \tag{5.33}$$

Here, it is always assumed that two observables which occur in the same sequence commute. Moreover, if we define

$$\zeta_N \equiv \sum_{i=1}^{N} \langle A_i A_{i+1} \rangle - \langle A_N A_1 \rangle, \tag{5.34}$$

then we have

$$\zeta_N \leq N - 2 \tag{5.35}$$

in two-dimensional systems, while for three-dimensional systems.

$$\begin{aligned} \zeta_3 \leq 1; \qquad & \zeta_4 \leq 2, \\ \zeta_5 \leq \sqrt{5}(4 - \sqrt{5}), \qquad & \zeta_6 \leq 1 + \sqrt{5}(4 - \sqrt{5}) = 4(\sqrt{5} - 1), \end{aligned} \tag{5.36}$$

holds.

Proof. If we consider η_N for $N = 2$ both observables commute and the claim $\langle A_1 \rangle + \langle A_1 A_2 \rangle - \langle A_2 \rangle \leq 1$ is clear, as it holds for any eigenvector. The bounds for general η_N follow by induction, where in each step of the induction $\langle A_N A_{N+1} \rangle - \langle A_{N+1} \rangle \leq 1 - \langle A_N \rangle$ is used, but this is nothing but the bound for $N = 2$.

The bounds for ζ_N are just the ones derived for the generalized KCBS inequalities, see Eq. (4.12) in Sects. 5.1.1 and 5.3.2. □

Proposition 8. *Consider the PM square with dichotomic observables on a three-dimensional system, where for one column and one row only the case (c) in Prop. 6 applies. Then, one cannot violate the classical bound and one has $\langle \chi_{\text{PM}} \rangle \leq 4$.*

Proof. Let us consider the case that the condition holds for the first column and the first row, the other cases are analogous. Then, none of the observables A, B, C, a, α is proportional to the identity since, otherwise, case (a) in Prop. 6 would apply. These observables can all be written as

$$A = \pm(\mathbb{1} - 2|A\rangle\langle A|), \tag{5.37}$$

with some vector $|A\rangle$, and the vector $|A\rangle$ characterizes the observable A up to the total sign uniquely. In this notation, two observables X and Y commute if and only if the corresponding vectors $|X\rangle$ and $|Y\rangle$ are the same or orthogonal. For our situation, it follows that the vectors $|A\rangle, |B\rangle$, and $|C\rangle$ form an orthonormal basis of the three-dimensional space, since if two of them were the same, then for the first row also the case (b) in Prop. 6 would apply. Similarly, the vectors $|A\rangle, |a\rangle$ and $|\alpha\rangle$ form another orthonormal basis of the three-dimensional space. We can distinguish two cases:

Case 1: The vector $|B\rangle$ is neither orthogonal nor parallel to $|a\rangle$. From this, it follows that $|B\rangle$ is also neither orthogonal nor parallel to $|\alpha\rangle$ and similarly, $|C\rangle$ is neither orthogonal nor parallel to $|a\rangle$ and $|\alpha\rangle$ and vice versa.

Let us consider the observable b in the PM square. This observable can be proportional to the identity, but if this is not the case, the corresponding vector $|b\rangle$ has to be parallel or orthogonal to $|B\rangle$ and $|a\rangle$. Since $|B\rangle$ and $|a\rangle$ are neither orthogonal nor parallel, it has to be orthogonal to both, which means that it is parallel to $|A\rangle$. Consequently, the observable b is either proportional to the identity or proportional to A. Similarly, all the other observables β, c, and γ are either proportional to the identity or proportional to A.

Let us now consider the expectation value of the PM operator $\langle\chi_{\rm PM}\rangle$ for some quantum state ϱ. We denote this expectation value as $\langle\chi_{\rm PM}\rangle_\varrho$ in order to stress the dependence on ϱ. The observable A can be written as $A = P_+ - P_-$, where P_+ and P_- are the projectors onto the positive or negative eigenspace. One of these projectors is one-dimensional and equals $|A\rangle\langle A|$, the other other one is two-dimensional. For definiteness, let us take $P_+ = |A\rangle\langle A|$ and $P_- = \mathbb{1} - |A\rangle\langle A|$.

Instead of ϱ, we may consider the depolarized state $\sigma = p_+\varrho_+ + p_-\varrho_-$, with $\varrho_\pm = P_\pm\varrho P_\pm/p_\pm$ and $p_\pm = \mathrm{tr}(P_\pm\varrho P_\pm)$. Our first claim is that, in our situation,

$$\langle\chi_{\rm PM}\rangle_\varrho = \langle\chi_{\rm PM}\rangle_\sigma = p_+\langle\chi_{\rm PM}\rangle_{\varrho+} + p_-\langle\chi_{\rm PM}\rangle_{\varrho-}. \tag{5.38}$$

It suffices to prove this for all rows and columns separately. Since the observables in each column or row commute, we can first measure observables which might be proportional to A. For the first column and the first row the statement is clear: We first measure A and the result is the same for ϱ and σ. After the measurement of A, however, the state ϱ is projected either onto ϱ_+ or ϱ_-. Therefore, for the following measurements it does not matter whether the initial state was ϱ or σ. As an example for the other rows and columns, we consider the second column. Here, we can first measure β and then b and finally B. If β or b are proportional to A, then the statement is again clear. If both β and b are proportional to the identity, then the measurement of $\langle\beta b B\rangle_\varrho$ equals $\pm\langle B\rangle_\varrho$. Then, however, one can directly calculate that $\langle B\rangle_\varrho = \langle B\rangle_\sigma$, since B and A commute.

Having established the validity of Eq. (5.38), we proceed by showing that for for each term $\langle\chi_{\rm PM}\rangle_{\varrho+}$ and $\langle\chi_{\rm PM}\rangle_{\varrho-}$ separately the classical bound holds. For $\langle\chi_{\rm PM}\rangle_{\varrho+}$ this is clear: Since $P_+ = |A\rangle\langle A|$, we have that $\varrho_+ = |A\rangle\langle A|$ and $|A\rangle$ is an eigenvector of all observables occurring in the PM square. Therefore, the results obtained in $\langle\chi_{\rm PM}\rangle_{\varrho+}$ correspond to a classical assignment of ±1 to all observables, and $\langle\chi_{\rm PM}\rangle_{\varrho+} \leq 4$ follows. For the other term $\langle\chi_{\rm PM}\rangle_{\varrho-}$, the problem is effectively a two-dimensional one, and we can consider the restriction of the observables to the two-dimensional space, e.g., $\bar{A} = P_- A P_-$, etc. In this restricted space we have that $\bar{A}, \bar{b}, \bar{\beta}, \bar{c}$, and $\bar{\gamma}$ are all of them proportional to the identity and, therefore, result in a classical assignment ±1 independent of ϱ_-. Let us denote these assignments by $\hat{A}, \hat{b}, \hat{\beta}, \hat{c}$, and $\hat{\gamma}$. Then, it remains to be shown that

$$\begin{aligned}\mathcal{Z} \;=\; & \hat{A}\big[\langle\bar{B}\bar{C}\rangle_{\varrho-} + \langle\bar{\alpha}\bar{a}\rangle_{\varrho-}\big] + \hat{b}\hat{c}\langle\bar{a}\rangle_{\varrho-} \\ & + \hat{\beta}\hat{\gamma}\langle\bar{\alpha}\rangle_{\varrho-} + \hat{b}\hat{\beta}\langle\bar{B}\rangle_{\varrho-} - \hat{c}\hat{\gamma}\langle\bar{C}\rangle_{\varrho-} \leq 4\end{aligned} \tag{5.39}$$

for all classical assignments and for all states ϱ_-. For observables $\bar{B}$ and $\bar{C}$ we have furthermore that $\bar{B}\bar{C} = \pm\mathbb{1}$ (see Prop. 3), hence $\bar{B} = \pm\bar{C}$ and similarly $\bar{a} = \pm\bar{\alpha}$. If one wishes to maximize $\mathcal{Z}$

for the case $\hat{A} = +1$, one has to choose $\bar{B} = \bar{C}$ and $\bar{a} = \bar{\alpha}$. Then, the product of the four last terms in $\mathcal{Z}$ equals -1, and $\mathcal{Z} \leq 4$ holds. For the case $\hat{A} = -1$ one chooses $\bar{B} = -\bar{C}$ and $\bar{a} = -\bar{\alpha}$, but still the product of the four last terms in $\mathcal{Z}$ equals -1, and $\mathcal{Z} \leq 4$. This finishes the proof of the first case.

