Università del Salento
Dipartimento di Matematica e Fisica
"Ennio de Giorgi"
Corso di Laurea Magistrale in Fisica

Master's Thesis

Exploring the boundaries of quantum correlations: Bell inequalities, convex duality and self-testing

Supervisor:
Prof. Luigi Martina

Co-Supervisor:
Prof. Jean-Daniel Bancal

Student:
Andrea Zingarofalo

Academic Year 2024/2025

My advice to young people entering science:
you should do it for the love of science...
You should enter because you are fascinated by it.
(J. Peebles)

Sommario

Questo lavoro di tesi esplora la geometria delle correlazioni quantistiche nello scenario CHSH, con particolare riguardo alla struttura estremale dell'insieme delle correlazioni quantistiche. Utilizzando strumenti propri della geometria convessa e tecniche di dualità, si studia come i punti quantistici estremali caratterizzano la frontiera dell'insieme e le correlazioni ottenibili da sistemi quantistici entangled.
La prima parte della tesi fornisce un'introduzione autocontenuta alla nonlocalità secondo Bell, comprendente una classificazione delle correlazioni, la geometria degli insiemi delle correlazioni locali, quantistiche e no-signaling, e il ruolo dei punti estremali nei protocolli device-independent e in questioni fondazionali correlate.
Successivamente, si presenta l'approccio duale alla geometria delle correlazioni quantistiche, passando in rassegna tecniche dell'analisi convessa e applicandole alla caratterizzazione analitica e numerica della faccia ortogonale associata al punto di Tsirelson.
Il contributo originale della tesi consiste nell'estendere questa analisi a una classe più ampia di punti quantistici estremali. Fissando le operazioni di misura effettuate sul sistema, e al contempo variando il parametro di entanglement, si individua una famiglia di punti quantistici estremali e se ne caratterizzano gli iperpiani di supporto ad essi associati.
I risultati ottenuti offrono nuove prospettive sulla struttura della frontiera quantistica e contribuiscono alla comprensione teorica della nonlocalità quantistica, sia dal punto di vista fondazionale che in ambito applicativo.

Abstract

This dissertation explores the geometry of quantum correlations in the CHSH scenario, with a focus on the extremal structure of the quantum set. Using tools from convex geometry and duality theory, we study how extremal quantum behaviors shape the boundary of the set and determine the correlations that can arise from entangled quantum systems.
The first part of the thesis provides a self-contained introduction to Bell nonlocality, including the classification of correlations, the geometry of the local, quantum, and no-signaling sets, the role of extremality in device-independent protocols and related foundational questions.
We then present the dual geometric approach to quantum correlations, reviewing techniques from convex analysis and applying them to characterize the orthogonal face of the Tsirelson point, both analytically and numerically.
The original contribution of the thesis consists in extending this analysis to a broader class of extremal quantum points. By fixing measurement settings and varying the entanglement parameter, we identify a family of extremal behaviors and characterize their supporting hyperplanes.
Our results offer new insights into the structure of the quantum boundary and contribute to the theoretical understanding of quantum nonlocality in both foundational and applied contexts.

Table of Contents

List of Symbols

Symbol	Description
Sets and Spaces	
$\mathbb{R}, \mathbb{C}$	Sets of real and complex numbers
$\mathcal{H}$	Hilbert space
$L(\mathcal{H})$	Set of bounded linear operators acting on $\mathcal{H}$
$\mathcal{P}$	Probability space
$\mathcal{K}$	Generic convex set
$\mathcal{L}$	Local set
$\mathcal{Q}$	Quantum set
$\mathcal{NS}$	No-signaling set
Operators and Measurements	
$\mathbb{1}$	Identity operator
$\hat{\rho}$	Density operator
$\hat{A}_x, \hat{B}_y$	Unitary measurement operators (CHSH scenario)
$\hat{M}^{(j)}_{a_j\|x_j}$	POVM element for party j, setting x_j, outcome a_j
$\hat{\Pi}^{(j)}_{a_j\|x_j}$	Projective measurement for party j, setting x_j, outcome a_j
$\hat{S}$	Bell operator
σ_i	Pauli operators
Bell Scenario Parameters	
n	Number of parties (generalized Bell scenario)
m	Number of measurements per party
k	Number of outcomes per measurement
x, y	Inputs (measurement settings)
a, b	Outputs (measurement outcomes)
λ	Local hidden variable
Correlations and Behaviors	
$P(a, b\|x, y)$	Joint probabilities (correlations)
$\boldsymbol{P}$	Correlation behavior in vector representation

Symbol	Description
Γ	Moment matrix
W	Dual of the moment matrix
Bell Expressions and Bounds	
β	Bell expression in polynomial representation
$\boldsymbol{\beta}$	Bell expression as coefficient vector
$\beta(\boldsymbol{P})$	Bell expression in linear functional representation
$\beta_{\mathcal{L}}$	Local bound of a Bell expression
$\beta_{\mathcal{Q}}$	Quantum (Tsirelson) bound

Preface

It is striking that it has so far not been possible to find a logically consistent theory that is close to quantum mechanics, other than quantum mechanics itself.

Steven Weinberg
Dreams of a Final Theory

Over the past century, quantum mechanics has established itself as one of the most successful and far-reaching theories in the history of science. Its predictive power has been confirmed with astonishing precision across a wide range of physical phenomena, from the behavior of subatomic particles to the properties of condensed matter systems. Additionally, new models widening the understanding of our surrounding were constructed from quantum physics. The Standard Model of particle physics, within the framework of quantum field theory, represents a milestone achievement for the natural sciences, it is the most complete account of the physical world that we have so far. From a pragmatic point of view, quantum theory (QT) led to the so-called *first quantum revolution*, where devices exploiting the laws of quantum mechanical systems were invented, like the transistor and the laser, which consequently led to further developments and applications in areas of science like chemistry, biology, medicine, engineering and computing, with a huge impact in the economy and culture of our society, and virtually every aspect of contemporary human life.

Beyond its empirical success, QT has fundamentally reshaped our understanding of the physical world, challenging classical notions of *determinism*, *locality*, and *realism*. The counterintuitive features of quantum mechanics—superposition, entanglement, and measurement-induced collapse—have stimulated profound conceptual debates and motivated some of the most important foundational questions in physics [1–3].

Among these, the nature of quantum correlations stands out as a central theme. In 1935, The seminal paper by Einstein, Podolsky, and Rosen (EPR) [4] famously questioned the completeness of quantum mechanics, highlighting a paradox that arises when considering entangled states shared between spatially separated systems. The EPR argument [5–8] rested on the assumption of *local realism*: the idea that physical properties exist prior to and independent of measurement, and that no influence can travel faster than light. If one accepts these assumptions, QT appears

to allow for instantaneous correlations between distant events, which Einstein famously referred to as "spooky action at a distance."

This tension was brought into sharp focus by John Bell in 1964 [9]. Bell formulated a precise mathematical inequality—now known as Bell's inequality—that must be satisfied by any theory obeying local realism. He then showed that quantum mechanics predicts violations of this inequality in certain entangled systems. Bell's theorem thus marks a fundamental turning point in our understanding of nature: it proves that no local hidden variable theory can reproduce all the predictions of quantum mechanics. These predictions have since been confirmed experimentally [10–13], with increasing precision and sophistication [14, 15], leaving little doubt that quantum theory exhibits an intrinsic form of *nonlocality*. [1]

The discovery of quantum nonlocality prompted a renewed interest in the structure of correlations allowed by different physical theories. In particular, researchers began to explore the space of possible correlations among distant parties performing local measurements [19, 20]. From this perspective, each class of physical theories—classical, quantum, and no-signaling—defines a distinct set of achievable correlations. The classical (or local) set forms a polytope characterized by deterministic strategies and convex combinations thereof. The no-signaling set, which includes all correlations consistent with relativistic causality, forms a larger polytope encompassing both classical and quantum correlations. The set of quantum correlations, however, occupies a rich intermediate position: it is convex but not a polytope, and exhibits a highly nontrivial boundary structure.

A powerful way to study quantum correlations is through their *geometric* representation in probability space. This approach was pioneered by B. S. Tsirelson in the 1980s, who characterized the set of quantum correlations in terms of the Hilbert space formalism of quantum theory. In what is now known as Tsirelson's theorem [21], he provided explicit bounds on the strength of quantum violations of Bell inequalities, showing that these bounds are strictly stronger than classical, but weaker than the no-signaling constraints. Tsirelson's work laid the foundation for a geometric approach to quantum nonlocality, enabling a deeper understanding of the shape and structure of the quantum set $\mathcal{Q}$.

This geometric perspective plays a central role in foundational investigations. One of the major open questions in the foundations of quantum theory is whether quantum mechanics is uniquely singled out by simple physical or information-theoretic principles. In recent decades, several reconstruction programs have attempted to derive the formalism of quantum theory from axioms such as continuity, reversibility, and tomographic locality [22–25]. Other approaches seek to understand quantum correlations by embedding them within a broader landscape of logically consistent theories, asking whether quantum theory is *an island in theoryspace* [26, 27].

[1] Throughout this thesis, we use the term *nonlocality* to refer specifically to the empirical violation of a Bell inequality. This notion is standard in the physics literature, where Bell nonlocality denotes statistical correlations that cannot be explained by any local hidden variable model. It is important to emphasize, however, that Bell's theorem—and the corresponding definition of locality—rests on precise assumptions, including *statistical independence*, *realism*, and a form of *relativistic causality*. The interpretation and physical justification of "locality" involved, remain subjects of ongoing debate. In particular, the extent to which Bell inequality violations conflict with the principles of special relativity depends on one's interpretative framework and assumptions about the ontology of quantum theory. These foundational questions, which have been explored extensively in both the physics and philosophy of science communities since Bell's original work, lie beyond the scope of this thesis. For further discussion, see e.g. [16–18].

Principles like *information causality* [28], *macroscopic locality* [29], and *local orthogonality* [30] have been proposed to distinguish quantum correlations from more general no-signaling ones. In this context, the structure of the set $\mathcal{Q}$ reflects deep physical constraints, and understanding its geometry provides a window into the fundamental properties of quantum theory itself.

In recent decades, this perspective has also proved to be of growing *practical* relevance. The nonlocal correlations predicted by quantum mechanics have become a key resource in quantum information science. *Device-independent* protocols [31, 32] exploit quantum nonlocality to guarantee security or functionality without relying on the internal workings of the devices used. These include device-independent quantum key distribution [33–35], randomness certification [36], and self-testing of quantum states and measurements [37]. Furthermore, the structure of quantum correlations plays a role in quantum computing and complexity theory, where it can lead to computational speed-ups [38, 39], and in the design of quantum networks [40] and the future quantum internet [41–43]. Understanding the geometry of the quantum set is thus not only essential for foundational clarity, but also for technological advancement.

Despite a sustained body of work [44–48], the structure of the quantum set remains mostly elusive. It is known that $\mathcal{Q}$ is convex but not a polytope. Its boundary is curved and often defined only implicitly by the constraints of quantum theory. Many of its extremal points are poorly understood, and explicit characterizations are scarce. While numerical methods such as the Navascués-Pironio-Acín (NPA) hierarchy of semidefinite programs [49, 50] provide powerful tools to approximate the set from the outside, an exact analytic description has long remained out of reach. Recently, a notable progress was obtained with the first complete characterization in terms of extremal points for the quantum set in the so-called CHSH scenario [51].

This thesis is situated within this broad context, and aims to contribute to the geometric understanding of the quantum set. Specifically, we focus on the hyperplanes that support the newly identified *extremal points* in the CHSH scenario, i.e. on their *orthogonal faces*. Building on earlier results concerning the Tsirelson point [52], we generalize the analysis to a family of nonlocal extremal points in the CHSH scenario that arise from states with a varying degrees of entanglement. We develop both analytical and numerical tools to identify the Bell expressions maximized by these points, determine local and quantum bounds, and investigate the role of sum-of-squares decompositions in *certifying* quantum optimality.

Outline of the thesis

Chapter 1 introduces the general framework of Bell nonlocality, including the classification of correlations, the geometry of the local, quantum, and no-signaling sets, and an overview of device-independent protocols and self-testing. We also briefly mention related foundational aspects.

Chapter 2 presents the dual geometric approach to the quantum set. After introducing tools from convex analysis and duality theory, we review the complete characterization of the orthogonal face of the Tsirelson point using both analytical and numerical techniques.

Chapter 3 constitutes the original contribution of this thesis. We generalize the methods of Chapter 2 to a wider class of extremal quantum points in the CHSH scenario. We then study the associated orthogonal faces, identifying the corresponding Bell functionals, and determining both local and quantum bounds through numerical and symbolic analysis.

Taken together, these results contribute to a more refined understanding of the quantum set's geometry and lay the groundwork for future studies of extremal correlations in more general Bell scenarios.

Chapter 1

Bell nonlocality

In this chapter, we explore the so-called device-independent approach to quantum theory, where conclusions about physical systems are drawn directly from observed statistics, without making assumptions about the internal functioning of the devices. We begin by introducing the basic structure of device-independent scenarios, focusing on the Bell and CHSH settings as paradigmatic examples. We then classify the different types of correlations that can arise—local, quantum, and no-signaling—and examine their rich geometric structure. This leads naturally to a discussion of Bell inequalities and their role in distinguishing these sets, including the use of Bell operators and the Navascués-Pironio-Acín hierarchy of semidefinite programs. Building on this framework, we introduce the powerful concept of device-independent self-testing, which allows one to certify quantum states and measurements from correlations alone. Finally, we briefly reflect on the broader foundational implications of this perspective, and what it reveals about the nature of nonlocality, causality, and the structure of quantum theory itself.

1.1 The device-independent paradigm

In many areas of science, it is often fruitful to analyze systems by focusing solely on their observable input-output behavior, abstracting away from their internal mechanisms. This strategy, commonly referred to as the *black-box approach*, is particularly powerful in quantum information theory, where it underpins the *device-independent (DI) paradigm* [31, 32]. In this framework, quantum devices are treated as opaque systems whose internal workings are unknown or untrusted. All that is accessible to the experimenter is the statistical behavior of the devices in response to various inputs.

In this thesis, we adopt this perspective to investigate the implications of quantum theory in bipartite Bell scenarios, where two spatially separated parties—traditionally named Alice and Bob—interact with uncharacterized devices. By analyzing the observable input-output correlations produced in such settings, we can uncover deep insights into the structure of quantum correlations, the limitations of classical models, and the potential for certifying quantum resources in a device-independent manner.

Consider a bipartite experimental setting involving two space-like separated parties, traditionally named Alice and Bob. Each party is located in a distinct laboratory, such that no communication can occur between them during the execution of the experiment. In particular, the choice of measurement setting by one party cannot be signaled to the other. This condition can be enforced by ensuring that the measurement events are space-like separated in accordance with special relativity: if Alice's choice of setting x and Bob's choice y are made after they are causally disconnected, then neither party can influence the other's measurement outcome.

In each run of the experiment, Alice and Bob each choose one of m possible measurement settings, denoted $x, y \in \{1, \ldots, m\}$, and obtain outcomes $a, b \in \{1, \ldots, k\}$, respectively. The experiment is characterized statistically by the joint conditional probability distribution

$$P(a, b|x, y), \tag{1.1}$$

which gives the probability of obtaining outcomes a and b when the measurement settings x and y are chosen.

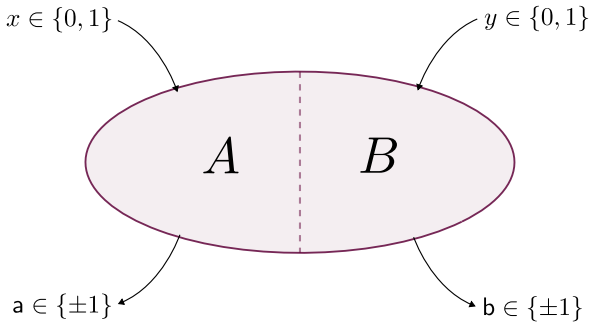

Figure 1.1: Pictorial representation of the CHSH scenario. From [51].

This abstract framework is known as a *Bell scenario*, and it encapsulates the notion of a *black-box model*, in which each party inputs a setting and receives an output. A Bell scenario with outcome alphabet of size k, and m measurement settings per party is specified by k^2m^2 conditional probabilities, one for each possible combination of inputs and outputs.