Case 2: The bases $|A\rangle, |B\rangle, |C\rangle$ and $|A\rangle, |a\rangle, |\alpha\rangle$ are (up to some permutations or signs) the same. For instance, we can have the case in which $|B\rangle = |a\rangle$ and $|C\rangle = |\alpha\rangle$; the other possibilities can be treated similarly.

In this case, since $|B\rangle$ and $|\alpha\rangle$ are orthogonal, the observable β has to be either proportional to the identity or proportional to A. For the same reason, c has to be either proportional to the identity or to A.

Let us first consider the case in which one of the observables β and c is proportional to A, say $\beta = \pm A$ for definiteness. Then, since $|\beta\rangle = |A\rangle$ and $|B\rangle$ are orthogonal, b can only be the identity or proportional to C. Similarly, γ can only be the identity or proportional to C. It follows that *all* nine observables in the PM square are diagonal in the basis $|A\rangle, |B\rangle, |C\rangle$, and all observables commute. Then, $\langle \chi_{\text{PM}} \rangle \leq 4$ follows, as this inequality holds in any eigenspace.

Second, let us consider the case in which β and c are both proportional to the identity. This results in fixed assignments $\hat{\beta}$ and $\hat{c}$ for them. Moreover, B and a differ only by a sign $\hat{\mu}$ (that is, $a = \hat{\mu}B$) and C and α differ only by a sign $\hat{\nu}$ (i.e., $\alpha = \hat{\nu}C$). So we have to consider

$$\begin{aligned} \mathcal{X} &= \langle ABC \rangle + \hat{\mu}\hat{\nu}\langle ABC \rangle + \hat{\beta}\langle Bb \rangle \\ &\quad + \hat{\mu}\hat{c}\langle Bb \rangle + \hat{\nu}\hat{\beta}\langle C\gamma \rangle - \hat{c}\langle C\gamma \rangle. \end{aligned} \tag{5.40}$$

In order to achieve $\mathcal{X} > 4$ one has to choose $\hat{\mu} = \hat{\nu}$, $\hat{\beta} = \hat{\mu}\hat{c}$, and $\hat{c} = -\hat{\nu}\hat{\beta}$. However, the later is equivalent to $\hat{\beta} = -\hat{\nu}\hat{c}$, showing that this assignment is not possible. Therefore, $\mathcal{X} \leq 4$ has to hold. This finishes the proof of the second case. $\square$

Proposition 9. *Consider the PM square with dichotomic observables on a three-dimensional system, where for one column (or one row) only the case (c) in Prop. 6 applies. Then, one cannot violate the classical bound and one has $\langle \chi_{\text{PM}} \rangle \leq 4$.*

Proof. We assume that the condition holds for the first column. Then, none of the observables A, a, and α are proportional to the identity, and the corresponding vectors $|A\rangle, |a\rangle$, and $|\alpha\rangle$ form an orthonormal basis of the three-dimensional space.

The idea of our proof is to consider possible other observables in the PM square, which are not proportional to the identity, but also not proportional to A, a, or α. We will see that there are not many possibilities for the observables, and in all cases the bound $\langle \chi_{\text{PM}} \rangle \leq 4$ can be proved explicitly.

First, consider the case that there *all* nontrivial observables in the PM square are proportional to A, a, or α. This means that all observables in the PM square are diagonal in the basis defined by $|A\rangle, |a\rangle$, and $|\alpha\rangle$, and all observables commute. But then the bound $\langle \chi_{\text{PM}} \rangle \leq 4$ is clear.

Second, consider the case that there are several nontrivial observables, which are *not* proportional to A, a, or α. Without losing generality, we can assume that the first of these observables is B. This implies that $|B\rangle$ is orthogonal to $|A\rangle$ and lies in the plane spanned by $|a\rangle$ and $|\alpha\rangle$, but $|a\rangle \neq |B\rangle \neq |\alpha\rangle$. It follows for the observables b and β that they can only be proportional to the identity or to A (see Case 1 in Prop. 8). We denote this as $b = \hat{b}[A]$, where $[A] = A$ or $\mathbb{1}$, and $\hat{b}$ denotes the proper sign, i.e., $b = \hat{b}A$ or $b = \hat{b}\mathbb{1}$. Similarly, we write $\beta = \hat{\beta}[A]$.

Let us assume that there is a second nontrivial observable which is not proportional to A, a, or α (but it might be proportional to B). We can distinguish three cases:

(i) First, this observable can be given by C and C is not proportional to B. Then, this is exactly the situation of Case 1 in Prop. 8, and $\langle \chi_{\text{PM}} \rangle \leq 4$ follows.

(ii) Second, this observable can be given by C. However, C is proportional to B. Then, $c = \hat{c}[A]$ and $\gamma = \hat{\gamma}[A]$ follows. Now the proof can proceed as in Case 1 of Prop. 8. One arrives to the same Eq. (5.39), with the extra condition that $\bar{B} = \pm\bar{C}$, which was deduced after Eq. (5.39) anyway. Therefore, $\langle\chi_{\mathrm{PM}}\rangle \leq 4$ has to hold.

(iii) Third, this observable can be given by c. Then, it cannot be proportional to B, since $|B\rangle$ is not orthogonal to $|a\rangle$. It first follows that $C = \hat{C}[a]$ and $\gamma = \hat{\gamma}[a]$. Combined with the properties of B, one finds that $C = \hat{C}\mathbb{1}$ and $b = \hat{b}\mathbb{1}$ has to hold. Then, the PM inequality reads

$$\begin{aligned}\mathcal{Y} &= \langle A\alpha a\rangle + \langle B(A\hat{C} + \hat{b}\hat{\beta}[A])\rangle \\ &\quad +\hat{\beta}\hat{\gamma}\langle\alpha[A][a]\rangle + \langle c(\hat{b}a - \hat{C}\hat{\gamma}[a])\rangle. \end{aligned} \tag{5.41}$$

In this expression, the observables B and c occur only in a single term and a single context. Therefore, for any quantum state, we can obtain an upper bound on $\mathcal{Y}$ by replacing $B \mapsto \pm\mathbb{1}$ and $c \mapsto \pm\mathbb{1}$ with appropriately chosen signs. However, with this replacement, all observables occurring in $\mathcal{Y}$ are diagonal in the basis defined by $|A\rangle, |a\rangle$, and $|\alpha\rangle$, and $\mathcal{Y} = \langle\chi_{\mathrm{PM}}\rangle \leq 4$ follows.

In summary, the discussion of the cases (i), (ii), and (iii) has shown the following: It is not possible to have three nontrivial observables in the PM square, which are all of them not proportional to A, a, or α. If one has two of such observables, then the classical bound has been proven.

It remains to be discussed what happen if one has only one observable (say, B), which is not proportional to A, a, or α. However, then the PM inequality can be written similarly as in Eq. (5.41), and B occurs in a single context. We can set again $B \mapsto \pm\mathbb{1}$ and the claim follows. □

Finally, we can prove our Observation 3:

Proof of Observation 3. Prop. 8 and Prop. 8 solve the problem, if case (c) in one column or row happens. Therefore, we can assume that in all columns and all rows only the cases (a) or (b) from Prop. 6 apply. However, in these cases, we obtain a simple replacement rule: For case (a), one of the observables has to be replaced with a classical value ± 1 and, for case (b), one of the observables can be replaced by a different one from the same row or column. In both cases, the PM inequality is simplified.

For case (a), there are six possible replacement rules, as one of the three observables must be replaced by ± 1. Similarly, for case (b), there are six replacement rules. Therefore, one obtains a finite number, namely $(6+6)^6$ possible replacements. As in the case of the KCBS inequality (see the alternative proof of Observation 1 in Sect. 5.3.1), some of them lead to contradictions (e.g., one may try to set $A = +\mathbb{1}$ from the first column, but $A = -\mathbb{1}$ holds due to the rule from the first row). Taking this into account, one can perform an exhaustive search of all possibilities, preferably by computer. For all cases, either the classical bound holds trivially (e.g., because the assignments require already, that one row is -1) or the PM inequality can be reduced, up to some constant, to one of the inequalities in Prop. 6. In most cases, one obtains the classical bound. However, in some cases, the PM inequality is reduced to $\langle\chi_{\mathrm{PM}}\rangle = \zeta_5 + 1$ or $\langle\chi_{\mathrm{PM}}\rangle = \zeta_6$. To give an example, one may consider the square

$$\begin{bmatrix} A & B & C \\ a & b & c \\ \alpha & \beta & \gamma \end{bmatrix} = \begin{bmatrix} A & \mathbb{1} & C \\ a & b & \mathbb{1} \\ \mathbb{1} & \beta & \gamma \end{bmatrix}, \tag{5.42}$$

which results in $\langle\chi_{\mathrm{PM}}\rangle = \zeta_6$ for appropriately chosen A_i. Therefore, from Prop. 7 follows that in three dimensions $\langle\chi_{\mathrm{PM}}\rangle = 4(\sqrt{5}-1) \approx 4.94$ holds and can indeed be reached. □

5.3.5 Imperfect measurements

In this section we discuss the noise robustness of Observation 4. In the first subsection, we prove that Observation 4 also holds for the model of noisy measurements explained above. In the second subsection, we discuss a noise model that reproduces the probabilities of the most general POVM.

Noisy measurements

In order to explain the probabilities from a noisy measurement, we first consider the following measurement model: Instead of performing the projective measurement A, one of two possible actions are taken:

(a) with a probability p_A the projective measurement is performed, or

(b) with a probability $1 - p_A$ a completely random outcome ± 1 is assigned independently of the initial state. Here, the results $+1$ and -1 occur with equal probability.

In case (b), after the assignment the physical system is left in one of two possible states ϱ^+ or ϱ^-, depending on the assignment. We will not make any assumptions on $\varrho^\pm$.

Before formulating and proving a bound on $\langle\chi_{\mathrm{PM}}\rangle$ in this scenario, it is useful to discuss the structure of $\langle\chi_{\mathrm{PM}}\rangle$ for the measurement model. A single measurement sequence $\langle ABC\rangle$ is split into eight terms: With a prefactor $p_A p_B p_C$ one has the value, which is obtained, if all measurements are projective; with a prefactor $p_A p_B(1-p_C)$ one has the value, where A and B are projective, and C is a random assignment, etc. It follows that the total mean value $\langle\chi_{\mathrm{PM}}\rangle$ is an affine function in the probability p_A (if all other parameters are fixed) and also in all other probabilities p_X for the other measurements. Consequently, the maximum of $\langle\chi_{\mathrm{PM}}\rangle$ is attained either at $p_A = 1$ or $p_A = 0$, and similarly for all the measurements. Therefore, for maximizing $\langle\chi_{\mathrm{PM}}\rangle$ it suffices to consider the finite set of cases where, for each observable, either always possibility (a) or always possibility (b) is taken. We can formulate:

Proposition 10. *Consider noisy measurements as described above. Then, the bound from Observation 4*

$$\langle\chi_{\mathrm{PM}}\rangle \leq 3\sqrt{3} \tag{5.43}$$

holds.

Proof. As discussed above, we only have to discuss a finite number of cases. Let us consider a single term $\langle ABC\rangle$. If C is a random assignment, then $\langle ABC\rangle = 0$, independently how A and B are realized. It follows that if C, β or a are random assignments, then $\langle\chi_{\mathrm{PM}}\rangle \leq 4$.

On the other hand, if A is a random assignment, then $\langle ABC\rangle = 0$ as well: (i) If B and C are projective, then the measurement of B and C results in the state independent mean value $\langle BC\rangle$. This value is independent of the state $\varrho^\pm$ remaining after the assignment of A, hence $\langle ABC\rangle = \langle AB\rangle - \langle AB\rangle = 0$. (ii) If B is a random assignment, one can also directly calculate that $\langle ABC\rangle = 0$ and the case that (iii) C is a random assignment has been discussed already. Consequently, if A, b, or γ are random assignments, then $\langle\chi_{\mathrm{PM}}\rangle \leq 4$.

It remains to discuss the case that B, c, or α are random assignments while all other measurements are projective. First, one can directly calculate that if A, C are projective, and B is a random assignment, then

$$\langle ABC\rangle = \mathrm{tr}(\varrho A)\mathrm{tr}(CX), \tag{5.44}$$

with $X = (\varrho^+ - \varrho^-)/2$. If X is expressed in terms of Pauli matrices, then the length of its Bloch vector does not exceed one, since the Bloch vectors of $\varrho^\pm$ are subnormalized.

The estimate of $\langle\chi_{\mathrm{PM}}\rangle$ can now proceed as in the proof of Observation 4, and one arrives at the situation of Prop. 5 in Sect.5.3.3, where now the vectors are subnormalized, and not necessarily normalized. But still the bound from Prop. 5 is valid: If the smallest vector in χ_6 has a length ω, one can directly see that $\chi_6 \geq \omega[-N\cos(\pi/N)] - (1-\omega)4$. This proves Prop. 10. $\square$

More general POVMs

Now we consider a general dichotomic positive operator valued measure (POVM) on a qubit system. This is characterized by two effects E^+ and E^-, where $E^+ + E^- = \mathbb{1}$ and the probabilities of the measurement results are $p^+ = \mathrm{tr}(\varrho E^+)$ and $p^- = \mathrm{tr}(\varrho E^-)$.

These effects have to commute and one can write $E^+ = \alpha|0\rangle\langle 0| + \beta|1\rangle\langle 1|$ and $E^- = \gamma|0\rangle\langle 0| + \delta|1\rangle\langle 1|$ in an appropriate basis. We can assume that $\alpha \geq \beta$ and consequently $\delta \geq \gamma$. Furthermore, it is no restriction to choose $\beta \leq \gamma$. Then, the effects can be written as $E^+ = \beta\mathbb{1} + (\alpha - \beta)|0\rangle\langle 0|$ and $E^- = \beta\mathbb{1} + (\gamma - \beta)\mathbb{1} + (\alpha - \beta)|1\rangle\langle 1|$. This means that one can interpret the probabilities of the POVM as coming from the following procedure: With a probability of 2β one assigns a random outcome, with a probability of $\gamma - \beta$ one assigns the fixed value -1, and with a probability of $(\alpha - \beta)$ one performs the projective measurement.