The collection of all such probabilities defines what is called a *behavior*:

$$\boldsymbol{P} \coloneqq \{P(a, b|x, y)\}_{a,b,x,y}. \tag{1.2}$$

Each behavior $\boldsymbol{P}$ is a point in the space $\mathbb{R}^{k^2m^2}$. However, physical realizability imposes two fundamental constraints:

$$\text{Positivity:} \qquad P(a, b|x, y) \geq 0, \quad \forall a, b, x, y \tag{1.3a}$$

$$\text{Normalization:} \quad \sum_{a,b} P(a, b|x, y) = 1, \quad \forall x, y. \tag{1.3b}$$

These conditions ensure that the behavior $\boldsymbol{P}$ is a valid probability distribution for each choice of settings. The set of all points $\boldsymbol{P}$ satisfying these constraints defines a convex subset of $\mathbb{R}^{k^2m^2}$ known as the *behavior space* $\mathcal{P}$, with

$$\dim \mathcal{P} = (k^2 - 1)m^2. \tag{1.4}$$

A particularly important case is the simplest nontrivial Bell scenario, known as the *CHSH scenario*, where $m = 2$ and $k = 2$. In this case, Alice and Bob each have two possible measurement settings, $x, y \in \{0, 1\}$, and each measurement yields one of two outcomes, conventionally labeled $a, b \in \{\pm 1\}$. This scenario has been extensively studied in both theoretical and experimental contexts, and plays a central role in the investigation of quantum nonlocality.

Remark 1.1. While the discussion above is restricted to the bipartite case ($n = 2$), the framework can be naturally generalized to *multipartite* or *generalized Bell scenarios*, where n spatially separated parties share a global system and independently choose local measurements. The corresponding behavior is then a conditional distribution

$$P(a_1, \ldots, a_n | x_1, \ldots, x_n) \coloneqq P(\boldsymbol{a}, \boldsymbol{x}), \tag{1.5}$$

and the behavior space $\mathcal{P}_n$ is embedded in $\mathbb{R}^{k^n m^n}$, subject to similar positivity and normalization constraints.

1.2 Different types of correlations

The existence of a given physical model behind the correlations obtained in a Bell scenario translates into additional constraints on the behaviors $\boldsymbol{P}$. We consider three main types of correlations:

1. no-signaling correlations
2. local correlations
3. quantum correlations.

1.2.1 No-signaling correlations

A foundational principle shared by both classical and quantum physics is the *no-signaling principle*, which asserts the impossibility of instantaneous information transfer between spatially separated systems:

Principle 1.1 (No-signaling principle). *The transmission of information requires a physical carrier that departs from the sender after the message has been chosen.*

This principle encapsulates the empirical fact that communication necessitates a physical medium—such as acoustic waves, electromagnetic radiation, or material particles—to convey information. Since such carriers propagate through space-time, and special relativity prohibits any influence from propagating faster than the speed of light in vacuum, causal signaling between space-like separated events is forbidden. Consequently, in scenarios involving two distant parties, if their respective inputs are chosen after they become space-like separated, the input selected by one party cannot influence the measurement outcome obtained by the other.

In operational terms, this implies that the marginal probability distribution of one party's outcomes must be independent of the other party's input choice:

$$P(a|x,y) \coloneqq \sum_{b=1}^{k} P(a,b|x,y) \overset{\text{NS}}{=} P(a|x), \qquad \forall a,x,y. \tag{1.6a}$$

$$P(b|x,y) \coloneqq \sum_{a=1}^{k} P(a,b|x,y) \overset{\text{NS}}{=} P(b|y), \qquad \forall b,x,y. \tag{1.6b}$$

Equivalently,

Definition 1.1. A conditional probability distribution $P(a,b|x,y)$ over outcomes $a,b \in \{1,\dots,k\}$ and inputs (measurement settings) $x,y \in \{1,\dots,m\}$ is said to be *no-signaling* if it satisfies the following constraints:

$$\sum_{b=1}^{k} P(a,b|x,y) = \sum_{b=1}^{k} P(a,b|x,y'), \qquad \forall a,x,y,y', \tag{1.7a}$$

$$\sum_{a=1}^{k} P(a,b|x,y) = \sum_{a=1}^{k} P(a,b|x',y), \qquad \forall b,x,x',y. \tag{1.7b}$$

The set of all such distributions is called the *no-signaling set*, denoted by $\mathcal{NS}$.

These constraints guarantee that neither Alice nor Bob can use their choice of measurement setting to influence, and thus communicate with, the other party.

The no-signaling principle plays a pivotal role in quantum information theory and the foundations of quantum mechanics. It serves as a universal constraint on physically realizable correlations and is essential for delineating the boundary between quantum and post-quantum theories. In particular, while quantum correlations always satisfy the no-signaling conditions, the converse is not true: some no-signaling correlations, such as those described by the *Popescu-Rohrlich (PR) box*, exceed the strength of quantum correlations and are not physically realizable within quantum theory [53]. The canonical PR box is defined by the conditional probabilities

$$P_{\text{PR}}(a,b|x,y) \coloneqq \begin{cases} \frac{1}{2} & \text{if } ab = (-1)^{xy}, \\ 0 & \text{otherwise,} \end{cases} \tag{1.8}$$

where $a,b \in \{-1,1\}$, $x,y \in \{0,1\}$.

Thus, the no-signaling principle not only encodes a fundamental causal structure of physical theories, but also imposes stringent operational constraints on any observed statistical correlations.

1.2.2 Local correlations

The concept of *local causality* was introduced by John Bell as a formulation of the relativistic principle that influences cannot travel faster than light. In Bell's words [54]:

> *A theory will be said to be locally causal if the probabilities attached to values of local beables*[1] *in a space-time region 1 are unaltered by specification of values in a space-like separated region 2, when what happens in the backward light cone of 1 is already sufficiently specified* (*op. cit.* pp. 239-240).

This principle expresses the idea that events in one region of spacetime should not be directly influenced by choices or events occurring in a distant, space-like separated region— provided that all relevant past information is already accounted for. More precisely, once a complete specification of all physical influences in the past light cone of region 1 is given, any additional information from region 2 should be redundant in determining the probabilities for events in region 1.

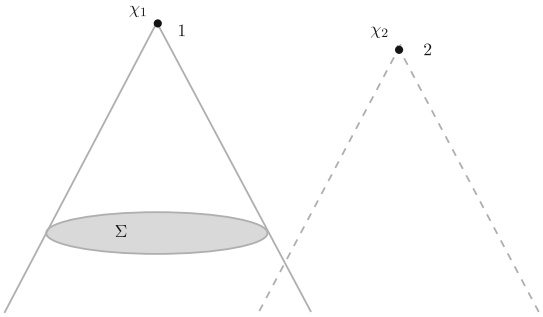

Figure 1.2: According to Bell's notion of local causality, the probability of an event χ_1 in region 1, conditioned on a complete specification C_Σ of events in a spacetime slice Σ within its backward light cone, should be unaffected by additional conditioning on a space-like separated event χ_2 outside the future of Σ. From [7].

In a typical Bell experiment, although Alice and Bob are located at distant, space-like separated regions, they may still share information from a common past. This shared information could consist of pre-established strategies, shared randomness, or any other correlated data originating from events within their mutual backward light cones. To formally account for such pre-existing correlations, one introduces a variable λ, commonly referred to as a *hidden variable*, which encodes all relevant data accessible to both parties before the measurements are chosen.

According to the principle of local causality, measurement outcomes should be determined solely by local causes. That is, the outcome observed by each party depends only on their local measurement setting and on the shared hidden variable λ, and not on the distant party's setting or outcome. This assumption leads to factorize of the conditional on λ joint probability distribution in the form

$$P(a, b|x, y, \lambda) = P(a|x, \lambda)P(b|y, \lambda), \tag{1.9}$$

where $P(a|x, \lambda)$ and $P(b|y, \lambda)$ are valid probability distributions, describing the local response of each party given the hidden variable. Then, it is natural to state that

[1] The term *beable*, introduced by John Bell in [55], refers to elements of a physical theory that are supposed to correspond to objective features of reality—what "is" rather than what is merely "observed." Unlike *observables* in standard quantum mechanics, which are associated with measurement outcomes, beables are intended to describe the physical state of a system independently of observation.

Definition 1.2. A joint probability distribution $P(a,b|x,y)$ is said to be *local* if it can be expressed as a convex mixture of product distributions conditioned on a shared hidden variable λ. That is,

$$P(a,b|x,y) = \int_\Lambda \mathrm{d}\lambda\, Q(\lambda) P(a|x,\lambda) P(b|y,\lambda), \tag{1.10}$$

where:

- λ is a hidden variable, drawn from a probability distribution $Q(\lambda)$ over some measurable space Λ;
- $P(a|x,\lambda)$ and $P(b|y,\lambda)$ are local response functions, representing valid conditional probability distributions for Alice and Bob, respectively.

The set of all probability distributions that admit such a decomposition defines the *local set* $\mathcal{L}$.

In this formulation, Alice's and Bob's outcomes depend only on their local settings and the shared variable λ, in line with the requirement of relativistic local causality.

Remark 1.2. It is essential to emphasize that locality is a property of the model, not of the hidden variable. The hidden variable λ may encode any kind of shared information, but the requirement of locality is that the responses of Alice and Bob depend only on their respective inputs, once λ is fixed.

A particularly relevant class of local models is that of the *deterministic strategies*, where for each value of λ the response functions $P(a|x,\lambda)$ and $P(b|y,\lambda)$ take values in $\{0,1\}$. That is, for each λ, the outputs are deterministically fixed functions of the inputs:

$$a = f_A(x,\lambda), \qquad b = f_B(y,\lambda). \tag{1.11}$$

We now prove that deterministic local hidden-variable (LHV) models are equivalent in their expressive power to general (stochastic) LHV models when it comes to describing local correlations. This means that any correlation achievable by a stochastic LHV model can also be achieved by a deterministic one, provided the hidden variable itself can encompass additional randomness.

Proposition 1.1. *Every local behavior $\boldsymbol{P} \in \mathcal{L}$ generated by a general (stochastic) local hidden-variable model can also be generated by a deterministic local hidden-variable model.*

Proof. Let us start with a general local behavior $\boldsymbol{P}$ defined by (1.10), with stochastic response functions $P(a|x,\lambda)$ and $P(b|y,\lambda)$, and a hidden variable distribution $Q(\lambda)$. We aim to construct an equivalent deterministic model.

First, define a new, augmented hidden variable

$$\lambda' \coloneqq (\lambda, \mu_A, \mu_B), \tag{1.12}$$

where $\mu_A, \mu_B \in [0,1]$ are independent uniformly distributed random numbers. These additional random numbers will serve to "determinize" the local probabilistic choices.

Next, we define new deterministic response functions, $P'(a|x,\lambda')$ for Alice and $P'(b|y,\lambda')$ for Bob. For Alice, we use her original stochastic response function $P(a|x,\lambda)$ to build a cumulative distribution function (CDF)

$$F(a|x,\lambda) \coloneqq \sum_{a_0 \leq a} P(a_0|x,\lambda). \tag{1.13}$$

Using this, we construct Alice's new deterministic response function

$$P'(a|x,\lambda') \coloneqq \begin{cases} 1 & \text{if } F(a-1|x,\lambda) \leq \mu_A < F(a|x,\lambda), \\ 0 & \text{otherwise.} \end{cases} \tag{1.14}$$

Here, $F(a-1|x,\lambda)$ for the smallest outcome a is taken to be 0. This construction ensures that for a fixed λ and a random μ_A, $P'(a|x,\lambda')$ assigns 1 to exactly one outcome a, thus making the response deterministic. A similar construction applies for Bob, using μ_B to define $P'(b|y,\lambda')$.

Finally, we define the new distribution over the augmented hidden variable λ' as

$$Q'(\lambda') \coloneqq Q(\lambda). \tag{1.15}$$

Now, let's verify that this deterministic model recovers the predictions of the original stochastic model. The joint probability $P'(a,b|x,y)$ in the new deterministic model is

$$P'(a,b|x,y) = \int_\Lambda \mathrm{d}\lambda \int_0^1 \mathrm{d}\mu_A \int_0^1 \mathrm{d}\mu_B \; P'(a|x,\lambda,\mu_A)P'(b|y,\lambda,\mu_B)Q(\lambda).$$

Due to the independence of μ_A and μ_B, we can separate the integrals:

$$P'(a,b|x,y) = \int_\Lambda \mathrm{d}\lambda \left(\int_0^1 \mathrm{d}\mu_A \; P'(a|x,\lambda,\mu_A)\right)\left(\int_0^1 \mathrm{d}\mu_B \; P'(b|y,\lambda,\mu_B)\right) Q(\lambda).$$

Consider the inner integral for Alice:

$$\int_0^1 \mathrm{d}\mu_A \; P'(a|x,\lambda,\mu_A).$$

By definition of $P'(a|x,\lambda')$ in (1.14), this integral is 1 if μ_A falls within the interval $[F(a-1|x,\lambda), F(a|x,\lambda))$, and 0 otherwise. The length of this interval is $F(a|x,\lambda) - F(a-1|x,\lambda)$, which by definition of the CDF (1.13) is precisely $P(a|x,\lambda)$. So,

$$\int_0^1 \mathrm{d}\mu_A \; P'(a|x,\lambda,\mu_A) = P(a|x,\lambda).$$

and a similar result holds for Bob.

Substituting these back into the expression for $P'(a,b|x,y)$, one obtains

$$P'(a,b|x,y) = \int_\Lambda \mathrm{d}\lambda \; Q(\lambda)P(a|x,\lambda)P(b|y,\lambda). \tag{1.16}$$

This is exactly the expression for the general stochastic local behavior $\boldsymbol{P}$ given by (1.10).

This construction, originally introduced by Fine [56], shows that every stochastic local model can be simulated by a deterministic one by incorporating the necessary randomness into the definition of the hidden variable itself. □

Correlations that do not admit a decomposition of the form (1.10) are called *nonlocal*. This can only occur in scenarios with at least two inputs and two outputs ($m, k \geq 2$); otherwise, all behaviors in $\mathcal{P}$ can be modeled locally.[2]

1.2.3 Quantum correlations

Finally, we consider the set of behaviors achievable in quantum mechanics.

Definition 1.3. For a given Bell scenario, the *quantum set* $\mathcal{Q}$ of quantum behaviors corresponds to the elements of $\mathcal{P}$ that can be written as

$$P(a, b|x, y) = \mathrm{Tr}\Big((\hat{M}^A_{a|x} \otimes \hat{M}^B_{b|y})\rho_{AB}\Big), \tag{1.17}$$

where $\hat{\rho}_{AB}$ is a density operator acting in the bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ of unspecified dimension, $\hat{M}^A_{a|x}$, ($\hat{M}^B_{b|y}$) are Positive Operator-Valued Measures (POVM) elements on $\mathcal{H}_A$ ($\mathcal{H}_B$) characterizing Alice's (Bob's) measurements.

Note that the operators $\hat{M}^A_{a|x}$ and $\hat{M}^B_{b|y}$ are not necessarily orthogonal projectors. The only requirement is that they are normalized positive semidefinite operators, i.e.,

$$\hat{M}^A_{a|x}, \hat{M}^B_{b|y} \succeq 0, \qquad \sum_{a=1}^{k} \hat{M}^A_{a|x} = \mathbb{1}_A, \quad \sum_{b=1}^{k} \hat{M}^B_{b|y} = \mathbb{1}_B. \tag{1.18}$$

Remark 1.3. The fact that the dimensions of the Hilbert space is left unspecified is significant for several reasons. First, it ensures that the players' strategies are not artificially constrained. In particular, the shared quantum state $\hat{\rho}_{AB}$ need not represent only entanglement between the parties: it may also include arbitrary local (separable) information.

Second, this flexibility allows us to adopt two important simplifications without loss of generality:

1. Purification of the state: any mixed state $\hat{\rho}_{AB}$ shared by the players can be viewed as the reduced state of a larger pure state $|\psi\rangle$ defined on an extended Hilbert space (purification theorem).

2. Projective measurements: any general measurement described by a Positive Operator-Valued Measurerement (POVM) can be implemented as a projective measurement (i.e., one using orthogonal projection operators) on a larger Hilbert space. This follows from Naimark's dilation theorem[3], which guarantees that for any POVM, there exists an extended space in which the same measurement statistics can be obtained via a standard projective measurement.

[2] Nonlocal correlations can only arise in Bell scenarios where both the number of measurement settings and the number of outcomes per party satisfy $m, k \geq 2$. If $m = 1$, each party performs only one measurement, so the observed behavior reduces to a fixed joint distribution $P(a, b)$. In this case, any correlation — no matter how strong — can be reproduced by a local model, since there is no dependence on measurement choices. Likewise, if $k = 1$, the outputs are fixed, rendering the correlations trivial ($a = a_0, b = b_0$ and $P(a_0, b_0|x, y) = 1 \quad \forall x, y$). Thus, both measurement choice and outcome variability are essential for revealing nonlocality.