This motivates the following measurement model: Instead of performing the projective measurement A, one of three possible actions are taken:

(i) with a probability p_1^A the projective measurement is performed, or

(ii) with a probability p_2^A a fixed outcome ± 1 is assigned independently of the initial state. After this announcement, the state is left in the corresponding eigenstate of A, or

(iii) with a probability p_3^A a completely random outcome ± 1 is assigned independently of the initial state.

As above, in case (iii), the physical system is left in one of two possible states ϱ^+ or ϱ^-, but we will not make any assumptions on $\varrho^\pm$. For this measurement model, we have:

Proposition 11. *In the noise model described above, the PM operator is bounded by*

$$\langle \chi_{\mathrm{PM}} \rangle \leq 1 + \sqrt{9 + 6\sqrt{3}} \approx 5.404. \tag{5.45}$$

Proof. As in the proof of Prop. 10, we only have to consider a finite set of cases. Let us first discuss the situation, where for each measurement only the possibilities (i) and (ii) are taken.

First, we have to derive some formulas for sequential measurements. The reason is that, if the option (ii) is chosen, then the original formula for sequential measurements, Eq. (5.8), is not appropriate any more and different formulas have to be used.

In the following, we write $A = (\pm)_A$ if A is a fixed assignment as described in possibility (ii) above. If not explicitly stated otherwise, the observables are measured as projective measurements. Then one can directly calculate that

$$\langle ABC \rangle = (\pm)_A \langle BC \rangle \text{ if } A = (\pm)_A, \tag{5.46a}$$
$$\langle ABC \rangle = \mathrm{tr}(\varrho A)\langle BC \rangle \text{ if } B = (\pm)_B, \tag{5.46b}$$
$$\langle ABC \rangle = (\pm)_C \langle AB \rangle \text{ if } C = (\pm)_C, \tag{5.46c}$$

Note that in Eq. (5.46b) there is no deviation from the usual formula Eq. (5.15). Furthermore, we have

$$\langle ABC \rangle = (\pm)_A (\pm)_B \mathrm{tr}(C|B^\pm\rangle\langle B^\pm|) = (\pm)_A \langle BC \rangle \quad \text{if } A = (\pm)_A \text{ and } B = (\pm)_B, \tag{5.47a}$$
$$\langle ABC \rangle = (\pm)_A (\pm)_C \mathrm{tr}(B|A^\pm\rangle\langle A^\pm|) = (\pm)_C \langle AB \rangle \quad \text{if } A = (\pm)_A \text{ and } C = (\pm)_C, \tag{5.47b}$$
$$\langle ABC \rangle = (\pm)_B (\pm)_C \mathrm{tr}(\varrho A) \quad \text{if } B = (\pm)_B \text{ and } C = (\pm)_C. \tag{5.47c}$$

In Eqs. (5.47a) and (5.47b), $|B^{\pm}\rangle$ and $|A^{\pm}\rangle$ denote the eigenstates of B and A, which are left after the fixed assignment.

Equipped with these rules, we can discuss the different cases. First, from Eqs. (5.46a), (5.46b), and (5.47a) it follows that the proof of Observation 4 does not change, if fixed assignments are made only on the observables which are measured at first or second position of a sequence (i.e., the observables A, b, γ, B, c, and α).

However, the structure of the inequality changes if one of the last measurements is a fixed assignment. To give an example, consider the case that the measurement β is a fixed assignment [case (ii) above], while all other measurements are projective [case (i) above]. Using Eq. (5.46c) we have to estimate

$$\begin{aligned} \mathcal{X} = &\langle A\rangle\langle BC\rangle + \langle A\rangle\langle \alpha a\rangle + \langle b\rangle\langle ca\rangle \\ &+ \langle bB\rangle(\pm)_\beta + \langle \gamma\alpha\rangle(\pm)_\beta - \langle\gamma\rangle\langle cC\rangle. \end{aligned} \tag{5.48}$$

On can directly see that it suffices to estimate

$$\begin{aligned} \mathcal{X}' = &\langle B|C\rangle + \langle \alpha|a\rangle + \langle \varrho|b\rangle\langle c|a\rangle \\ &+ \langle b|B\rangle + \langle \gamma|\alpha\rangle - \langle\varrho|\gamma\rangle\langle c|C\rangle, \end{aligned} \tag{5.49}$$

where all expressions should be understood as scalar products of the corresponding Bloch vectors. Then, a direct optimization over the three-dimensional Bloch vectors proves that here

$$\mathcal{X}' \leq 1 + \sqrt{9 + 6\sqrt{3}} \approx 5.404 \tag{5.50}$$

holds. In general, the observables β, C, or a are the possible third measurements in a sequence. One can directly check that, if one or several of them are fixed assignments, then an expression analogue to Eq. (5.48) arises and the bound of Eq. (5.50) holds. Finally, if some of the β, C, or a are fixed assignments and, in addition, some of the A, b, γ, B, c, and α are fixed assignments, then the comparison between Eq. (5.46c) and Eqs. (5.47b) and (5.47c) shows that no novel types of expressions occur.

It remains to discuss the case where not only the possibilities (i) and (ii) occur, but for one or more measurements also a random assignment [possibility (iii)] is realized. As in the proof of Prop. 10, one finds that only the cases where the second measurements (B, c, and α) are random are interesting. In addition to Eq. (5.44) one finds that $\langle ABC\rangle = (\pm_A)\mathrm{tr}(CX)$ if B is random and A is a fixed assignment, and $\langle ABC\rangle = 0$ if B is random and C is a fixed assignment. This shows that no new expressions occur, and proves the claim. □

Finally, we would like to add two remarks. First, it should be stressed that the presented noise model still makes assumptions about the measurement, especially about the post measurement state. Therefore, it is not the most general measurement, and we do not claim that the resulting dimension witnesses are device-independent.

Second, we would like to emphasize that the chosen order of the measurements in the definition in Eq. (5.10) is important for the proof of the bounds for noisy measurements: For other orders, it is not clear whether the dimension witnesses are robust against imperfections. In fact, for some choices one finds that the resulting inequalities are *not* robust against imperfections: Consider, for instance, a measurement order, where one observable (say, γ for definiteness) is the second observable in one context and the third observable in the other context. Furthermore, assume that γ is an assignment [case (iii) above], while all other measurements are projective. Then, we have to use Eq. (5.46b) for the first context of γ, and Eq. (5.46c) for the second context. In Eq. (5.46b) there is no difference to the usual formula, especially the formula does not depend on the value assigned to γ. Eq. (5.46c), however, depends on this value. This means that, for one term in the PM inequality, the sign can be changed arbitrarily and so $\langle\chi_{\mathrm{PM}}\rangle = 6$ can be reached.

5.3.6 Asymptotic value of the Leggett-Garg correlator for the precessing spin model

We here derive the expression for the correlation functions for the spin model with measurement times $\Omega t_1 = \pi$, $\Omega t_2 = \frac{3}{2}\pi$ and $\Omega t_3 = 2\pi$. Defining $R = e^{-i\frac{\pi}{2}J_x}$, the relevant time-evolution operators can be written $U(t_1) = R^2$, $U(t_2) = R^3$, and $U(t_2 - t_1) = U(t_3 - t_2) = R$. Starting in state $|-j\rangle$ (we use the shorthand $|m\rangle \equiv |m; j\rangle$ here), the correlation functions read

$$
\begin{aligned}
C_{21} &= \sum_{n,m=-j}^{j} q_n q_m |\langle m|R|n\rangle|^2 |\langle n|R^2|-j\rangle|^2; \\
C_{31} &= \sum_{n,m=-j}^{j} q_n q_m |\langle m|R^2|n\rangle|^2 |\langle n|R^2|-j\rangle|^2; \\
C_{32} &= \sum_{n,m=-j}^{j} q_n q_m |\langle m|R|n\rangle|^2 |\langle n|R^3|-j\rangle|^2.
\end{aligned}
\tag{5.51}
$$

The matrix R^2 has matrix elements such that $R^2|-j\rangle = (-i)^{2j}|+j\rangle$ and $R^2|+j\rangle = (-i)^{2j}|-j\rangle$. Thus, we obtain

$$
\begin{aligned}
C_{21} &= \sum_{m=-j}^{j} q_m |\langle m|R|j\rangle|^2; \qquad C_{31} = -1 \\
C_{32} &= \sum_{n,m=-j}^{j} q_n q_m |\langle m|R|n\rangle|^2 |\langle n|R^3|-j\rangle|^2.
\end{aligned}
\tag{5.52}
$$

Using the explicit representation of measurement assigments, $q_m = 1 - 2\delta_{m,-j}$, we can write

$$
\begin{aligned}
C_{21} &= \left(\sum_{m=-j}^{j} |\langle m|R|j\rangle|^2 \right) - 2|\langle -j|R|j\rangle|^2 \\
&= 1 - 2|\langle -j|R|j\rangle|^2.
\end{aligned}
\tag{5.53}
$$

The relevant matrix elements are

$$
|\langle n|R|-j\rangle| = \frac{1}{2^j}\sqrt{\binom{2j}{n+j}},
\tag{5.54}
$$

such that

$$
C_{21} = 1 - 2^{1-2j}.
\tag{5.55}
$$

The final term can evaluated as

$$
\begin{aligned}
C_{32} &= 1 - 2|\langle -j|R^3|-j\rangle|^2 \\
&\quad +4|\langle j|R|-j\rangle|^2 |\langle -j|R^3|-j\rangle|^2 \\
&\quad -2\sum_n |\langle -j|R|n\rangle|^2 |\langle n|R^3|-j\rangle|^2 \\
&= 1 - 2\frac{1}{2^{2j}} + 4\frac{1}{2^{4j}} - 2\frac{(4j)!}{4^{2j}[(2j)!]^2}.
\end{aligned}
\tag{5.56}
$$

We have therefore

$$K_3 = 3 - 4^{1-j} + 4^{1-2j} - \frac{2^{1-4j}(4j)!}{[(2j)!]^2}. \tag{5.57}$$

For large j, the latter term can be approximated as $-\sqrt{2/\pi j}$ which then dominates the j-dependence. In the large-spin limit, we have therefore

$$K_3 \sim 3 - \sqrt{\frac{2}{\pi j}}, \tag{5.58}$$

which obviously reaches the value 3 in the $j \to \infty$ limit.

5.4 Discussion

In the first part of the chapter, we have shown that the two main noncontextuality inequalities - the KCBS inequality (Observation 1) and the Peres-Mermin inequality (Observation 3 and 4) - can be used as dimension witnesses. In particular, Observation 4 shows that the the Peres-Mermin inequality can be used to certify the dimension of a Hilbert space independently of the state preparation and in a noise robust way. Our methods allow the application of other inequalities, showing that contextuality can be used as a resource for dimension tests of quantum systems. The resulting tests are state-independent, in contrast to the existing tests. This can be advantageous in experimental implementations, moreover it shows that one can bound the dimension of quantum systems without using the properties of the quantum state.

In the second part of the chapter, we applied a similar analysis to the Leggett-Garg inequality. First, we showed that higher violations of the Leggett-Garg inequality are possible within the framework of standard quantum theory plus projective measurements. Independently of its application as dimension witness, this is of fundamental importance since classical theories reproducing, or exceeding, the quantum correlations for temporal scenarios are conceivable and they do not violate any physical principle, as opposed to Bell scenarios where such classical theories involve faster-than-light communication between space-like separated experiments. In fact, in a temporal scenario a classical device with memory, keeping track of the performed measurements and outcomes, can easily saturate the algebraic bound. However, such a device cannot be considered in Leggett-Garg tests since it contradicts the hypothesis of non-invasiveness of the measurement: The memory must be stored on a (possibly auxiliary) physical system, in such a way that the subsequent dynamics is evidently modified. The same argument applies also to the quantum mechanical description of such a device, which is only possible with POVMs [23]. Such measurement schemes are, therefore, not meaningful in a Leggett-Garg test.

From an information-theoretic perspective, it is interesting to relate temporal correlations to the amount of information transmitted through sequential measurements [23]. While classical devices with memory, and their quantum counterparts based on POVMs, can easily saturate the algebraic bound $K_3 = 3$, the amount of information transmitted trough sequential projective measurements, subjected to Lüders rule, has been proven to obey stricter bounds, independent of the system size. Our analysis shows that degeneracy-breaking projective measurements, as those in von Neumann's scheme, are able to transmit more information, which is encoded in the different evolution paths in the set of quantum state, and can give rise to perfect correlations (or anticorrelations) in the limit of an infinite number of projectors. This is in stark contrast with Bell inequalities, which do not show any higher violation when tested with more general type of quantum measurements and are typically saturated only in the framework of post-quantum theories [37].

Again, here we propose the application of our results a a dimension witness [128]: an experimenter can certify that she is able to manipulate at least M levels of a quantum system, if she can violate the bound for $M-1$. Obviously, also the condition of projective measurement must be verified.

In contrast to the analysis of Sect. 5.1, and also the other proposal of dimension witnesses based on Bell inequalities [128] and the prepare-and-measure scenario [131], where specific inequalities violated only by high-dimensional systems and involving more complex measurement schemes must be found, here we need only the simplest Leggett-Garg inequality.

We also recall that, a further interesting application of the result of this section is the discrimination between Lüders' and von Neumann's state-update rules [148], i.e., which one, if any, correctly represents the measurement scenario. A violation of the bound corresponding to $M=2$ shows a contradiction with Lüders rule. Intermediate cases are possible and can also be investigated with our method.

Conclusions

In this thesis we discussed the characterization of sets of classical and quantum probabilities for different measurement scenarios and different, physically motivated, hidden variable models. More precisely, we discussed local [3], noncontextual [5] and macrorealist [8] hidden variable models and the characterization of their corresponding sets of allowed probabilities, whereas in the quantum case we focused on the characterization of probabilities arising from sequences of projective measurements.

In the classical case, the possible values for probabilities form a convex polytope whose vertices are given by the deterministic assignments, i.e., the $\{0, 1\}$-valued probability measures. Notwithstanding the existence of algorithms for completely characterizing a convex polytope starting from its vertices, the time required for such a computation grows exponentially in the number of settings of the measurement scenario and it can be directly performed only in the simplest cases.

In Chapt. 2, we developed an alternative method for the characterization of polytopes arising in the analysis of hidden variable models based on some results on the extension of probability measures. We then applied our method to several Bell and noncontextuality scenarios providing both computational and analytical results and showing the advantages of our method with respect to the existing ones.