[3] See Appendix A for details.

That is, we can equivalently write a quantum behavior as

$$P(a,b|x,y) = \langle\psi|\,\hat{\Pi}^A_{a|x} \otimes \hat{\Pi}^B_{b|y}\,|\psi\rangle\,, \tag{1.19}$$

where $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ and

$$\hat{\Pi}^A_{a|x}\hat{\Pi}^A_{a'|x} = \delta_{aa'}\hat{\Pi}^A_{a|x}, \qquad \sum_a \hat{\Pi}^A_{a|x} = \mathbb{1}_A, \tag{1.20a}$$

$$\hat{\Pi}^B_{b|y}\hat{\Pi}^B_{b'|y} = \delta_{bb'}\hat{\Pi}^B_{b|y}, \qquad \sum_b \hat{\Pi}^B_{b|y} = \mathbb{1}_B. \tag{1.20b}$$

A different definition of quantum behaviors is also possible, where instead of imposing a tensor product structure between Alice's and Bob's systems, we merely require that their local operators commute [57]:

Definition 1.4. A behavior $\boldsymbol{P} \in \mathcal{P}$ belongs to the set $\mathcal{Q}'$ of *commuting-operator quantum correlations* if there exist:

- a Hilbert space $\mathcal{H}$,
- a quantum state $|\psi\rangle \in \mathcal{H}$,
- families of positive operators $\{\hat{\Pi}^A_{a|x}\}$, $\{\hat{\Pi}^B_{b|y}\} \subset L(\mathcal{H})$,

such that

$$P(a,b|x,y) = \langle\psi|\,\hat{\Pi}^A_{a|x}\hat{\Pi}^B_{b|y}\,|\psi\rangle\,, \qquad \text{and} \qquad [\hat{\Pi}^A_{a|x}, \hat{\Pi}^B_{b|y}] = 0 \quad \forall a,b,x,y. \tag{1.21}$$

In contrast to the standard tensor-product definition of quantum behaviors (defining the set $\mathcal{Q}$), this commuting-operator model does not assume an explicit partitioning of the Hilbert space into local subsystems.

It is easy to see that $\mathcal{Q} \subseteq \mathcal{Q}'$: in fact, any behavior realized via a tensor-product representation also satisfies the commuting condition, since operators of the form

$$\{\hat{\Pi}^A_{a|x} \otimes \mathbb{1}_B\}, \qquad \{\mathbb{1}_A \otimes \hat{\Pi}^B_{b|y}\}$$

commute.

Whether every commuting-operator model admits a tensor-product representation was long known as Tsirelson's problem [58, 59]. In general,

$$\mathcal{Q} \subseteq \mathcal{Q}' \subseteq \mathcal{NS}, \tag{1.22}$$

but neither inclusion is an equality. For specific scenarios, such as the CHSH case in finite dimensions, it is known that $\mathcal{Q} = \mathcal{Q}'$. However, in the infinite-dimensional setting, this equivalence fails. In particular, Slofstra [60] showed that $\mathcal{Q}$ is not closed, meaning that there exist sequences of quantum behaviors that converge to a point outside $\mathcal{Q}$, although still in $\overline{\mathcal{Q}} \subseteq \mathcal{Q}'$.

This left open the question of whether the closure $\overline{\mathcal{Q}}$ of the quantum set coincides with the commuting-operator set $\mathcal{Q}'$. This question was conclusively resolved by Ji et al. [61], who proved that

$$\overline{\mathcal{Q}} \subsetneq \mathcal{Q}', \tag{1.23}$$

thereby showing that there exist correlations that can be realized using commuting measurements on a single Hilbert space, but that cannot even be approximated by any tensor-product quantum strategy — not even in infinite dimensions. This result demonstrates that the commuting-operator model is strictly more general than the tensor-product model, and implies that the maximum winning probability of certain multiplayer nonlocal games can differ depending on the chosen quantum model [61].

1.3 Geometry of correlations

Let us now consider some geometrical properties of the sets of correlations introduced so far, namely the local set $\mathcal{L}$, the quantum set $\mathcal{Q}$, and the no-signaling set $\mathcal{NS}$.

One can show that any local behavior, i.e., any probability distribution that admits a decomposition as in Eq. (1.10), also admits a quantum realization of the form (1.17), and thus belongs to $\mathcal{Q}$ [62]. In this sense, local correlations form a subset of quantum correlations, namely

$$\mathcal{L} \subseteq \mathcal{Q}.$$

Moreover, quantum correlations automatically satisfy the no-signaling constraints, since measurements on separate subsystems of a quantum system cannot transmit information instantaneously (*no-communication theorem*[4]). Therefore, it holds

$$\mathcal{Q} \subseteq \mathcal{NS}.$$

However, both inclusions are known to be strict:

- Bell famously showed that there exist quantum correlations that cannot be reproduced by any local model, i.e., $\mathcal{L} \subsetneq \mathcal{Q}$ [9].
- Later, Popescu and Rohrlich demonstrated that there exist hypothetical no-signaling correlations (such as PR-box correlations) that exceed the bounds imposed by quantum theory, i.e., $\mathcal{Q} \subsetneq \mathcal{NS}$ [53].

In general, we thus have the strict inclusions

$$\mathcal{L} \subsetneq \mathcal{Q} \subsetneq \mathcal{NS}. \tag{1.24}$$

From a physical perspective, this hierarchy reflects the increasingly general assumptions made about the nature of correlations:

- $\mathcal{L}$ contains all behaviors that can be explained by the only presence of shared classical information, i.e., local hidden variable models;
- $\mathcal{Q}$ allows for the use of quantum entanglement and measurements on composite systems;
- $\mathcal{NS}$ includes all correlations that obey the relativistic no-signaling constraint, regardless of their physical realizability.

[4] See Appendix A for further details.

Furthermore, it can be shown that $\dim \mathcal{L} = \dim \mathcal{Q} = \dim \mathcal{NS} = t$ [63], where

$$t \coloneqq 2(k-1)m + (k-1)^2 m^2. \tag{1.25}$$

In the case of binary outcomes ($k = 2$), a useful parametrization is given in terms of $2m + m^2$ quantities, collectively called *correlators*, which include both the *correlations*

$$\langle A_x B_y \rangle \coloneqq \sum_{a,\,b} ab\, P(a,b|x,y), \tag{1.26a}$$

and the marginals

$$\langle A_x \rangle \coloneqq \sum_{a} a\, P(a|x), \quad \langle B_y \rangle \coloneqq \sum_{b} b\, P(b|y). \tag{1.26b}$$

In the CHSH scenario, this change of variables enables the study of a behavior $\boldsymbol{P} \in \mathbb{R}^8$, by representing it in table form, structured as follows:

$$\boldsymbol{P} = \begin{array}{c|c|c} & \langle B_0 \rangle & \langle B_1 \rangle \\ \hline \langle A_0 \rangle & \langle A_0 B_0 \rangle & \langle A_0 B_1 \rangle \\ \langle A_1 \rangle & \langle A_1 B_0 \rangle & \langle A_1 B_1 \rangle \end{array}, \tag{1.27}$$

where each entry corresponds to a specific expectation value or joint probability.

For quantum behaviors, in the CHSH scenario, it is customary to express the measurements performed by Alice and Bob in terms of Hermitian[5] operators $\hat{A}_x \in L(\mathcal{H}_A)$ and $\hat{B}_y \in L(\mathcal{H}_B)$, each with eigenvalues in $\{-1, 1\}$ and defined as

$$\hat{A}_x \coloneqq \hat{\Pi}^A_{1|x} - \hat{\Pi}^A_{-1|x}, \qquad \hat{B}_y \coloneqq \hat{\Pi}^B_{1|y} - \hat{\Pi}^B_{-1|y}, \tag{1.28}$$

and satisfy

$$\hat{A}_x^2 = \mathbb{1}_A, \qquad \hat{B}_y^2 = \mathbb{1}_B, \qquad [\hat{A}_x, \hat{B}_y] = 0 \qquad \forall x, y \in \{0, 1\}. \tag{1.29}$$

These operators provide a convenient expression for the correlators:

$$\langle A_x B_y \rangle = \mathrm{Tr}\left[(\hat{A}_x \otimes \hat{B}_y)\hat{\rho}_{AB}\right], \tag{1.30a}$$

$$\langle A_x \rangle = \mathrm{Tr}\left[(\hat{A}_x \otimes \mathbb{1}_B)\hat{\rho}_{AB}\right], \tag{1.30b}$$

$$\langle B_y \rangle = \mathrm{Tr}\left[(\mathbb{1}_A \otimes \hat{B}_y)\rho_{AB}\right]. \tag{1.30c}$$

We refer to $(\hat{\rho}, \hat{A}_x, \hat{B}_y)$ as a *realization*, where each realization uniquely determines a correlator vector via the equations above.

Remarkably, the sets $\mathcal{L}$, $\mathcal{Q}$, and $\mathcal{NS}$ are all closed, bounded, and convex [62]. This means that if $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are elements of any one of these sets, then any convex combination $\mu\boldsymbol{P}_1 + (1-\mu)\boldsymbol{P}_2$ with $0 \leq \mu \leq 1$ also belongs to the same set. The convexity of the non-signaling set $\mathcal{NS}$ follows directly from the fact that it is defined by linear constraints, which are preserved under convex combinations. As for the local set $\mathcal{L}$, its convexity arises from its construction: it consists of

[5] Strictly speaking, quantum observables are represented by self-adjoint operators. In infinite-dimensional Hilbert spaces, Hermitian and self-adjoint operators are not equivalent; however, in the finite-dimensional case considered throughout this thesis, the two notions coincide.

all convex mixtures of deterministic (separable) probability distributions, as described in Eq. (1.10).

We now show that $\mathcal{Q}$ is convex, i.e., if $P(a,b|x,y)$ and $Q(a,b|x,y)$ belong to $\mathcal{Q}$, then their convex combination

$$R(a,b|x,y) \coloneqq \lambda P(a,b|x,y) + (1-\lambda)Q(a,b|x,y), \quad \text{for } 0 \leq \lambda \leq 1, \tag{1.31}$$

also belongs to $\mathcal{Q}$.

Proof. Since $P(a,b|x,y)$ and $Q(a,b|x,y)$ are in $\mathcal{Q}$, there exist the quantum realizations

$$P(a,b|x,y) = \operatorname{Tr}\left[(\hat{M}^A_{a|x} \otimes \hat{M}^B_{b|y})\hat{\rho}_1\right], \tag{1.32a}$$

and

$$Q(a,b|x,y) = \operatorname{Tr}\left[(\hat{N}^A_{a|x} \otimes \hat{N}^B_{b|y})\hat{\rho}_2\right], \tag{1.32b}$$

for some states $\hat{\rho}_1$, $\hat{\rho}_2$ and POVMs $\{\hat{M}^A_{a|x}, \hat{M}^B_{b|y}\}$, $\{\hat{N}^A_{a|x}, \hat{N}^B_{b|y}\}$.
To construct a quantum realization of $R(a,b|x,y)$, consider the new bipartite state:

$$\hat{\rho} \coloneqq \lambda\hat{\rho}_1 \oplus (1-\lambda)\hat{\rho}_2, \tag{1.33}$$

acting on the direct sum Hilbert space

$$\mathcal{H}_A \otimes \mathcal{H}_B \coloneqq (\mathcal{H}^1_A \otimes \mathcal{H}^1_B) \oplus (\mathcal{H}^2_A \otimes \mathcal{H}^2_B). \tag{1.34}$$

Similarly, define the block-diagonal measurement operators:

$$\hat{O}^A_{a|x} \coloneqq \hat{M}^A_{a|x} \oplus \hat{N}^A_{a|x}, \qquad \hat{O}^B_{b|y} \coloneqq \hat{M}^B_{b|y} \oplus \hat{N}^B_{b|y}. \tag{1.35}$$

By construction, these operators satisfy the completeness conditions:

$$\sum_a \hat{O}^A_{a|x} = \mathbb{1}_A, \qquad \sum_b \hat{O}^B_{b|y} = \mathbb{1}_B. \tag{1.36}$$

Now, computing the probability distribution using this realization:

$$\begin{aligned} R(a,b|x,y) &\coloneqq \operatorname{Tr}\left[(\hat{O}^A_{a|x} \otimes \hat{O}^B_{b|y})\hat{\rho}\right] \\ &= \operatorname{Tr}\left[(\hat{M}^A_{a|x} \oplus \hat{N}^A_{a|x}) \otimes (\hat{M}^B_{b|y} \oplus \hat{N}^B_{b|y})(\lambda\hat{\rho}_1 \oplus (1-\lambda)\hat{\rho}_2)\right] \\ &= \lambda \operatorname{Tr}\left[(\hat{M}^A_{a|x} \otimes \hat{M}^B_{b|y})\hat{\rho}_1\right] + (1-\lambda)\operatorname{Tr}\left[(\hat{N}^A_{a|x} \otimes \hat{N}^B_{b|y})\hat{\rho}_2\right]. \end{aligned}$$

Using the definitions of $P(a,b|x,y)$ and $Q(a,b|x,y)$, we conclude that

$$R(a,b|x,y) = \lambda P(a,b|x,y) + (1-\lambda)Q(a,b|x,y). \tag{1.37}$$

Since we have explicitly constructed a quantum state and measurement scheme that produces $R(a,b|x,y)$, it follows that $\boldsymbol{R} \in \mathcal{Q}$. This proves that the quantum set is convex. □

While discussing the geometry of the correlation sets, it is natural to employ standard tools from convex geometry —most notably the notions of boundary points, extremal points, and exposed points[6]. Let $\mathcal{K} \subset \mathbb{R}^d$ be a compact convex set (i.e., closed and bounded). A point $\boldsymbol{u} \in \mathcal{K}$ is called a *boundary point* if it belongs to the topological boundary of $\mathcal{K}$; that is, it does not lie in the interior of the set. A boundary point is said to be *extremal* if it cannot be written as a non-trivial convex combination of other points in $\mathcal{K}$. The set of all extremal points of $\mathcal{K}$ is denoted $\mathcal{K}_{\text{ext}}$. A stronger condition is that of *exposedness*: a point $\boldsymbol{u} \in \mathcal{K}$ is *exposed* if there exists a vector $\boldsymbol{v} \in \mathbb{R}^d$ such that $\boldsymbol{u}$ uniquely maximizes the inner product $\boldsymbol{v} \cdot \boldsymbol{w}$ over all $\boldsymbol{w} \in \mathcal{K}$. That is, the set

$$F_{\boldsymbol{v}} \coloneqq \left\{ \boldsymbol{w} \in \mathcal{K} \colon \boldsymbol{v} \cdot \boldsymbol{w} = \max_{\boldsymbol{z} \in \mathcal{K}} \boldsymbol{v} \cdot \boldsymbol{z} \right\} \tag{1.38}$$

is reduced to the singleton $\{\boldsymbol{u}\}$. In this case, the hyperplane defined by $\boldsymbol{v} \cdot \boldsymbol{w} = \max_{\boldsymbol{z} \in \mathcal{K}} \boldsymbol{v} \cdot \boldsymbol{z}$ is a *supporting hyperplane* of $\mathcal{K}$, and the set $F_{\boldsymbol{v}}$ is referred to as the *exposed face of* $\mathcal{K}$ *associated to* $\boldsymbol{v}$.

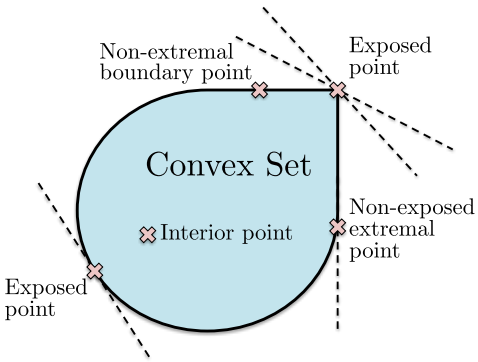

Figure 1.3: Different types of points of a compact convex set. From [46].

By construction, we always have the chain of inclusions:

$$\mathcal{K}_{\text{exp}} \subseteq \mathcal{K}_{\text{ext}} \subseteq \mathcal{K}_{\text{bnd}} \subseteq \mathcal{K}, \tag{1.39}$$

and in general, each inclusion is strict. This hierarchy is illustrated in Fig. 1.3. An important structural result in convex geometry, known as the *Krein-Milman theorem* [64], states that any compact convex set $\mathcal{K}$ is equal to the convex hull of its extremal points:

$$\mathcal{K} = \text{conv}(\mathcal{K}_{\text{ext}}). \tag{1.40}$$

Hence, optimization of a linear functional over $\mathcal{K}$ always attains its maximum at an extremal point. However, the maximizer need not be unique, nor does it need to be an exposed point. This leads to the distinction between extremal and exposed points: although every exposed point is extremal, the converse is not true in general. Still, a classical result due to Straszewicz [64] shows that in finite-dimensional spaces, the set of exposed points is dense in the set of extremal points. That is, extremal but non-exposed points exist, but they are in a certain sense exceptional.