In Chapt. 3, we analyzed the measurement scenarios where quantum correlations are stronger than noncontextual ones for every quantum state (state-independent contextuality). Given the complexity of such scenarios, a direct computation of the corresponding noncontextual polytope is not feasible. We exploited the fact that for such scenarios the measurement settings are usually known (e.g., they come from a proof of the Kochen-Specker theorem) and developed a method for computing optimal inequalities, in the sense of the maximal gap between classical and quantum predictions. Our method is based on linear programming, which allows optimal inequalities to be efficiently computed. Moreover, we applied our method to the most fundamental noncontextuality scenarios and showed that, for all the examples considered, our optimal inequalities are also facets inequalities of the corresponding noncontextual polytope.

In Chapt. 4, we considered quantum probabilities arising in the temporal scenario. Namely, the possible strength of correlations among sequence of projective measurements performed on the same system. Such a scenario is relevant for the test of noncontextuality as well as Leggett-Garg inequalities. We developed a method that provides a complete characterization of the maximal correlations allowed by quantum mechanics in the sequential measurement scenario. Our method is based on semidefinite programming and, thus, maximal correlations can be efficiently computed. We applied our method to the most fundamental Leggett-Garg and noncontextuality scenarios.

In Chapt. 5, we discussed possible application of the previous results as dimension witness, namely, as a certification of the minimal dimension of the Hilbert space needed to explain the arising of certain quantum correlations. As opposed to previous approaches based on Bell inequalities (spatial scenario), or prepare-and-measure scenario, we focused on the sequential measurement scenarios and tests of noncontextuality and Leggett-Garg inequalities. We analyzed the most fundamental noncontextuality inequalities, both with state-dependent and state-independent quantum violations, and show how they can be used to discriminate between different dimensions of the Hilbert space. Most notably, noncontextuality inequalities allows for the implementation of dimension witnesses that do not require the preparation of any specific state. We also discussed the robustness of our dimension witnesses against experimental imperfections. We then applied a similar analysis to the Leggett-Garg inequality, but considering a more general measurement scheme that may be interpreted as a coarse-graining of the measurement outcomes. Our analysis, not only provides new dimension witnesses, but also clarifies the role of the dimension in the temporal scenario. In fact, in the spatial scenario, it is known that quantum correlations, e.g., those appearing in a Bell inequality, obey a fundamental bound known as Tsirelson bound [74], which is independent of the dimension of the system. On the other hand, we proved that temporal correlations obey similar bounds, but such bounds strongly depend on the dimension D of the system, and they can reach the maximal algebraic value in the limit $D \to \infty$.

We believe that our results on the existence of fundamental bounds for temporal correlations closely related to the Tsirelson bound for spatial correlations, but with fundamental differences (e.g., the dependence on the dimension), open the possibility for the investigation of physical or information-theoretic principles explaining their existence. The principles proposed to explain the existence of quantum bounds for Bell and noncontextuality inequalities within the framework of general propobabilistic theories are certainly an interesting starting point (cf. [41, 43, 106, 107]). We leave this question for further research. Moreover, we hope that our results will be a catalyst for the experimental tests on high-dimensional systems, both for the our dimension witnesses and for the tests of macrorealist versus quantum theory, i.e., the Leggett-Garg inequality. In fact, so far all tests of the Leggett-Garg inequality, even when performed on highly-dimensional systems, have been designed and performed to reach the quantum bound that we now know to be valid only for two-dimensional systems.

Bibliography

[1] A. Peres, *Quantum Theory: Concepts and Methods*, (Kluwer, Dordrecht, 1993).

[2] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[3] J. S. Bell, Physics **1**, 195 (1964).

[4] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).

[5] S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).

[6] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).

[7] I. Pitowsky, J. Math. Phys. **27**, 1556 (1986).

[8] A. J. Leggett and A. Garg, Phys. Rev. Lett. **54**, 857 (1985).

[9] P. W. Shor, SIAM J. Comput., **26** (5), 148 (1997).

[10] S. Loyd, Science **273**, 1073 (1996).

[11] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007).

[12] R. Raussendorf, Phys. Rev. A **88**, 022322 (2013).

[13] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Nature (London) **510**, 351 (2014).

[14] I. Pitowsky, *Quantum Probability – Quantum Logic*, Lecture Notes in Physics (Springer-Verlag, Berlin, 1989).

[15] A. N. Kolmogorov, *Foundations of the Theory of Probability*, (Chelsea Publishing Company, New York, 1950).

[16] M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Vol. I, (Academic Press, New York,1972).

[17] L. de Broglie, *An Introduction to the Study of Wave Mechanics*, (E. P. Dutton and Company, Inc. , New York, 1930).

[18] D. Bohm, Phys. Rev. **85**, 166,180 (1952).

[19] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932).

[20] J. M. Jauch and C. Piron, Helv. Phys. Acta, **36**, 827 (1963).

[21] A. M. Gleason, J. Math. Mech. **6**, 885 (1957).

[22] R. W. Spekkens, arXiv:1312.3667 (2013).

[23] T. Fritz, New J. Phys. **12**, 083055 (2010).

[24] C. Budroni, T. Moroder, M. Kleinmann, and O. Gühne, Phys. Rev. Lett. **111**, 020403 (2013).

[25] C. Budroni and C. Emary, Phys. Rev. Lett. **113**, 050401 (2014).

[26] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[27] S. Freedman, and J. Clauser, Phys. Rev. Lett. **28**, 938 (1972).

[28] E. Fry and R. Thompson, Phys. Rev. Lett. **37**, 465 (1976).

[29] A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. 47, 460 (1981).

[30] A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. 49, 91 (1982).

[31] A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. 49 , 1804 (1982).

[32] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland, Nature (London) **409**, 791 (2001).

[33] M. Giustina *et al.*, Nature (London) 497, 227 (2014).

[34] A. A. Klyachko, M. A. Can, S. Binicioğlu, and A. S. Shumovsky, Phys. Rev. Lett. **101**, 020403 (2008).

[35] G. Birkhoff and J. von Neumann, Ann. Math. **37**, (1936).

[36] C. Budroni and G. Morchio, J. Math. Phys. **51**, 122205 (2010).

[37] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).

[38] P. Kurzyński, R. Ramanathan, and D. Kaszlikowski, Phys. Rev. Lett. **109**, 020404 (2012)

[39] E. P. Specker, Dialectica **14**, 239 (1960).

[40] A. Cabello, arXiv:1212.1756 (2012).

[41] A. Cabello, Phys. Rev. Lett. **110**, 060402 (2013).

[42] B. Yan, Phys. Rev. Lett. **110**, 260406 (2013)

[43] T. Fritz, A. B. Sainz, R. Augusiak, J. Bohr Brask, R. Chaves, A. Leverrier, A. Acín, Nature Comm. **4**, 2263 (2013).

[44] J. Henson, arXiv:1210.5978 (2012).

[45] A. Peres, J. Phys. A **24** 175 (1991).

[46] M. Kernaghan, J. Phys. A **27**, 829 (1994).

[47] A. Cabello , J. M. Estebaranz, G. García-Alcaine, Phys. Lett. A **212**, 183 (1996).

[48] A. Peres, Phys. Lett. A **151**, 107 (1990).

[49] N. D. Mermin, Phys. Rev. Lett. **65**, 3373 (1990).

[50] S. Givant and P. Halmos, *Introduction to Boolean Algebras* (Springer, New York, 2009).

[51] A. Cabello, Phys. Rev. Lett. **101**, 210401 (2008).

[52] P. Badziąg, I. Bengtsson, A. Cabello, and I. Pitowsky, Phys. Rev. Lett. **103** 050401 (2009).