Remark 1.4. For polytopes, i.e. convex hulls of finitely many points, this hierarchy collapses: all extremal points are exposed, and they correspond to the vertices of the polytope.

[6] Formal definitions are collected in Appendix B.

1.4 Bell inequalities

Since the sets $\mathcal{L}$, $\mathcal{Q}$, and $\mathcal{NS}$ are convex, the *hyperplane separation theorem*[7] ensures that any behavior $\boldsymbol{P}' \in \mathbb{R}^t$ lying outside one of these sets can be separated from it by a linear inequality. That is, if $\boldsymbol{P}' \notin \mathcal{K}$, where $\mathcal{K}$ denotes either $\mathcal{L}$, $\mathcal{Q}$, or $\mathcal{NS}$, then there exists a vector $\boldsymbol{\beta} \in \mathbb{R}^t$ and a constant $\beta_{\mathcal{K}} \in \mathbb{R}$ such that

$$\boldsymbol{\beta} \cdot \boldsymbol{P} \leq \beta_{\mathcal{K}} \quad \text{for all } \boldsymbol{P} \in \mathcal{K}, \qquad \text{while} \qquad \boldsymbol{\beta} \cdot \boldsymbol{P}' > \beta_{\mathcal{K}}. \tag{1.41}$$

In the case where $\mathcal{K} \equiv \mathcal{L}$, the inequality

$$\boldsymbol{\beta} \cdot \boldsymbol{P} \leq \beta_{\mathcal{L}} \tag{1.42}$$

is known as a *Bell inequality*. Such inequalities delineate the boundary of the local set and provide a criterion for detecting nonlocality: any behavior that violates a Bell inequality is necessarily nonlocal.

The quantity $\beta_{\mathcal{L}}$ is called the *local bound* of the Bell expression $\boldsymbol{\beta}$, and it represents the maximum value that this expression can attain over all local correlations. If the inequality

$$\boldsymbol{\beta} \cdot \boldsymbol{P} \leq \beta_{\mathcal{L}}$$

is violated by a quantum behavior $\boldsymbol{P} \in \mathcal{Q}$, then quantum theory exhibits a departure from classical locality. In such cases, the Bell inequality is said to be *violated by quantum mechanics*.

More generally, any inequality of the form

$$\boldsymbol{\beta} \cdot \boldsymbol{P} \leq \beta_{\mathcal{Q}}$$

that holds for all $\boldsymbol{P} \in \mathcal{Q}$ is referred to as a *Tsirelson inequality*. Similarly, one defines the *quantum bound* (or *Tsirelson bound*) $\beta_{\mathcal{Q}}$ as the maximal value of the same expression over the quantum set. These inequalities characterize the boundary of the quantum set within the no-signaling polytope and provide a natural generalization of Bell inequalities when the focus is shifted from classical to quantum correlations.

1.4.1 Bell expressions and Bell operators

We now provide a precise formulation of the relevant concepts. For a given a Bell scenario, let $\mathcal{P}$ denote the set of all admissible behaviors.

Definition 1.5. A *Bell expression* β is a linear functional on $\mathcal{P}$, which associates to each behavior $\boldsymbol{P} \in \mathcal{P}$ a real number called the *Bell score*

$$\beta(\boldsymbol{P}) \coloneqq \sum_{a,b,x,y} \beta_{a,b,x,y}, P(a,b|x,y), \tag{1.43}$$

defined by a set of real coefficients $\{\beta_{a,b,x,y}\}$

[7] See Appendix B for details.

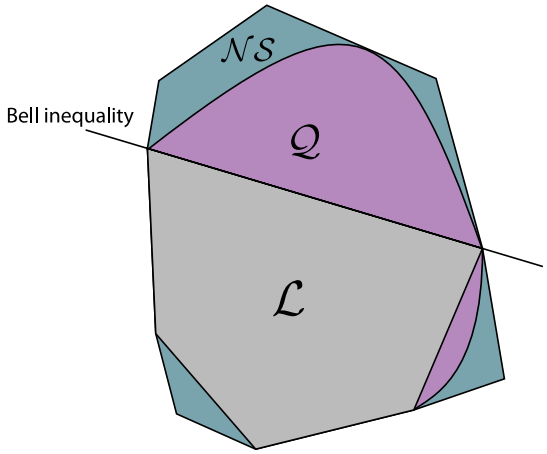

Figure 1.4: Relations between the no-signaling, local and quantum set. From [19].

If the space of behaviors $\mathcal{P}$ is represented as a submanifold of the real vector space $\mathbb{R}^t$, we can associate to each Bell expression β a vector in $\boldsymbol{\beta} \in \mathbb{R}^t$, where the components $\beta_{a,b,x,y}$ are ordered consistently with the representation of behaviors. In this sense, a Bell expression defines a linear map (a covector) on $\mathbb{R}^t$, and evaluating the Bell score translates in computing the Euclidean scalar product

$$\beta(\boldsymbol{P}) = \boldsymbol{\beta} \cdot \boldsymbol{P}. \tag{1.44}$$

If the behavior $\boldsymbol{P}$ arises from a quantum realization $(|\psi\rangle, \{\hat{\Pi}^A_{a|x}\}, \{\hat{\Pi}^B_{b|y}\})$, then associated with the Bell expression $\beta(\boldsymbol{P})$ and the measurement choices, the corresponding *Bell operator* $\hat{S} \in L(\mathcal{H}_A \otimes \mathcal{H}_B)$ is defined as

$$\hat{S} \coloneqq \sum_{a,b,x,y} \beta_{a,b,x,y} \hat{\Pi}^A_{a|x} \otimes \hat{\Pi}^B_{b|y}. \tag{1.45}$$

This operator captures the structure of the Bell expression for a fixed choice of measurement operators. The quantum value of the Bell expression is then obtained by computing its expectation value on the given state:

$$\boldsymbol{\beta} \cdot \boldsymbol{P} = \langle\psi| \hat{S} |\psi\rangle . \tag{1.46}$$

As we have seen, in the CHSH scenario ($t = 8$), a behavior $\boldsymbol{P}$ can be parametrized by (1.27), seen as a point in $\mathbb{R}^8$. Then, a general Bell expression is represented as a linear combination of expectation values:

$$\boldsymbol{\beta} \cdot \boldsymbol{P} = \sum_x \alpha_x \langle A_x \rangle + \sum_y \beta_y \langle B_y \rangle + \sum_{x,y} \gamma_{xy} \langle A_x B_y \rangle,$$

where $\alpha_x, \beta_y, \gamma_{xy} \in \mathbb{R}$ are the components of $\boldsymbol{\beta} \in \mathbb{R}^8$. The corresponding Bell operator, for a given choice of measurements $\{\hat{A}_x\}$, $\{\hat{B}_y\}$, reads

$$\hat{S} \coloneqq \sum_x \alpha_x \hat{A}_x \otimes \mathbb{1}_B + \sum_y \beta_y \mathbb{1}_A \otimes \hat{B}_y + \sum_{x,y} \gamma_{xy} \hat{A}_x \otimes \hat{B}_y. \tag{1.47}$$

Remark 1.5. Note that:

- In writing $\boldsymbol{\beta} \cdot \boldsymbol{P}$, the vector $\boldsymbol{\beta}$ encodes the functional form of the Bell expression (independent of any physical implementation), and $\boldsymbol{P}$ contains the full set of expectation values generated by a specific behavior —whether classical, quantum, or otherwise.

- In the expression $\langle\psi|\,\hat{S}\,|\psi\rangle$, the Bell operator $\hat{S}$ already includes the information on the measurement choices $\{\hat{A}_x\}, \{\hat{B}_y\}$, and the quantum state $|\psi\rangle$ provides the complementary information relative of the quantum realization.

In particular, for a fixed Bell expression β, different choices of observables $\{\hat{A}_x\}, \{\hat{B}_y\}$ give rise to different Bell operators S, which in turn lead to different quantum expectation values. Conversely, for a fixed behavior $\boldsymbol{P}$, the value $\boldsymbol{\beta} \cdot \boldsymbol{P}$ is completely determined by the underlying correlators.

1.4.2 The CHSH inequality

The CHSH inequality, introduced by Clauser, Horne, Shimony, and Holt in 1969 [65], is a Bell inequality formulated for the simplest nontrivial Bell scenario, where $(n, m, k) = (2, 2, 2)$. In this setting, the CHSH inequality is expressed as

$$\langle A_0 B_0\rangle + \langle A_0 B_1\rangle + \langle A_1 B_0\rangle - \langle A_1 B_1\rangle \leq 2. \tag{1.48}$$

This inequality is satisfied by all probability distributions that admit a local hidden-variable (LHV) model, i.e., all $\boldsymbol{P} \in \mathcal{L}$.

Proof. Consider a CHSH scenario. In a local theory, the outcome of Alice's measurement depends only on her setting and eventually on a hidden variable λ, and similarly for Bob. Thus, we can represent the outcomes as deterministic functions of the hidden variable λ, $A_x(\lambda), B_y(\lambda) \in \{-1, +1\}$. We define the CHSH quantity $S_{\text{CHSH}}(\lambda)$ as:

$$S_{\text{CHSH}}(\lambda) \coloneqq A_0(\lambda)B_0(\lambda) + A_1(\lambda)B_0(\lambda) + A_0(\lambda)B_1(\lambda) - A_1(\lambda)B_1(\lambda). \tag{1.49}$$

This can be rearranged as:

$$S_{\text{CHSH}}(\lambda) = A_0(\lambda)(B_0(\lambda) + B_1(\lambda)) + A_1(\lambda)(B_0(\lambda) - B_1(\lambda)). \tag{1.50}$$

Let's analyze the possible values of $S_{\text{CHSH}}(\lambda)$ for any given λ. Since $B_0(\lambda)$ and $B_1(\lambda)$ can only take values of $+1$ or -1,

- If $B_0(\lambda) = B_1(\lambda)$, $(B_0 - B_1)(\lambda) = 0$, and $(B_0 + B_1)(\lambda) = \pm 2$. In this case, $S_{\text{CHSH}}(\lambda) = \pm 2A_1(\lambda)$.

- If $B_0(\lambda) = -B_1(\lambda)$, then $(B_0 - B_1)(\lambda) = \pm 2$, and $(B_0 + B_1)(\lambda) = 0$. In this case, $S(\lambda) = \pm 2A_0(\lambda)$.

Therefore, in all possible scenarios,

$$|S_{\text{CHSH}}(\lambda)| \leq 2 \tag{1.51}$$

Let $Q(\lambda)$ be the probability distribution of the hidden variable $\lambda \in \Lambda$, satisfying $\int_\Lambda \mathrm{d}\lambda\, Q(\lambda) = 1$. The expectation value of S_{CHSH} is given by:

$$\langle S_{\text{CHSH}} \rangle \coloneqq \int_\Lambda \mathrm{d}\lambda\, S_{\text{CHSH}}(\lambda) Q(\lambda) \tag{1.52}$$

Due to the linearity of expectation values, this can be written as:

$$\langle S_{\text{CHSH}} \rangle = \langle A_0 B_0 \rangle + \langle A_1 B_0 \rangle + \langle A_0 B_1 \rangle - \langle A_1 B_1 \rangle \tag{1.53}$$

where $\langle A_x B_y \rangle = \int_\Lambda \mathrm{d}\lambda\, A_x(\lambda) B_y(\lambda) \rho(\lambda)$.

Since $|S(\lambda)| \leq 2$ for all λ, we can integrate this inequality over λ:

$$\int_\Lambda \mathrm{d}\lambda\, |S_{\text{CHSH}}(\lambda)| Q(\lambda) \leq 2 \int_\Lambda \mathrm{d}\lambda\, Q(\lambda), \tag{1.54}$$

that simplifies to

$$|\langle S_{\text{CHSH}} \rangle| \leq 2. \tag{1.55}$$

Thus, for any local hidden variable theory, the expectation value of the CHSH quantity must satisfy:

$$\langle A_0 B_0 \rangle + \langle A_1 B_0 \rangle + \langle A_0 B_1 \rangle - \langle A_1 B_1 \rangle \leq 2. \tag{1.56}$$

□

In quantum mechanics, the CHSH inequality can be violated. The maximal value achievable by quantum correlations is given by the *Tsirelson bound* [21]:

$$\langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle \leq 2\sqrt{2}. \tag{1.57}$$

Proof. Let $\hat{A}_0$, $\hat{A}_1$, $\hat{B}_0$, and $\hat{B}_1$ be Hermitian operators in the CHSH scenario, each with eigenvalues in ± 1. This implies that:

$$\hat{A}_0^2 = \hat{A}_1^2 = \hat{B}_0^2 = \hat{B}_1^2 = \mathbb{1}.$$

Assume furthermore that Alice's observables commute with Bob's ones:

$$[\hat{A}_x, \hat{B}_y] = 0 \qquad \text{for all } x, y \in \{0, 1\},$$

while $[\hat{A}_0, \hat{A}_1] \neq 0$, $[\hat{B}_0, \hat{B}_1] \neq 0$. Define the CHSH Bell operator[8]:

$$\hat{S}_{\text{CHSH}} \coloneqq \hat{A}_0 \hat{B}_0 + \hat{A}_0 \hat{B}_1 + \hat{A}_1 \hat{B}_0 - \hat{A}_1 \hat{B}_1.$$

[8] To simplify notation, we omit explicit tensor products between operators acting on different subsystems when there is no ambiguity. For instance, expressions like $\hat{A}_x \hat{B}_y$ should be understood as $\hat{A}_x \otimes \hat{B}_y$. This convention will be used throughout the thesis.

We wish to compute the maximum possible expectation value $\langle\psi|\hat{S}_{\text{CHSH}}|\psi\rangle$ over all quantum states $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. To this end, consider the operator norm:

$$\|\hat{C}\|_{\text{sup}} \coloneqq \sup_{|\psi\rangle} \frac{\|\hat{C}\,|\psi\rangle\,\|}{\|\,|\psi\rangle\,\|}.$$

We evaluate an upper bound on $\|\hat{S}_{\text{CHSH}}\|_{\text{sup}}$ by computing $\hat{S}^2_{\text{CHSH}}$. A direct algebraic manipulation yields:

$$\hat{S}^2_{\text{CHSH}} = 4\mathbb{1} \otimes \mathbb{1} - [\hat{A}_0, \hat{A}_1] \otimes [\hat{B}_0, \hat{B}_1].$$

Using the submultiplicativity of the operator norm, and the fact that all operators involved are bounded with $\|\hat{A}_x\|_{\text{sup}}, \|\hat{B}_y\|_{\text{sup}} \leq 1$, we obtain:

$$\|\hat{S}^2_{\text{CHSH}}\|_{\text{sup}} \leq 4 + \|[\hat{A}_0, \hat{A}_1]\|_{\text{sup}} \cdot \|[\hat{B}_0, \hat{B}_1]\|_{\text{sup}} \leq 8.$$

Since $\hat{S}_{\text{CHSH}}$ is Hermitian,

$$\|\hat{S}_{\text{CHSH}}\|^2_{\text{sup}} = \|\hat{S}^2_{\text{CHSH}}\|_{\text{sup}},$$

and thus:

$$\|\hat{S}_{\text{CHSH}}\|_{\text{sup}} \leq 2\sqrt{2}. \tag{1.58}$$

The latter expression is called a *quantum Bell inequality* or a *Tsirelson inequality*. □

The Tsirelson bound is tight: it can be achieved within quantum mechanics. In particular, consider the bipartite maximally entangled state:

$$\left|\phi^+\right\rangle \coloneqq \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \tag{1.59}$$

Let Alice and Bob perform the following projective measurements on their respective subsystems:

$$\hat{A}_0 = \hat{Z}_A, \qquad \hat{A}_1 = \hat{X}_A, \tag{1.60a}$$

$$\hat{B}_0 = \frac{1}{\sqrt{2}}(\hat{Z}_B + \hat{X}_B), \qquad \hat{B}_1 = \frac{1}{\sqrt{2}}(\hat{Z}_B - \hat{X}_B). \tag{1.60b}$$

Here, $\hat{X}_A$ and $\hat{Z}_A$ denote the Pauli $\hat{X}$ and $\hat{Z}$ operators acting on Alice's qubit, while $\hat{X}_B$ and $\hat{Z}_B$ act similarly on Bob's qubit. A straightforward calculation then shows that:

$$\left\langle\phi^+\right| \hat{S}_{\text{CHSH}} \left|\phi^+\right\rangle = 2\sqrt{2}. \tag{1.61}$$

Thus, the quantum correlations arising from these measurements on the entangled state $|\phi^+\rangle$ saturate the Tsirelson bound. The corresponding behavior is represented by the following table

$$\boldsymbol{P}_T \coloneqq \begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \hline 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{array} \tag{1.62}$$

and it is known as the *Tsirelson behavior* or *Tsirelson point*.

Remark 1.6. In order to obtain nonlocal correlations from measurements on a quantum state, it is necessary that the latter is entangled. That is, the state cannot be written in the separable form

$$\hat{\rho}_{AB} = \sum_{\lambda} p_\lambda \hat{\rho}_A^\lambda \otimes \hat{\rho}_B^\lambda. \tag{1.63}$$

Indeed, if a state is of such a form, the correlations obtained by performing local measurements on it are given by

$$\begin{aligned} P(a,b|x,y) &= \mathrm{Tr}\left[\sum_{\lambda} p_\lambda (\hat{\rho}_A^\lambda \otimes \hat{\rho}_B^\lambda) \hat{M}_{a|x}^A \otimes \hat{M}_{b|y}^B\right] \\ &= \sum_{\lambda} p_\lambda \,\mathrm{Tr}\left(\hat{\rho}_A^\lambda \hat{M}_{a|x}^A\right) \mathrm{Tr}\left(\hat{\rho}_B^\lambda \hat{M}_{b|y}^B\right) \\ &= \sum_{\lambda} p_\lambda P(a|x,\lambda) P(b|y,\lambda). \end{aligned} \tag{1.64}$$

which is of the local form (1.10). Hence the observation of nonlocal correlations implies the presence of entanglement.

It is interesting to investigate whether this link can be reversed. That is, do all entangled states lead to nonlocality? In the case of pure states, the answer is positive. Specifically, for any entangled pure state, it is possible to find local measurements such that the resulting correlations violate a Bell inequality, in particular, the CHSH inequality. Therefore, all pure entangled states are nonlocal [66]. The only pure states that do not violate Bell inequalities are the product states

$$|\Psi\rangle = |\psi\rangle_A \otimes |\phi\rangle_B .$$

1.4.3 No-signaling bounds

We now consider the problem of computing bounds on Bell expressions for no-signaling correlations. Contrary to the case of local and quantum correlations, this turns out to be a rather easy task. To understand why note that once the no-signaling constraints Eq. (1.7) are taken into account, e.g., by introducing a parametrization of the relevant affine subspace $\mathbb{R}^t$, the set $\mathcal{NS}$ of no-signaling behaviors is uniquely determined by the set of k^2m^2 positivity inequalities

$$P(a,b|x,y) \geq 0. \tag{1.65}$$

Since $\mathcal{NS}$ is defined by a finite number of linear inequalities, it is a polytope and can thus be described as the convex hull of a finite set of vertices.

1.4.4 Local bounds

Let us now explore how Bell inequalities, i.e., the hyperplanes that define the boundary of the local set $\mathcal{L}$, can be systematically derived. A useful starting point is the observation that local correlations can be represented using *deterministic local hidden-variable models*, as was proved in Prop. 1.1.

Moreover, since we assume a finite number of measurement inputs and outputs, there are only finitely many deterministic response functions. Each hidden variable λ in a deterministic

model corresponds to a complete assignment of outputs to inputs: that is, to a fixed function $x \mapsto a_x$ and $y \mapsto b_y$. Hence, we can identify λ with a tuple

$$\lambda \coloneqq (a_1, \ldots, a_m, b_1, \ldots, b_m), \tag{1.66}$$

which specifies a deterministic output a_x and b_y for each measurement setting $x, y = 1, \ldots, m$.

The corresponding deterministic behavior $\boldsymbol{D}_\lambda$ is given by:

$$D_\lambda(a, b|x, y) \coloneqq \delta_{aa_x}\delta_{bb_y}, \tag{1.67}$$

which has value 1 only if a and b are the outputs assigned by λ for the inputs x and y.

There are k^{2m} such deterministic behaviors, where k is the number of possible outputs per party, and m the number of inputs per party. A general behavior $\boldsymbol{P}$ is local if and only if it can be written as a convex combination of these deterministic points:

$$\boldsymbol{P} = \sum_\lambda q_\lambda \boldsymbol{D}_\lambda, \quad \text{with } q_\lambda \geq 0, \quad \sum_\lambda q_\lambda = 1. \tag{1.68}$$

This formulation is powerful, as it transforms the problem of testing locality into a *linear programming (LP) problem*: determining whether a given point $\boldsymbol{P}$ lies within the convex hull of known deterministic points $\boldsymbol{D}_\lambda$ under the constraints above. However, because the number of variables (k^{2m}) grows exponentially, solving the LP may become computationally intractable in practice.

Remark 1.7. Note that since $\mathcal{L}$ is the convex hull of a finite set of points, it is a polytope. By Minkowski's theorem,[9] any polytope can also be represented as the intersection of a finite number of half-spaces. Therefore, there exists a finite set of inequalities of the form:

$$\boldsymbol{\beta}^i \cdot \boldsymbol{P} \leq \beta^i_{\mathcal{L}}, \quad \forall i \in I, \tag{1.69}$$

such that $\boldsymbol{P} \in \mathcal{L}$ if and only if it satisfies all inequalities in this set. These inequalities are precisely the Bell inequalities, and they define the *facets* of the *local polytope* $\mathcal{L}$. Moreover, the fact that the list of extremal local points is known means that the local bound $\beta_{\mathcal{L}}$ of a Bell inequality can be computed directly, by taking the maximum value over a finite number of deterministic points.

Every linear program admits a dual formulation.[10] The dual LP associated with equation (1.68) has a natural physical interpretation: it corresponds to searching for a Bell inequality violated by $\boldsymbol{P}$. The dual form can be written as:

$$\begin{aligned} \max_{(\boldsymbol{\beta}, \beta_{\mathcal{L}})} \quad & \boldsymbol{\beta} \cdot \boldsymbol{P} - \beta_{\mathcal{L}} \\ \text{subject to} \quad & \boldsymbol{\beta} \cdot \boldsymbol{D}_\lambda - \beta_{\mathcal{L}} \leq 0 \quad \forall \lambda, \\ & \boldsymbol{\beta} \cdot \boldsymbol{P} - \beta_{\mathcal{L}} \leq 1. \end{aligned} \tag{1.70}$$

Then,

[9] See Appendix B for details.

[10] A minimal review of convex programs, including linear and semidefinite programs, is provided in Appendix C.

- If $\boldsymbol{P} \in \mathcal{L}$, then $\boldsymbol{P}$ satisfies all Bell inequalities, and the optimal value S of the dual problem satisfies $S \leq 0$.

- If $\boldsymbol{P} \notin \mathcal{L}$, then the dual program returns a vector $\boldsymbol{\beta}$ and bound $\beta_{\mathcal{L}}$ such that

$$\boldsymbol{\beta} \cdot \boldsymbol{P} = \beta_{\mathcal{L}} + 1 > \beta_{\mathcal{L}}, \tag{1.71}$$

  i.e., $\boldsymbol{P}$ violates the Bell inequality $\boldsymbol{\beta} \cdot \boldsymbol{P} \leq \beta_{\mathcal{L}}$, which is satisfied by all local points (since it's satisfied by all $\boldsymbol{d}_\lambda$).

This demonstrates that solving the dual LP provides an effective method for certifying nonlocality and for explicitly constructing a Bell inequality that $\boldsymbol{P}$ violates.

1.4.5 Quantum bounds

Unlike the sets of local and no-signaling correlations, the set $\mathcal{Q}$ of quantum correlations is generally not a polytope. This implies that $\mathcal{Q}$ cannot be described by a finite number of extremal points or a finite set of linear inequalities. However, some structural features of $\mathcal{L}$ persist within $\mathcal{Q}$.

For instance, all extremal points of $\mathcal{L}$ —i.e., the local deterministic behaviors— are also extremal in $\mathcal{Q}$. Moreover, certain faces of $\mathcal{L}$ are shared with $\mathcal{Q}$. An example is provided by the $(k-1)$-dimensional face defined by the hyperplanes

$$P(a, b|x, y) = 0. \tag{1.72}$$

Although the associated Bell inequalities $P(a, b|x, y) \geq 0$ are never violated by physical correlations, they still define flat regions on the boundary of $\mathcal{Q}$.

In [67], it is shown that some faces of $\mathcal{L}$ correspond to Bell inequalities that are violated by no-signaling correlations but not by quantum ones. These inequalities define common faces between $\mathcal{L}$ and $\mathcal{NS}$ (but not necessarily $\mathcal{Q}$). Previously, an open question remained whether such shared faces could be of maximal dimension, i.e., whether there existed facets of $\mathcal{L}$ that are violated by some $\mathcal{NS}$ behaviors but not by any quantum behavior. This question has since been affirmatively answered in the multipartite (e.g., tripartite) scenario; for instance, the work in [68] demonstrates Bell inequalities (related to the "guess your neighbor's input" game) that correspond to facets of $\mathcal{L}$, are violated by no-signaling correlations, yet yield no quantum advantage (i.e., are not violated by quantum correlations).

1.4.5.1 Maximizing the norm of the Bell operator

Let us now focus more specifically on the problem of computing the quantum bound of a Bell expression. Recall that $\mathcal{Q}$ as any convex compact set can be described by an infinite system of linear inequalities. Given a Bell expression defined by a vector $\boldsymbol{\beta}$, the quantum set $\mathcal{Q}$ can be characterized as the set of conditional probability distributions $\boldsymbol{P} = \{P(a, b|x, y)\}$ satisfying all Tsirelson inequalities of the form:

$$\boldsymbol{\beta} \cdot \boldsymbol{P} \leq \beta_{\mathcal{Q}}. \tag{1.73}$$

The quantity $\beta_{\mathcal{Q}}$ denotes the quantum bound of the Bell expression and is defined as

$$\beta_{\mathcal{Q}} \coloneqq \max_{\boldsymbol{P}\in\mathcal{Q}} \boldsymbol{\beta}\cdot\boldsymbol{P} = \max_{\hat{S}} \|\hat{S}\|_{\sup}, \tag{1.74}$$

where again $\|S\|_{\sup}$ is the spectral norm of the associated Bell operator

$$\hat{S} \coloneqq \sum_{a,b,x,y} \beta_{a,b,x,y}\hat{\Pi}^A_{a|x}\hat{\Pi}^B_{b|y}. \tag{1.75}$$

The maximization is performed over all possible quantum realizations of the Bell operator $\hat{S}$, i.e., over all valid measurement operators $\{\hat{\Pi}^A_{a|x}\}$, $\{\hat{\Pi}^B_{b|y}\}$ acting on a bipartite Hilbert space $\mathcal{H}_A\otimes\mathcal{H}_B$.

A straightforward approach to estimate $\beta_{\mathcal{Q}}$ is to fix a Hilbert space of bounded dimension and optimize the spectral norm $\|\hat{S}\|_{\sup}$ numerically over a parametrized family of Bell operators. However, this method generally provides only a lower bound on $\beta_{\mathcal{Q}}$, since:

- it may miss the global optimum;
- there is no guarantee that the optimal violation can be realized in the chosen Hilbert space.

Despite these limitations, such lower bounds are useful for constructing explicit quantum violations of classical Bell inequalities.

1.4.5.2 The NPA hierarchy

The previous discussion focused on techniques for establishing lower bounds on $\beta_{\mathcal{Q}}$. Now, we shift our attention to methods for determining upper bounds.

For the purpose of device-independent certification, it is advantageous to define *supersets* or *relaxations* of $\mathcal{Q}$. This way, any property proven for one of these supersets automatically applies to the entire quantum set. This section introduces a systematic construction of such supersets, characterized by compact and computationally efficient conditions for membership. These conditions were systematically introduced by Navascués, Pironio, and Acín [49, 50] and are thus known as *NPA relaxations*.

The foundation of NPA relaxations lies in the following observation:

Lemma 1.1. *Let $\mathcal{O} \coloneqq \{\hat{O}_1,\dots,\hat{O}_n\}$ be a set of linear operators, $\hat{O}_i \in L(\mathcal{H}_{AB})$. Then, for any state $|\psi\rangle$, the Hermitian matrix $\Gamma(|\psi\rangle,\mathcal{O})$ of entries*

$$\Gamma_{ij} \coloneqq \langle\psi|\hat{O}_i^\dagger\hat{O}_j|\psi\rangle \tag{1.76}$$

is positive semidefinite.

Proof. For any vector $\boldsymbol{v}\in\mathbb{C}^n$ it holds

$$\boldsymbol{v}^\dagger\Gamma\boldsymbol{v} = \langle\psi|\left(\sum_i v_i^*\hat{O}_i^\dagger\right)\left(\sum_j v_j\hat{O}_j\right)|\psi\rangle \succeq 0, \tag{1.77}$$

because any operator of the form $O^\dagger O$ is positive semidefinite. □

Consider a behavior $\boldsymbol{P} \in \mathcal{Q}$. We can construct the matrix $\Gamma_1(|\psi\rangle, \mathcal{O}_1)$ using the state $|\psi\rangle$ and the collection of $2mk$ projectors:

$$\mathcal{O}_1 \coloneqq \Big\{\mathbb{1}, \{\hat{\Pi}^A_{a|x}\}, \{\hat{\Pi}^B_{b|y}\}\Big\}. \tag{1.78}$$

If the specific state and measurements were known, all entries of Γ_1 could be directly computed. However, in a device-independent scenario, we only assume knowledge of the statistics (the state and projectors are unknown a priori). Therefore, certain entries of Γ_1 are directly determined by the observed behavior. In fact, it results

$$\langle\psi|\,\hat{\Pi}^A_{a|x}\hat{\Pi}^B_{b|y}\,|\psi\rangle \equiv P(a,b|x,y), \qquad \langle\psi|\,\hat{\Pi}^A_{a|x}\,|\psi\rangle \equiv P(a|x), \qquad \langle\psi|\,\hat{\Pi}^B_{b|y}\,|\psi\rangle \equiv P(b|y). \tag{1.79}$$

Other entries, which involve two operators from Alice's side or two from Bob's, do not correspond to directly observed quantities in a Bell test. Nonetheless, Lemma 1.1 guarantees that these entries can be completed such that $\Gamma_1 \succeq 0$. Conversely, if, after populating the known entries with the observed behavior, it becomes impossible to fill the remaining entries to satisfy $\Gamma_1 \succeq 0$, then it can be conclusively stated that $\boldsymbol{P} \notin \mathcal{Q}$.

This describes a necessary condition for a behavior $\boldsymbol{P}$ to be part of the quantum set $\mathcal{Q}$. For some specific subsets of the quantum set, this condition is also sufficient. However, generally, it represents only the first step in the NPA hierarchy of criteria. The set of behaviors that satisfy the condition $\Gamma_1 \succeq 0$ is denoted as $\mathcal{Q}_1$.

For the subsequent level of the hierarchy, denoted $\nu = 2$, the matrix Γ_2 is constructed using an expanded collection of operators:

$$\mathcal{O}_2 = \mathcal{O}_1 \cup \Big\{\{\hat{\Pi}^A_{a|x}\hat{\Pi}^A_{a'|x'}\}, \{\hat{\Pi}^B_{b|y}\hat{\Pi}^B_{b'|y'}\}, \{\hat{\Pi}^A_{a|x}\hat{\Pi}^B_{b|y}\}\Big\}. \tag{1.80}$$

It's understood that all $3(mk)^2$ newly formed operators are added to $\mathcal{O}_1$. The set of behaviors for which $\Gamma_2 \succeq 0$ is denoted $\mathcal{Q}_2$. The positivity of Γ_2 is a necessary condition for $\boldsymbol{P} \in \mathcal{Q}$, providing a tighter constraint than Γ_1 positivity, though still not generally sufficient. Notably, none of the entries involving operators like $\hat{\Pi}^A_{a|x}\hat{\Pi}^A_{a'|x'}$ are determined by the observed behavior for $a \neq a'$. Nevertheless, these terms impose specific structural requirements on the matrix. For instance:

$$\langle\psi|\,\hat{\Pi}^A_{a|x}\hat{\Pi}^A_{a'|x'}\hat{\Pi}^A_{a''|x'}\,|\psi\rangle = \langle\psi|\,\hat{\Pi}^A_{a|x}\hat{\Pi}^A_{a'|x'}\,|\psi\rangle\,\delta_{a'a''}. \tag{1.81}$$

It can be inferred that the n-th step of this hierarchy involves constructing the matrix Γ_n from the collection of all products of n or fewer projectors. The set of behaviors for which $\Gamma_n \succeq 0$ is denoted $\mathcal{Q}_n$. This construction yields a nested sequence of sets:

$$\mathcal{Q}_1 \supseteq \mathcal{Q}_2 \supseteq \mathcal{Q}_3 \supseteq \cdots \supseteq \mathcal{Q}_n \supseteq \mathcal{Q}'. \tag{1.82}$$

NPA demonstrated the convergence of this hierarchy, proving that $\lim_{n\to\infty} \mathcal{Q}_n = \mathcal{Q}'$. It's important to note that this convergence is established for $\mathcal{Q}'$, not $\mathcal{Q}$. In many practical cases, the quantum result is achieved at a finite level n of the hierarchy.