[53] G. Kirchmair, F. Zähringer, R. Gerritsma, M. Kleinmann, O. Gühne, A. Cabello, R. Blatt, and C.F. Roos, Nature (London) **460**, 494 (2009).

[54] E. Amselem, M. Rådmark, M. Bourennane, and A. Cabello, Phys. Rev. Lett. **103**, 160405 (2009).

[55] O. Moussa, C.A. Ryan, D.G. Cory, and R. Laflamme, Phys. Rev. Lett. **104**, 160501 (2010).

[56] S. Yu and C.H. Oh, Phys. Rev. Lett. **108**, 030402 (2012).

[57] G. Lüders, Ann. Phys. (Leipzig) **8**, 322 (1951).

[58] G. Lüders (English translation and discussion by K. A. Kirkpatrick), Ann. Phys. **15** , 663 (2006).

[59] A. Palacios-Laloy, F. Mallet, F. Nguyen, P. Bertet, D. Vion, D. Esteve, and A. N. Korotkov, Nature Phys. **6**, 442 (2010).

[60] M. E. Goggin, M. P. Almeida, M. Barbieri, B. P. Lanyon, J. L. OBrien, A. G. White, and G. J. Pryde, Proc. Natl. Acad. Sci. USA **108**, 1256 (2011).

[61] J.-S. Xu, C.-F. Li, X.-B. Zou, and G.-C. Guo, Scientific Reports **1**, 101 (2011).

[62] J. Dressel, C. J. Broadbent, J. C. Howell, and A. N. Jordan, Phys. Rev. Lett. **106**, 040402 (2011).

[63] Y. Suzuki, M. Iinuma, and H. F. Hofmann, New J. Phys. **14**, 103022 (2012).

[64] G. Waldherr, P. Neumann, S. F. Huelga, F. Jelezko, and J. Wrachtrup, Phys. Rev. Lett. **107**, 090401 (2011).

[65] R. E. George, L. M. Robledo, O. J. E. Maroney, M. S. Blok, H. Bernien, M. L. Markham, D. J. Twitchen, J. J. L. Morton, G. A. D. Briggs, and R. Hanson, Proc. Natl. Acad. Sci. USA **110**, 3777 (2013).

[66] V. Athalye, S. S. Roy, and T. S. Mahesh, Phys. Rev. Lett. **107**, 130402 (2011).

[67] A. M. Souza, I. S. Oliveira, and R. S. Sarthour, New J. Phys. **13**, 053023 (2011).

[68] M. M. Deza and M. Laurent, *Geometry of Cuts and Metrics. Volume 15 of Algorithms and Combinatorics* (Springer, Berlin 1997).

[69] J. F. Clauser and M. A. Horne, Phys. Rev. D **10**, 526 (1974).

[70] `cdd`, `http://www.ifor.math.ethz.ch/~fukuda/cdd_home/cdd.html`

[71] `lrs`, `http://cgm.cs.mcgill.ca/~avis/C/lrs.html`

[72] `porta`, `http://www.zib.de/Optimization/Software/porta/`

[73] I. Pitowsky, Math. Programming **50**, 395 (1991).

[74] B. C. Cirel'son, Lett. Math. Phys. **4**, 93 (1980).

[75] L. J. Landau, Phys. Lett. A **120**, 54 (1987.

[76] S. Wehner, Phys. Rev. A **73**, 022110 (2006).

[77] L. Vandenberghe and S. Boyd, SIAM Rev. **38**, 49 (1996).

[78] D. Avis, H. Imai, and T. Ito, J. Phys. A: Math. Gen. **39** 11283 (2006).

[79] M. Navascués, S. Pironio, and A. Acín, Phys. Rev. Lett. **98** (2007).

[80] M. Navascués, S. Pironio, and A. Acín, New J. Phys. **10** (2008).

[81] A Cabello, S Severini, and A Winter, arXiv:1010.2163 (2010).

[82] A Cabello, S Severini, and A Winter, Phys. Rev. Lett. **112**, 040401 (2014).

[83] L. Lovász, IEEE Trans. Inf. Theory **25**, 1 (1979).

[84] L. Lovász, *Geometric Representations of Graphs*, `www.cs.elte.hu/lovasz/geomrep.pdf` (unpublished).

[85] S. Boyd and L. Vandenberghe, *Convex Optimization*, (Cambridge University Press 2004).

[86] C. Budroni and A. Cabello, J. Phys. A: Math. Theor. **45**, 385304 (2012).

[87] M. Araújo, M. T. Quintino, C. Budroni, M. Terra Cunha, and A. Cabello, Phys. Rev. A **88**, 022118 (2013).

[88] C. Śliwa Phys. Lett. A **317** 165 (2003).

[89] D. Collins and N. Gisin J. Phys. A: Math. Gen. **37** 1775 (2004).

[90] G. M. Ziegler, *Lectures on Polytopes* (Springer-Verlag, New York 1995)

[91] C. Budroni and G. Morchio, Found. Phys. **42**, 544 (2012).

[92] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, Phys. Rep. **506**, 1 (2011).

[93] S. L. Braunstein and C. M. Caves, Nucl. Phys. B, Proc. Suppl. **6**, 211 (1989).

[94] M. Kleinmann, C. Budroni, J.-Å. Larsson, O. Gühne, and A. Cabello, Phys. Rev. Lett. **109**, 250402 (2012).

[95] E. Amselem, M. Bourennane, C. Budroni, A. Cabello, O. Gühne, M. Kleinmann, J.-Å. Larsson, and M. Wieśniak, Phys. Rev. Lett. **110**, 078901 (2013).

[96] B. Jungnitsch, S. Niekamp, M. Kleinmann, O. Gühne, H. Lu, W.-B. Gao, Y.-A. Chen, Z.-B. Chen, and J.-W. Pan, Phys. Rev. Lett. **104**, 210401 (2010).

[97] `CVXOPT`, `http://cvxopt.org/`

[98] A. Cabello, arXiv:1112.5149 (2011).

[99] R. Ramanathan and P. Horodecki, Phys. Rev. Lett. 112, 040404 (2014).

[100] A. Cabello, Phys. Rev. A **82**, 032110 (2010).

[101] S. Yu and C. H. Oh, arXiv:1112.5513 (2011).

[102] O. Gühne, M. Kleinmann, A. Cabello, J.-Å Larsson, G. Kirchmair, F. Zähringer, R. Gerritsma, and C. F. Roos, Phys. Rev. A **81**, 022121 (2010).

[103] C. Zu, Y.-X. Wang, D.-L. Deng, X.-Y. Chang, K. Liu, P.-Y. Hou, H.-X. Yang, and L.-M. Duan, Phys. Rev. Lett. 109, 150401 (2012).

[104] X. Zhang, M. Um, J. Zhang, S. An, Y. Wang, D.-L. Deng, C. Shen, L.-M. Duan, and K. Kim, Phys. Rev. Lett. 110, 070401 (2013).

[105] X. Zhang, M. Um, J. Zhang, S. An, Y. Wang, D.-L. Deng, C. Shen, L.-M. Duan, and K. Kim, arXiv:1209.3831 (2012).

[106] G. Brassard, H. Buhrman, N. Linden, A. A. Méthot, A. Tapp, and F. Unger, Phys. Rev. Lett. **96**, 250401 (2006).

[107] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Nature (London) **461**, 1101 (2009).

[108] D. Gross ,M. Müller, R. Colbeck, and C. O. Dahlsten, Phys. Rev. Lett. **104** 080402 (2010).