Remark 1.8. Note that:

- The inherent flexibility in choosing the operator collection $\mathcal{O}$ within Lemma 1.1 is a key aspect of this hierarchy.

- Adding even a single operator to an existing $\mathcal{O}$ can lead to a tighter bound. The notation $\mathcal{Q}_{\mathcal{O}}$ is used to denote the set of behaviors where $\Gamma_{\mathcal{O}} \succeq 0$ for a given collection of operators $\mathcal{O}$.

The fact that the NPA conditions are expressed as semidefinite requirements ($\Gamma \succeq 0$) makes them particularly well-suited for use in *semidefinite programs (SDPs)*. SDPs are the next level of complexity in convex optimization after linear programs. Similar to linear programs, for any given SDP, a dual SDP can be algorithmically constructed, providing simultaneous lower and upper bounds for the desired solution.

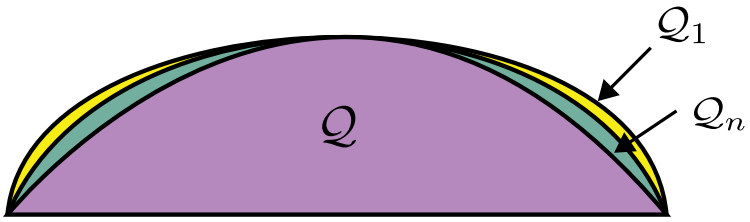

Figure 1.5: Hierarchy of sets $\mathcal{Q}_n$ generated by the NPA recipe. Each set in the hierarchy approximates better the set of quantum correlations $\mathcal{Q}$. From [19].

The membership problem. Let us describe how the NPA techniques are employed to address the problem of quantum set membership, i.e., determining whether a given behavior $\boldsymbol{P}$ belongs to $\mathcal{Q}$. For any chosen set of operators $\mathcal{O}$, the membership problem can be relaxed to the following optimization:

$$\begin{aligned} \Lambda := \max \quad & \lambda \\ \text{s. t.} \quad & \Gamma - \lambda \mathbb{1} \succeq 0 \\ & \Gamma_{(x,a)} = P(a|x) \\ & \Gamma_{(y,b)} = P(b|y) \\ & \Gamma_{(x,a)(y,b)} = P(a,b|x,y) \end{aligned} \tag{1.83}$$

Here, $\Gamma_{(x,a)}$, $\Gamma_{(y,b)}$, and $\Gamma_{(x,a)(y,b)}$ refers to specific entries of Γ corresponding to the marginal probabilities $P(a|x)$, $P(b|y)$, and the joint probabilities $P(a,b|x,y)$, respectively. The optimization variables are $\lambda \in \mathbb{R}$ and all other entries of the matrix Γ.

Remark 1.9. These additional entries can be restricted to be real without loss of generality: if a solution Γ_{sol} is found to be positive semidefinite and satisfies the constraints, potentially with complex entries, then its complex conjugate Γ^*_{sol} also satisfies the constraints. Consequently, their real part, $\Gamma = \frac{1}{2}(\Gamma_{\text{sol}} + \Gamma^*_{\text{sol}})$, is also positive semidefinite and satisfies all the imposed constraints.

If the solution to this optimization, Λ, is found to be negative ($\Lambda < 0$), it implies that no positive semidefinite matrix Γ can be constructed while satisfying the given constraints. In such a scenario, it can be definitively concluded that the behavior $\boldsymbol{P} \notin \mathcal{Q}$. Conversely, if a $\Gamma \succeq 0$ is found, it indicates that $\boldsymbol{P} \in \mathcal{Q}_{\mathcal{O}}$. Since $\mathcal{Q}_{\mathcal{O}}$ is a superset of $\mathcal{Q}$ (i.e., $\mathcal{Q}_{\mathcal{O}} \supset \mathcal{Q}$), this condition alone does not definitively prove that $\boldsymbol{P} \in \mathcal{Q}$. The tightness of this assessment can be improved by selecting a larger collection of operators $\mathcal{O}$. The only conclusive method to

formally demonstrate that $\boldsymbol{P} \in \mathcal{Q}$ is to explicitly provide a quantum realization of the behavior, specifying the underlying quantum state and measurement projectors (of course, knowing such a realization, the semidefinite program becomes redundant for membership verification).

Maximal violation of a Bell inequality. Next, we consider the problem of determining the maximal violation of a Bell inequality within quantum theory, which represents one of the earliest applications of semidefinite programs in the study of nonlocality, as highlighted in [69]. The goal is to find

$$\begin{aligned} \beta_{\mathcal{Q}} \coloneqq \max \quad & \beta(\boldsymbol{P}) \\ \text{s. t.} \quad & \boldsymbol{P} \in \mathcal{Q}. \end{aligned} \tag{1.84}$$

Unless analytical arguments are available, such as those used to derive the Tsirelson bound, this problem cannot be solved efficiently. At best, one can employ heuristic optimization strategies over a restricted class of states and measurements, effectively searching over a subset of $\mathcal{Q}$. Due to the heuristic nature of such optimizations, possibly converging to a local maximum, the obtained outcome will satisfy $\beta_{\text{heur}} \leq \beta_{\mathcal{Q}}$.

The NPA techniques provide a systematic way to relax the optimization problem stated in (1.84) into the SDP:

$$\begin{aligned} \beta_{\Gamma} \coloneqq \max \quad & \beta(\boldsymbol{P}) \\ \text{s. t.} \quad & \Gamma(\boldsymbol{P}) \succeq 0. \end{aligned} \tag{1.85}$$

In this SDP, all entries of the matrix Γ are treated as optimization variables. To ensure a non-trivial solution, all probabilities $P(a, b|x, y)$ that appear in the Bell expression $\beta(\boldsymbol{P})$ must correspond to entries within Γ. This requirement implicitly defines the minimal collection of operators $\mathcal{O}$ necessary for the problem. The SDP typically yields both an upper and a lower bound for $\beta(\boldsymbol{P})$, which usually align within numerical precision. Crucially, by optimizing over $\mathcal{Q}_{\mathcal{O}} \supseteq \mathcal{Q}$, we are guaranteed that the obtained value β_{Γ} satisfies $\beta_{\Gamma} \geq \beta_{\mathcal{Q}}$. Increasing the size of the operator collection $\mathcal{O}$ will generally lead to a β_{Γ} that converges more closely to the true $\beta_{\mathcal{Q}}$. If β_{Γ} is found to be equal to an explicitly achievable quantum value (either a numerical β_{heur} from a specific quantum realization or by an exact computation), then we have successfully determined $\beta_{\mathcal{Q}}$.

Dual approach. An alternative, dual approach to this hierarchy was developed in [70], drawing upon concepts from *sum-of-squares (SOS) polynomial optimization* [71]. For a given Bell operator $\hat{S}$, the quantum bound $\beta_{\mathcal{Q}}$ satisfies the condition

$$\hat{\xi} \coloneqq \beta_{\mathcal{Q}} \mathbb{1} - \hat{S} \succeq 0, \tag{1.86}$$

if and only if $\hat{\xi}$ is a positive operator. In the SOS methodology, the objective is to express $\hat{\xi}$ as a sum of squares of polynomials in the measurement operators. By constraining these polynomials to a maximum degree of 2ℓ, this problem transforms into an SDP, yielding progressively tighter bounds as ℓ increases. This SOS hierarchy converges to the quantum set in the limit $\ell \to \infty$.

For certain Bell expressions, finite convergence can be established. When this occurs, it becomes possible to reconstruct a quantum realization of the optimal bound. That is, to explicitly determine the state $|\psi\rangle$ and the measurements $\{\hat{\Pi}^A_{a|x}\}$, $\{\hat{\Pi}^B_{b|y}\}$ that achieve $\beta_{\mathcal{Q}}$ [50].

Finally, the optimality at a finite level can be confirmed by comparing the upper bound derived from the NPA hierarchy with a lower bound obtained through direct optimization over Bell operators, as previously discussed.

1.4.6 Geometry of correlations in the CHSH scenario

Already in the minimal Bell scenario ($n = k = m = 2$), the structure of the sets of correlations exhibits a remarkably rich geometry. The local set $\mathcal{L}$ consists of the local deterministic behaviors, forming a polytope with 16 vertices (corresponding to deterministic assignments of outputs to inputs) and 24 facets. Among these, 16 correspond to trivial positivity constraints, while the remaining 8 correspond to violations of the CHSH inequality, up to relabeling of inputs and outputs.

The no-signaling set $\mathcal{NS}$, defined by the requirement that neither party can influence the marginal statistics of the other through their choice of input, also forms a polytope, but with a different geometric structure. It has 24 extremal points: the 16 local deterministic ones, and 8 additional nonlocal vertices, all equivalent under symmetries to the PR box (1.8). This behavior maximally violates the CHSH inequality up to the algebraic maximum

$$\boldsymbol{\beta}_{\text{CHSH}} \cdot \boldsymbol{P}_{\text{PR}} = 4. \tag{1.87}$$

This value clearly exceeds the Tsirelson bound $2\sqrt{2}$.

The CHSH expression thus delineates three distinct regions in correlation space:

$$\boldsymbol{\beta}_{\text{CHSH}} \cdot \boldsymbol{P} \leq \begin{cases} 4 & \text{if } \boldsymbol{P} \in \mathcal{NS}, \\ 2\sqrt{2} & \text{if } \boldsymbol{P} \in \mathcal{Q}, \\ 2 & \text{if } \boldsymbol{P} \in \mathcal{L}, \end{cases} \tag{1.88}$$

providing a rare example where a single Bell inequality fully separates three operationally distinct sets of interest. This tri-level structure is visually represented in simplified two-dimensional slices, as in Fig. 1.6.

While this coarse-grained representation of the geometry is helpful, the actual structure of the quantum set $\mathcal{Q}$ is far more complex. A fundamental result in this direction was the discovery that the quantum set $\mathcal{Q}$ contains extremal points that are *non-exposed*, i.e., points which cannot be obtained as the unique maximizers of any Bell inequality over $\mathcal{Q}$ [46]. This contrasts sharply with the classical and no-signaling sets, where every extremal point is exposed. The existence of non-exposed points implies that the boundary of $\mathcal{Q}$ is not fully accessible via Bell inequality violations: some quantum behaviors are extremal yet remain invisible to any linear optimization procedure over the set. This fact underscores a fundamental limitation of the Bell inequality framework in fully characterizing quantum correlations.

Additional structural richness is revealed by the existence of so-called *quantum voids*—open regions of the boundary of the no-signaling polytope $\mathcal{NS}$ that remain entirely inaccessible to quantum correlations [72]. Although the quantum set $\mathcal{Q}$ is dense in the interior of $\mathcal{NS}$, meaning that quantum correlations can approximate any interior no-signaling point arbitrarily well, this is not the case for the full boundary. In particular, there exist boundary faces of $\mathcal{NS}$ of high dimension—up to six in the CHSH scenario—that are completely disjoint from $\mathcal{Q}$. These quantum voids underscore the sparsity of the quantum set relative to $\mathcal{NS}$, and illustrate that quantum theory does not saturate all the physically allowed no-signaling correlations, even in the limit.

This feature challenges the common intuition that quantum mechanics nearly saturates the space of physically consistent probabilistic theories compatible with relativistic causality.

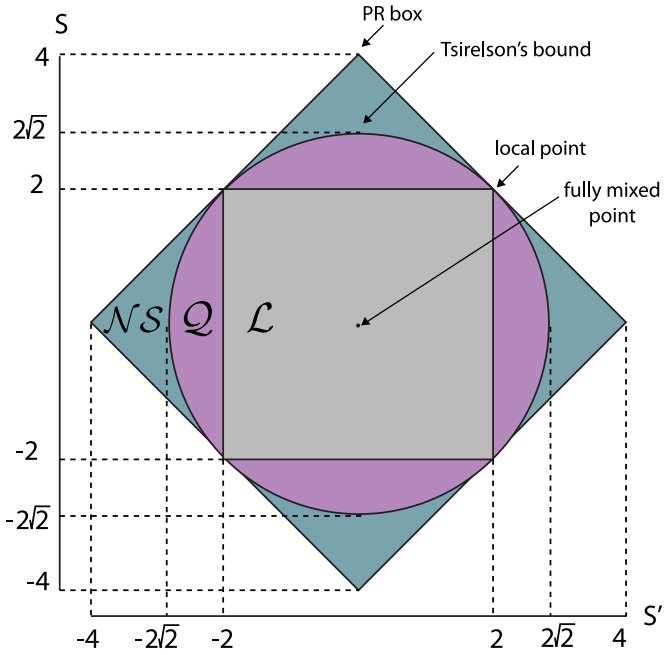

Figure 1.6: A two-dimensional section of the no-signaling polytope in the CHSH scenario ($m = k = 2$). The vertical axis represents the CHSH value $S := \langle \hat{S}_{\text{CHSH}} \rangle$, while the horizontal axis represents a symmetry of the CHSH expression S'. Local correlations satisfy $|S| \leq 2$ and $|S'| \leq 2$. The PR box achieves the algebraic maximum $S = 4$, while Tsirelson's bound corresponds to $S = 2\sqrt{2}$, the maximum achievable quantum value. From [19].

Surprisingly, despite the general curvature of $\mathcal{Q}$, recent work has shown that some of its facets are flat and shared with the no-signaling polytope [47]. That is, there exist supporting hyperplanes that simultaneously define a face of both $\mathcal{Q}$ and $\mathcal{NS}$, indicating that the boundary of the quantum set aligns exactly with the no-signaling polytope in certain directions. These shared facets demonstrate that not all of the boundary of $\mathcal{Q}$ is curved, and some portions admit a polyhedral structure —another departure from the naive image of a uniformly smooth body.

Complementary to the phenomenon of non-exposed extremal points, recent results have also uncovered "pointy" extremal points of $\mathcal{Q}$, which lie at the intersection of several tight Bell inequalities [52]. These points locally resemble vertices of a polytope, and although $\mathcal{Q}$ itself is not a polytope, these features provide local structures where quantum correlations are highly constrained. This polyhedral-like structure implies that some quantum behaviors can, after all, be fully characterized by a finite set of Bell inequalities—suggesting that the curved geometry of $\mathcal{Q}$ coexists with strongly "cornered" extremal features.

Understanding this geometry is not merely an academic exercise: it is crucial to identifying the principles that distinguish quantum theory from more general no-signaling theories. The subtle interplay between exposed and non-exposed points, between flat and curved facets, and between accessible and inaccessible regions of $\mathcal{Q}$, suggests that any candidate physical principle aiming to recover quantum theory must be sensitive not only to the set's algebraic constraints, but also to its geometric subtleties.

1.5 Device-independent self-testing

In quantum information theory, a central goal is to verify whether a physical system possesses certain properties (such as entanglement, measurement incompatibility, or high-dimensional structure) based solely on observed data. This task falls under the broad umbrella of *quantum property testing* [73], which aims to infer meaningful features of a quantum system while making minimal assumptions about its internal structure.

In many scenarios, we must further consider the constraint that the quantum devices themselves may be uncharacterized or untrusted. This is a consequence of the fact that quantum systems are typically probed using other quantum devices, which may themselves be flawed or adversarial. The device-independent approach is urged wherein one seeks to certify physical properties using only classical inputs and outputs, without relying on any assumptions about the internal workings of the devices. Such methods are especially relevant in cryptographic and adversarial contexts, where robust certifications are essential [35].

A key tool in this framework is the violation of a Bell inequality, which enables the *device-independent certification of entanglement* [74, 75]. If Alice and Bob observe correlations $\boldsymbol{P}$ such that

$$\boldsymbol{\beta} \cdot \boldsymbol{P} > \beta_{\mathcal{L}},$$

then the source must necessarily produce an entangled state. This conclusion depends solely on the observed statistics $\{P(a, b|x, y)\}$, not on the internal details of the devices. This is exactly the content of Remark 1.6: entanglement is a necessary ingredient for generating nonlocal correlations. Bell inequality violations thus act as entanglement detectors that are inherently robust to implementation imperfections. As long as a violation is observed, entanglement is guaranteed, independently of how the system is realized.

However, such certification does not reveal the specific form of the underlying quantum state or measurements: it merely confirms that the observed correlations cannot arise from any separable system. As the study of Bell nonlocality developed, it became evident [57] that certain nonlocal correlations not only require entanglement and measurement incompatibility but can only be realized through specific entangled states and measurements.

This observation leads to a stronger goal: instead of merely detecting entanglement, Alice and Bob aim to identify the specific entangled state produced by their source. Even without direct access to the physical system, such identification becomes possible when the observed statistics achieve the maximal quantum violation of a Bell inequality. In these rare cases, the observed data uniquely determine the underlying state and measurements (up to local isometries), enabling what is known as a *device-independent self-test*.

The field of *self-testing* [37] emerged from this idea. Pioneered by Mayers and Yao [76], it studies the conditions under which, from observed statistics alone, an entirely classical verifier can certify that untrusted quantum devices share a particular entangled state. Self-testing thus represents the strongest form of device-independent certification and plays a central role in secure quantum communication, randomness generation, and delegated quantum computation.

Remark 1.10. In a scenario involving only a single device, where one has access only to the conditional probabilities $P(a|x)$ and state and measurements are jointly unknown, it is fundamentally impossible to determine the underlying quantum state $|\psi\rangle$. Indeed, one cannot

exclude the possibility that the device is simply a preprogrammed classical system designed to reproduce the observed statistics $P(a|x)$. In the device-independent framework, which makes no assumptions about the internal workings of the device, such classical simulations cannot be ruled out. Consequently, no genuinely quantum properties of $|\psi\rangle$ can be certified in a single-device scenario. To overcome this limitation, it is essential to consider a multipartite setting. It is precisely the phenomenon of Bell nonlocality—i.e., the inability of local classical models to reproduce certain quantum correlations—that enables the certification of non-classical states in a device-independent manner.

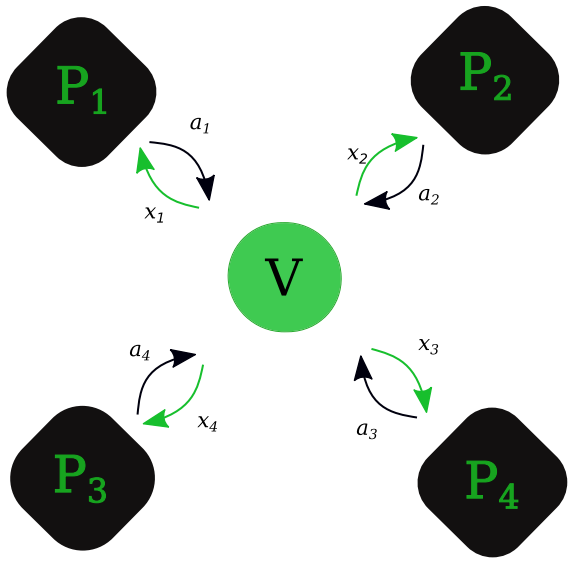

Figure 1.7: In a self-testing scenario, a classical verifier V seeks to certify a specific quantum property of a shared resource held by non-communicating provers P_i. The verifier provides each prover with classical inputs x_i, to which the provers respond with classical outputs a_i. If the resulting correlations between the outputs exhibit nonlocality, the verifier can infer nontrivial information about the underlying quantum system—such as the presence of entanglement or the structure of the implemented measurements. From [77].

Remark 1.11. In realistic experiments, Alice and Bob will not observe exact maximal violations due to noise and finite statistics. To make self-testing practically applicable, two additional components are required:

1. *Robust self-testing* techniques, which quantify how close the implemented state and measurements are to the ideal ones, based on the level of observed Bell violation.

2. Statistical tools, such as *Chernoff bounds*, which provide confidence intervals and error estimates for the inferred probabilities.

Together, these allow self-testing to function as a practical protocol for device-independent certification of quantum states and measurements.

Self-testing also has deep theoretical significance. While the Born rule defines a forward map from quantum states and measurements to probability distributions, this map is generally not invertible: different physical realizations can produce the same statistics. Self-testing identifies the exceptional cases where an inverse exists (up to local isometries), thereby characterizing extremal quantum correlations with unique physical realizations. These insights have proven instrumental in advancing our understanding of the geometry and structure of the quantum set.

1.5.1 Introduction to self-testing

Let us suppose we are interested in certifying a specific pure bipartite quantum state $|\psi'\rangle_{A'B'}$ along with a set of projective measurements $\{\hat{\Pi}'^{A}_{a|x}\}$ and $\{\hat{\Pi}'^{B}_{b|y}\}$. These are referred to as the *reference state* and *reference measurements*. On the other hand, the actual experimental implementation involves a quantum state $\hat{\rho}_{AB}$ and measurement operators $\{\hat{\Pi}^{A}_{a|x}\}$ and $\{\hat{\Pi}^{B}_{b|y}\}$, collectively called the *physical state* and *physical measurements*.

Accordingly, we define the *reference experiment* as

$$\left(|\psi'\rangle_{A'B'}, \{\hat{\Pi}'^{A}_{a|x}\}, \{\hat{\Pi}'^{B}_{b|y}\}\right),$$

and the *physical experiment* as

$$\left(\hat{\rho}_{AB}, \{\hat{\Pi}^{A}_{a|x}\}, \{\hat{\Pi}^{B}_{b|y}\}\right).$$

Remark 1.12. It is important to recognize that it is generally impossible to reconstruct the reference state and measurements uniquely from the observed statistics alone. Two main reasons explain this:

- *Local unitaries.* Due to the invariance of measurement statistics under local unitary transformations, one can reproduce the same correlations by applying local unitaries $\hat{U}$ and $\hat{V}$ to the state and measurement operators:

  $$(\hat{U}\otimes\hat{V})\,|\psi'\rangle, \quad \{\hat{U}\hat{\Pi}'^{A}_{a|x}\hat{U}^\dagger\}, \quad \{\hat{V}\hat{\Pi}'^{B}_{b|y}\hat{V}^\dagger\}.$$

  As such, the state needs not be exactly $|\psi'\rangle$, but only locally equivalent to it.

- *Additional ancilla systems.* It is possible that the physical system includes extra degrees of freedom not involved in the measurement process. That is, one may instead have the state $|\psi'\rangle\otimes|\xi\rangle_C$ and measurement operators of the form:

  $$\{\hat{\Pi}'^{A}_{a|x}\otimes\mathbb{1}_C\}, \quad \{\hat{\Pi}'^{B}_{b|y}\otimes\mathbb{1}_C\}.$$

  These give rise to the same observable statistics as the original reference experiment.

To meaningfully compare the physical and reference experiments in a device-independent manner, we must define an appropriate equivalence between quantum states that accounts for both local unitaries and auxiliary systems. This is achieved via the notion of a local isometry.

Definition 1.6. An *isometry* is a linear map

$$\Phi\colon \mathcal{H}_{A_1}\to\mathcal{H}_{A_2}$$

that preserves inner products. In physical terms, an isometry can be viewed as the embedding of a quantum system into a larger Hilbert space. Such embeddings can typically be realized by adding an ancillary system and applying a unitary operation on the joint system, namely $\Phi[|\psi\rangle_{A_1}] = \hat{U}\,|\psi\rangle_{A_2}$.

Explicitly, for a pure state $|\psi\rangle_{A_1}$, the action of an isometry can be represented as

$$\Phi(|\psi\rangle_{A_1}) \coloneqq \hat{U}\,|\psi\rangle_{A_2}\,,$$

where $\hat{U}$ is a unitary acting on an extended space.

Definition 1.7. A *local isometry* is a tensor product of isometries acting independently on each subsystem:

$$\Phi_{A_1} \otimes \Phi_{B_1} \colon \mathcal{H}_{A_1} \otimes \mathcal{H}_{B_1} \to \mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2}.$$

Such transformations can be implemented using local quantum operations alone.

1.5.2 Self-testing of pure states

We are now equipped to define what it means to self-test a quantum state based on observed statistics alone.

Definition 1.8 (Self-testing of pure states). A set of correlations $P(a,b|x,y)$ *self-tests the state* $|\psi'\rangle_{A'B'}$ if for any state $\hat{\rho}_{AB}$ compatible with these statistics (for some choice of local measurements) and for any purification $|\psi\rangle_{ABP}$ of $\hat{\rho}_{AB}$ there exists a local isometry

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

and a state $|\xi\rangle_{\bar{A}\bar{B}P}$ such that

$$(\Phi_A \otimes \Phi_B \otimes \mathbb{1}_P)\,|\psi\rangle_{ABP} = |\psi'\rangle_{A'B'} \otimes |\xi\rangle_{\bar{A}\bar{B}P}\,.$$

1.5.3 Self-testing of measurements

While the self-testing of quantum states seeks to certify the underlying state solely from observed correlations, a natural extension is to ask whether the measurement operators themselves can also be certified in a device-independent manner.

Definition 1.9 (Self-testing of states and measurements). The correlations $P(a,b|x,y)$ self-test the realization

$$\left(|\psi'\rangle_{A'B'}\,, \{\hat{\Pi}'^{A}_{a|x}\}, \{\hat{\Pi}'^{B}_{b|y}\}\right)$$

if for any state and measurements

$$\left(\hat{\rho}_{AB}, \{\hat{\Pi}^{A}_{a|x}\}, \{\hat{\Pi}^{B}_{b|y}\}\right)$$

compatible with $P(a,b|x,y)$ and for any purification $|\psi\rangle_{ABP}$ of $\hat{\rho}_{AB}$ there exists a local isometry $\Phi_A \otimes \Phi_B$ such that

$$(\Phi_A \otimes \Phi_B \otimes \mathbb{1}_P)\left(\left[\hat{\Pi}^{A}_{a|x} \otimes \hat{\Pi}^{B}_{b|y} \otimes \mathbb{1}_P\right]|\psi\rangle_{ABP}\right) = \left[\left(\hat{\Pi}'^{A}_{a|x} \otimes \hat{\Pi}'^{B}_{b|y}\right)|\psi'\rangle_{A'B'}\right] \otimes |\xi\rangle_{\bar{A}\bar{B}P} \tag{1.89}$$

for all a, x, b, y and for some state $|\xi\rangle_{\bar{A}\bar{B}P}$.

Note that, as previously noticed in Remark 1.12, we are only be able to say something about the part of the measurement that acts on the support of the physical state, as any action outside this subspace is operationally irrelevant.

For completeness, we recall the notion of self-testing of measurements, as introduced in [37]:

Definition 1.10 (Self-testing of measurements). The correlations $P(a,b|x,y)$ *self-test the measurements* $\{\hat{\Pi}'_{a|x}\}$ for Alice if for any measurements $\{\hat{\Pi}_{a|x}\}$ and state $\hat{\rho}_{AB}$ compatible with $P(a,b|x,y)$ there exists a unitary operator $\hat{U}$ such that

$$\hat{U}_A \hat{\Omega}_{a|x} \hat{U}_A^\dagger = \hat{\Pi}'_{a|x} \otimes \mathbb{1}$$

for all a, x, where

$$\hat{\Omega}_{a|x} \coloneqq \hat{E}\hat{\Pi}_{a|x}\hat{E}$$

and $\hat{E}$ is the projector onto the support of the reduced state

$$\hat{\rho}_A = \mathrm{Tr}_B(\hat{\rho}_{AB}).$$

where $\mathrm{Tr}_B(\cdot)$ is the partial trace over B.[11]

It is worth noting that measurement self-testing can also be formulated in a way that does not explicitly involve the measured state, but rather focuses directly on the measurement operators themselves [78]. This alternative definition is arguably more natural and parallels the standard formulation of state self-testing.

Beyond the self-testing of states and measurements, recent developments have extended device-independent certification to include quantum operations and channels [79], as well as fundamental resources such as randomness [36] and shared secrecy [33, 80, 81]. This allows for the robust verification of coherent processes—such as storage, processing, and transmission of quantum information——and even the certification of quantum error-correcting codes [82], thereby completing the set of tools required to certify all building blocks of a quantum computer in a fully device-independent manner [79].

It has been shown that only extremal points of the quantum set $\mathcal{Q}$ can give rise to self-testing correlations [46]. While one might conjecture that every extremal quantum behavior in a given Bell scenario corresponds to a self-test of some quantum state, no general proof of this statement currently exists. It is important to note that different extremal behaviors may self-test the same state. For instance, the maximally entangled two-qubit state is self-tested not only by the maximal CHSH violation $\langle \hat{S} \rangle = 2\sqrt{2}$, but also by the Mayers-Yao correlations and, more broadly, by all nonlocal points lying on the so called *Tsirelson-Landau-Masanes (TLM) boundary* [83].

[11] Let $\mathcal{H}_A \otimes \mathcal{H}_B$ be a composite Hilbert space, and let $\{|\beta_i\rangle\}$ be an orthonormal basis for $\mathcal{H}_B$. Then, for any operator $\hat{O}_{AB} \in L(\mathcal{H}_A \otimes \mathcal{H}_B)$, the partial trace over B is defined by

$$\mathrm{Tr}_B(\hat{O}_{AB}) \coloneqq \sum_i \langle \beta_i | \hat{O}_{AB} | \beta_i \rangle .$$

An important consequence of this characterization is that mixed states cannot be self-tested. The behaviors they give rise to are not extremal in $\mathcal{Q}$, as they can be decomposed into convex combinations of behaviors arising from pure states measured in the same way. This fact retrospectively justifies the standard practice of formulating self-testing definitions in terms of pure states.

It is known that every pure entangled bipartite state admits a self-testing protocol [84]. In the multipartite setting, numerous examples of self-testable states are also established, including GHZ states, all graph states, the three-qubit $|W\rangle$ state, and various generalizations [37]. However, in the general multipartite case, the notion of self-testing must be treated with additional care: this is because there exist entangled states that are not equivalent under local unitaries, yet give rise to identical state-behavior pairs in a given Bell scenario, making them operationally indistinguishable through device-independent means.

It is known that every pure entangled bipartite state admits a self-testing protocol [84]. In the multipartite setting, numerous examples of self-testable states have also been established, including GHZ states, all graph states, the three-qubit $|W\rangle$ state, and various generalizations [37]. However, in the general multipartite case, the notion of self-testing must be treated with additional care: there exist entangled states that are not equivalent under local unitaries, yet give rise to identical state-behavior pairs in a given Bell scenario, rendering them operationally indistinguishable through device-independent means. Despite this limitation, recent work has shown that, up to this additional degree of freedom, all multi-qubit states are self-testable in principle, thereby significantly extending the reach of device-independent certification in the multipartite regime [85].

1.5.4 Self-testing via a Bell inequality

It is known that only those correlations that are extremal points of the quantum set of correlations and are achievable with finite dimensional quantum systems can be used to self-test both a state and measurements [46]. Such points can often be witnessed by the maximum violation of some Bell inequality over the set of quantum correlations. As a result, one often does not need the full set of probabilities $P(a, b|x, y)$ in order to prove self-testing statements; the maximum quantum violation of a Bell inequality may already imply the existence of the desired isometry. One can thus consider self-testing relative to a Bell inequality by replacing the observation of the correlations by the value of a Bell inequality $\beta(\boldsymbol{P})$ in the previous definitions. Many of the well known Bell inequalities, such as CHSH have been used to this effect.

In this light one might ask if the maximal violation of every nontrivial Bell inequality, i.e. one which can be violated in quantum theory, is also a self-test of some entangled state. Or even more generally, do all extremal points of the set of quantum correlations self-test some state? These questions are examined in [46], where it was shown that the relation between self-testing, maximisers of non-trivial Bell inequalities and the boundary of the quantum set is not as simple as one might hope for.