[109] M. Navascués, Y. Guryanova, M. J. Hoban, and A. Acín arXiv:1403.4621 (2014).

[110] M. Barbieri, Phys. Rev. A **80**, 034102 (2009).

[111] J. Kofler and C. Brukner, Phys. Rev. Lett. **101**, 090403 (2008).

[112] T. Fritz, J. Math. Phys. **51**, 052103 (2010).

[113] A.C. Doherty, Y.-C. Liang, B. Toner, and S. Wehner, in *Proceedings of IEEE Conference on Computational Complexity, College Park, MD, 2008* (IEEE, New York 2008) p.199.

[114] S. Pironio, M. Navascués, and A. Acín, SIAM J. Optim. **20(5)**, 2157 (2010).

[115] O. Gühne, C. Budroni, A. Cabello, M. Kleinmann, J.-Å. Larsson, Phys. Rev. A **89**, 062107 (2014).

[116] K. F. Pál and T. Vértesi, Phys. Rev. A **82**, 022116 (2010).

[117] T. Moroder, J.-D. Bancal, Y.-C. Liang, M. Hofmann, and O. Gühne, Phys. Rev. Lett. **111**, 030501 (2013).

[118] M. L. Almeida, J.-D Bancal, N. Brunner, A. Acín, N. Gisin, and S. Pironio, Phys. Rev. Lett. **104**, 230404 (2010).

[119] R. M. Gray, Foundations and Trends in Communications and Information Theory, **2** 155 (2006).

[120] B. S. Tsirel'son, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **142**, 174 (1985), (Russian, English translation [121]).

[121] B. S. Tsirel'son, J. Soviet Math. **36**, 557 (1987).

[122] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics Vol. 2*, (Springer 1997).

[123] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).

[124] O. Gühne and G. Tóth, Phys. Rep. **474**, 1 (2009).

[125] G. Molina-Terriza, J.P. Torres, and L. Torner, Nature Physics **3**, 305 (2007).

[126] A. C. Dada, J. Leach, G. S. Buller, M. J. Padgett, and E. Andersson, Nature Physics **7**, 677 (2011).

[127] P. Neumann, N. Mizuochi, F. Rempp, P. Hemmer, H. Watanabe, S. Yamasaki, V. Jacques, T. Gaebel, F. Jelezko, and J. Wrachtrup, Science **320**, 1326 (2008).

[128] N. Brunner, S. Pironio, A. Acín, N. Gisin, A. Méthot, and V. Scarani, Phys. Rev. Lett. **100**, 210503 (2008).

[129] T. Vértesi and K. F. Pál, Phys. Rev. A **79**, 042106 (2009).

[130] S. Wehner, M. Christandl, and A.C. Doherty, Phys. Rev. A **78**, 062112 (2008).

[131] R. Gallego, N. Brunner, C. Hadley, and A. Acín, Phys. Rev. Lett. **105**, 230501 (2010).

[132] N. Brunner, M. Navascus, and T. Vrtesi, Phys. Rev. Lett. **110**, 150501 (2013).

[133] M. Hendrych, R. Gallego, M. Mičuda, N. Brunner, A. Acín, and J. P. Torres, Nature Phys. **8**, 588 (2012).

[134] J. Ahrens, P. Badziąg, A. Cabello, and M. Bourennane, Nature Phys. **8**, 592 (2012).

[135] M. M. Wolf and D. Pérez-García, Phys. Rev. Lett. **102**, 190504 (2009).

[136] R. Łapkiewicz, P. Li, C. Schaeff, N.K. Langford, S. Ramelow, M. Wieśniak, and A. Zeilinger, Nature (London) **474**, 490 (2011).

[137] J. Ahrens, E. Amselem, A. Cabello, and M. Bourennane, Sci. Rep. **3**, 2170 (2013).

[138] J. Kofler and Č. Brukner, Phys. Rev. Lett. **99**, 180403 (2007).

[139] N. Lambert, R. Johansson, and F. Nori, Phys. Rev. B **84**, 245421 (2011).

[140] M. M. Wilde, J. M. McCracken, and A. Mizel, Proc. R. 5 Soc. A **466**, 1347 (2010).

[141] B. Dakić, T. Paterek, and Č. Brukner, New J. Phys. **16** 023028 (2014).

[142] M. Markiewicz, P. Kurzynski, J. Thompson, S.-Y. Lee, A. Soeda, T. Paterek, and D. Kaszlikowski, Phys. Rev. A **89**, 042109 (2014).

[143] S. Das, S. Aravinda, R. Srikanth, and D. Home, Europhys. Lett. **104**, 60006 (2013).

[144] Č. Brukner, S. Taylor, S. Cheung, and V. Vedral, quant- ph/0402127 (2004).

[145] S. Marcovitch and B. Reznik, arXiv:1107.2186 (2011).

[146] O. Oreshkov, F. Costa, and Č. Brukner, Nat. Commun. **3**, 1092 (2012).

[147] J. Fitzsimons, J. Jones, and V. Vedral, arXiv:1302.2731 (2013).

[148] G. C. Hegerfeldt and R. Sala Mayato, Phys. Rev. A **85**, 032116 (2012).

List of publications

[1] C. Budroni and G. Morchio, *The extension problem for partial Boolean structures in quantum mechanics*, J. Math. Phys. **51**, 122205 (2010).

[2] C. Budroni and G. Morchio, *Bell inequalities as constraints on unmeasurable correlations*, Found. Phys. **42**, 544 (2012).

[3] C. Budroni and A. Cabello, *Bell inequalities from variable elimination methods*, J. Phys. A: Math. Theor. **45**, 385304 (2012).

[4] M. Kleinmann, C. Budroni, J.-Å. Larsson, O. Gühne, and A. Cabello, *Optimal inequalities for state-independent contextuality*, Phys. Rev. Lett. **109**, 250402 (2012).

[5] J.-Å. Larsson, M. Kleinmann, C. Budroni, O. Gühne, and A. Cabello, *Maximal violation of state-independent contextuality inequalities*, AIP Conf. Proc. **1508**, 265 (2012).

[6] E. Amselem, M. Bourennane, C. Budroni, A. Cabello, O. Gühne, M. Kleinmann, J.-Å. Larsson, and M. Wieśniak, *Comment on "State-Independent Experimental Test of Quantum Contextuality in an Indivisible System"*, Phys. Rev. Lett. **110**, 078901 (2013).

[7] C. Budroni, T. Moroder, M. Kleinmann, and O. Gühne, *Bounding temporal quantum correlations*, Phys. Rev. Lett. **111**, 020403 (2013).

[8] M. Araújo, M. T. Quintino, C. Budroni, M. Terra Cunha, and A. Cabello, *All noncontextuality inequalities for the n-cycle scenario*, Phys. Rev. A **88**, 022118 (2013).

[9] O. Gühne, C. Budroni, A. Cabello, M. Kleinmann, J.-Å. Larsson, *Bounding the quantum dimension with contextuality*, Phys. Rev. A **89**, 062107 (2014).

[10] C. Budroni and C. Emary, *Temporal quantum correlations and Leggett-Garg inequalities in multi-level systems*, Phys. Rev. Lett. **113**, 050401 (2014).

Publications [1] and [2] were the result of an elaboration of the work and ideas of my master thesis. Publications [3]-[10] were written in the course of this thesis.

Chapter 2 is based on publications [3] and [8]. Chapter 3 is based on publications [4], [5], and [6]. Chapter 4 is based on publications [7] and [8]. Chapter 5 is based on publications [9] and [10].