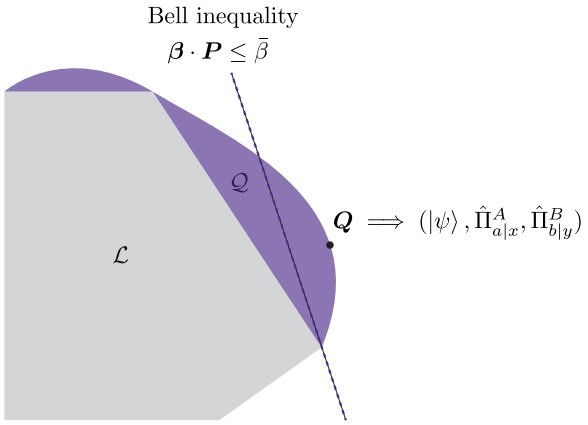

Figure 1.8: Pictorial representation of self-testing. The grey region represents the local set $\mathcal{L}$. A Bell inequality defines a boundary (dotted line) such that all local correlations lie within the half-space including $\mathcal{L}$. Quantum correlations, produced by entangled states, can go beyond this local region, as shown by points in the purple area. Certain extremal quantum correlations (e.g., $\boldsymbol{Q}$) that maximally violate a Bell inequality admit, up to local transformations, a unique realization in terms of a specific state and measurements. Self-testing consists in identifying such correlations and proving this uniqueness. From [37].

1.5.5 Self-testing proofs

The maximally entangled two-qubit state

$$\left|\phi^+\right\rangle := \tfrac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

plays a central role in quantum information theory, underpinning protocols such as quantum teleportation and quantum key distribution (see e.g., [33]). In this section, we present a formal self-testing proof for this state and a set of locally anticommuting observables, within the CHSH scenario. Many techniques used for self-testing more complex quantum systems can be viewed as generalizations of this foundational example.

We work in the CHSH scenario, where two parties each choose between two binary observables with outcomes ± 1. Let $\hat{A}_0, \hat{A}_1$ and $\hat{B}_0, \hat{B}_1$ denote Alice's and Bob's observables, respectively. It is convenient to adopt the observable (rather than projector) formalism here. We follow the self-testing definition given in Definition 1.8, and assume that the physical state $\hat{\rho}_{AB}$ is purified as $|\psi\rangle_{ABP}$, where the measurements act trivially on the purification space $\mathcal{H}_P$. For brevity, we omit identity operators on $\mathcal{H}_P$ in expressions such as $\hat{A}_x |\psi\rangle \equiv (\hat{A}_x \otimes \mathbb{1}_P) |\psi\rangle$.

We focus on behaviors that achieve the Tsirelson bound of the CHSH inequality,

$$\boldsymbol{\beta}_{\text{CHSH}} \cdot \boldsymbol{P} = \langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle = 2\sqrt{2}.$$

As previously mentioned (see Sec. 1.4.2), the canonical realization attaining this value consists of the state $\left|\phi^+\right\rangle$, with Alice measuring $\hat{A}_0 = \sigma_x, \hat{A}_1 = \sigma_z$, and Bob measuring

$$\hat{B}_0 = \frac{\sigma_z + \sigma_x}{\sqrt{2}}, \qquad \hat{B}_1 = \frac{\sigma_x - \sigma_z}{\sqrt{2}}.$$

This setup yields the maximum quantum value $\beta_{\mathrm{CHSH}} = 2\sqrt{2}$. It was shown early on [57, 86–88] that any other quantum realization reaching this value must be locally equivalent to the above. This constitutes a self-testing statement.

Proposition 1.2 (Self-testing from maximal CHSH violation). *Let $|\psi\rangle_{ABP}$ be a purification of a bipartite state on which ±1-valued observables $\hat{A}_0, \hat{A}_1$ and $\hat{B}_0, \hat{B}_1$ act, such that the CHSH value satisfies*

$$\langle\psi|\,\hat{S}_{\mathrm{CHSH}}\,|\psi\rangle = 2\sqrt{2}.$$

Then there exists a local isometry $\Phi := \Phi_A \otimes \Phi_B$ and an auxiliary state $|\xi\rangle_{\bar{A}\bar{B}P}$ such that

$$\Phi\,|\psi\rangle_{ABP} = \left|\phi^+\right\rangle_{A'B'} \otimes |\xi\rangle_{\bar{A}\bar{B}P}\,,$$

and

$$\begin{aligned}
\Phi(\hat{A}_0\,|\psi\rangle) &= \left(\frac{\sigma_x+\sigma_z}{\sqrt{2}} \otimes \mathbb{1}\right)\left|\phi^+\right\rangle \otimes |\xi\rangle\,,\\
\Phi(\hat{A}_1\,|\psi\rangle) &= \left(\frac{-\sigma_x+\sigma_z}{\sqrt{2}} \otimes \mathbb{1}\right)\left|\phi^+\right\rangle \otimes |\xi\rangle\,,\\
\Phi(\hat{B}_0\,|\psi\rangle) &= (\mathbb{1} \otimes \sigma_z)\left|\phi^+\right\rangle \otimes |\xi\rangle\,,\\
\Phi(\hat{B}_1\,|\psi\rangle) &= (\mathbb{1} \otimes \sigma_x)\left|\phi^+\right\rangle \otimes |\xi\rangle\,.
\end{aligned}$$

Proof. The key to the proof is to exploit the structure of the CHSH operator:

$$\hat{S}_{\mathrm{CHSH}} = \hat{A}_0 \otimes (\hat{B}_0 + \hat{B}_1) + \hat{A}_1 \otimes (\hat{B}_0 - \hat{B}_1),$$

and consider the shifted operator

$$2\sqrt{2}\,\mathbb{1} - \hat{S}_{\mathrm{CHSH}}.$$

Using the identities $\hat{A}_i^2 = \hat{B}_j^2 = \mathbb{1}$ and Hermiticity of the observables, one can write this as a sum-of-squares (SOS) decomposition:

$$2\sqrt{2}\mathbb{1} - \hat{S}_{\mathrm{CHSH}} = \frac{1}{\sqrt{2}}\left[\left(\frac{\hat{A}_0+\hat{A}_1}{\sqrt{2}} - \hat{B}_0\right)^2 + \left(\frac{\hat{A}_0-\hat{A}_1}{\sqrt{2}} - \hat{B}_1\right)^2\right].$$

Thus, for any state $|\psi\rangle$ saturating Tsirelson's bound, both square terms must vanish:

$$\left(\frac{\hat{A}_0 \pm \hat{A}_1}{\sqrt{2}} - \hat{B}_{0/1}\right)|\psi\rangle = 0,$$

which implies

$$\hat{B}_0\,|\psi\rangle = \frac{\hat{A}_0+\hat{A}_1}{\sqrt{2}}\,|\psi\rangle\,, \qquad \hat{B}_1\,|\psi\rangle = \frac{\hat{A}_0-\hat{A}_1}{\sqrt{2}}\,|\psi\rangle\,.$$

From these relations, one computes

$$\{\hat{B}_0, \hat{B}_1\}\,|\psi\rangle = 0, \tag{1.90}$$

i.e., Bob's observables anticommute on the support of $|\psi\rangle$. Similarly, it follows that

$$\{\hat{A}_0, \hat{A}_1\} |\psi\rangle = 0. \tag{1.91}$$

These anticommutation relations form the algebraic foundation of the proof, and enable the construction of a local isometry that maps the physical state and measurements to the reference realization.

We now prove a formal self-testing statement for the state $|\phi^+\rangle$, following Definition 1.8. This involves constructing an explicit isometry Φ such that

$$\Phi(|\psi\rangle_{ABP}) = \left|\phi^+\right\rangle_{A'B'} \otimes |\xi\rangle_{\bar{A}\bar{B}P},$$

for some auxiliary state $|\xi\rangle$. In most self-testing proofs, including the CHSH case, the isometry takes the form of a partial *swap gate*, depicted in Fig. 1.9.

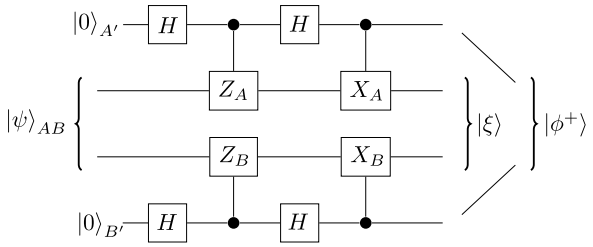

Figure 1.9: Partial swap isometry used to self-test the maximally entangled two-qubit state. Hadamard gates (H) act on ancilla qubits. After the circuit is applied, the entangled state is coherently extracted from $|\psi\rangle$ onto the ancillas. From [37].

The key idea is to define observables that emulate Pauli operators on the support of $|\psi\rangle$. Given that Alice's and Bob's observables anticommute on the support of $|\psi\rangle$, we define:

$$\hat{Z}_A = \frac{\hat{A}_0 + \hat{A}_1}{\sqrt{2}}, \quad \hat{X}_A = \frac{\hat{A}_0 - \hat{A}_1}{\sqrt{2}}, \qquad \hat{Z}_B = \hat{B}_0, \quad \hat{X}_B = \hat{B}_1. \tag{1.92}$$

Using the relations (1.90) and (1.91), one shows that

$$\{\hat{Z}_A, \hat{X}_A\} |\psi\rangle = 0, \qquad \{\hat{Z}_B, \hat{X}_B\} |\psi\rangle = 0, \qquad \hat{Z}_A |\psi\rangle = \hat{Z}_B |\psi\rangle, \quad \hat{X}_A |\psi\rangle = \hat{X}_B |\psi\rangle.$$

To ensure $\hat{Z}_A$ and $\hat{X}_A$ are unitary, we apply a standard regularisation procedure: zero eigenvalues are lifted to 1 and the resulting operator is normalized. The regularised operators still act on $|\psi\rangle$ as in (1.92), so we continue using the same notation.

The swap isometry acts on $|\psi\rangle$ as follows:

$$\Phi(|\psi\rangle) \coloneqq \sum_{i,j\in\{0,1\}} |ij\rangle_{A'B'} \otimes \hat{f}_{ij} |\psi\rangle,$$

with

$$\hat{f}_{ij} \coloneqq \frac{1}{4} \hat{X}_A^i (\mathbb{1} + (-1)^i \hat{Z}_A) \hat{X}_B^j (\mathbb{1} + (-1)^j \hat{Z}_B).$$

Due to the relations $\hat{Z}_A |\psi\rangle = \hat{Z}_B |\psi\rangle$, we find

$$\hat{f}_{01} |\psi\rangle = \hat{f}_{10} |\psi\rangle = 0, \qquad \hat{f}_{00} |\psi\rangle = \hat{f}_{11} |\psi\rangle .$$

Hence, the output becomes

$$\Phi(|\psi\rangle) = \left|\phi^+\right\rangle_{A'B'} \otimes |\xi\rangle , \qquad \text{where} \quad |\xi\rangle = \sqrt{2}\hat{f}_{00} |\psi\rangle .$$

The state $|\xi\rangle$ is normalized, since the isometry Φ is constructed from unitary operations.

The swap isometry also enables self-testing of the measurement observables. As an example, consider the observable $\hat{B}_0 = \hat{Z}_B$. Applying the isometry to $\hat{B}_0 |\psi\rangle$, we find:

$$\Phi(\hat{B}_0 |\psi\rangle) = \sum_{i,j\in\{0,1\}} |ij\rangle \otimes \hat{g}_{ij} |\psi\rangle , \qquad \hat{g}_{ij} = \hat{f}_{ij}\hat{Z}_B.$$

Using the anticommutation $\{\hat{Z}_B, \hat{X}_B\} = 0$, we compute:

$$\begin{aligned}\hat{g}_{ij} &= \frac{1}{4}\hat{X}_A^i(\mathbb{1} + (-1)^i\hat{Z}_A)\hat{X}_B^j((-1)^j\mathbb{1} + \hat{Z}_B) \\ &= (-1)^j\hat{f}_{ij},\end{aligned}$$

so that

$$\Phi(\hat{B}_0 |\psi\rangle) = \sum_{i,j}(-1)^j |ij\rangle \otimes \hat{f}_{ij} |\psi\rangle .$$

Using $\hat{f}_{01} |\psi\rangle = \hat{f}_{10} |\psi\rangle = 0$, and $\hat{f}_{11} = \hat{f}_{00}$, we obtain:

$$\Phi(\hat{B}_0 |\psi\rangle) = \left(\mathbb{1} \otimes \sigma_z \left|\phi^+\right\rangle\right) \otimes |\xi\rangle .$$

Thus, $\hat{B}_0$ acts as σ_z on the support of $|\phi^+\rangle$, completing the self-test of Bob's observable. Similar computations apply to $\hat{B}_1$, $\hat{A}_0$, and $\hat{A}_1$.

Since $\hat{B}_0$ corresponds to a measurement operator $\hat{M}^B_{b|0} = (\mathbb{1} + b\hat{B}_0)/2$, by linearity we find:

$$\Phi(\hat{M}^B_{b|0} |\psi\rangle) = (\mathbb{1} \otimes \hat{\tilde{M}}^B_{b|0}) \left|\phi^+\right\rangle \otimes |\xi\rangle , \qquad \hat{\tilde{M}}^B_{b|0} = \frac{1}{2}(\mathbb{1} + b\sigma_z).$$

We have thus established a full state-and-measurement self-testing statement for the CHSH scenario. The Swap isometry maps the unknown physical state $|\psi\rangle$ and measurement observables to a reference realization $|\phi^+\rangle$ and Pauli operators, up to local isometries and ancillary degrees of freedom. Notably, the isometry is constructed solely from measurement operators, which by assumption act trivially on the purification space $\mathcal{H}_P$. This confirms that the self-test conforms to the fully device-independent definition of self-testing. □

1.6 Foundational aspects

1.6.1 Quantum mechanics and special relativity

As previously said, quantum mechanics permits nonlocal correlations, yet strictly forbids faster-than-light communication. This apparent tension is reconciled by the *no-communication theorem*

[89, 90], which ensures that the marginal statistics observed by one party are independent of the measurement choices made by a distant, space-like separated party. As emphasized by Shimony [91], this guarantees a "peaceful coexistence" between quantum nonlocality and the causal structure of special relativity. In this sense, quantum theory violates locality while preserving relativistic causality (see also [92–94]).

At a deeper level, however, foundational questions persist concerning the compatibility between quantum theory and relativity, particularly in relation to realism, separability, and the ontology of quantum states [16, 95–101]. For comprehensive philosophical and technical treatments, see [7, 18, 102], and for a concise overview of the implications of Bell nonlocality for special relativity, see [17].

1.6.2 Nonlocality beyond quantum mechanics

Crucially, the no-signaling condition is not unique to quantum theory—it is a generic constraint obeyed by a wide range of probabilistic theories. This raises a foundational question: is quantum mechanics the most nonlocal theory consistent with the no-signaling principle? As we have already seen, the answer is negative. PR boxes (1.8) are hypothetical devices that achieve the algebraic maximum violation of the CHSH inequality while still respecting no-signaling. These *super-quantum correlations* demonstrate that quantum correlations form a strict subset of all non-signaling correlations [25].

The existence of super-quantum correlations prompts a fundamental inquiry: why does nature realize precisely the set of correlations allowed by quantum mechanics, and not more extreme ones? To address this, various information-theoretic and physical principles [22, 27, 103–105] have been proposed in an attempt to recover the quantum set from broader constraints:

- *Information causality* [28] generalizes the no-signaling principle by bounding the information gain achievable by one party about another's data, given a limited amount of classical communication. While respected by quantum correlations, PR-box correlations violate it.

- *Macroscopic locality* [29] posits that coarse-grained measurements over many systems should yield statistics compatible with local hidden variable models. This principle holds in quantum mechanics but is violated by many post-quantum theories.

- *Local orthogonality* [30] constrains the joint probability of mutually exclusive events in multipartite scenarios. Like the others, it rules out some super-quantum correlations without fully characterizing the quantum set.

Although these principles constrain the space of physically plausible correlations, none has been proven sufficient to uniquely recover quantum mechanics. The search for an axiomatic characterization of quantum theory remains one of the central open problems in the foundations of physics.

This quest is formally approached through the framework of *generalized probabilistic theories* (GPTs) [106], which abstracts the operational structure of physical theories in terms of states, measurements, and transformations, without assuming the Hilbert space formalism. Developed in [107], GPTs enable the comparative study of classical, quantum, and hypothetical theories on a

common footing. Within this framework, one can investigate which operational principles—such as no-signaling, tomographic locality, purification, or information causality—are necessary or sufficient to single out quantum mechanics.

The GPT approach has also clarified that several theorems commonly associated with quantum theory, such as the *no-cloning theorem* [108, 109] and the *no-broadcasting theorem* [110], actually follow from the no-signaling principle. This underscores that such phenomena are not uniquely quantum, but rather generic features of a broad class of *non-classical* no-signaling theories.[12]

The implications of post-quantum correlations extend beyond foundational concerns into computational and physical plausibility. For instance:

- *Communication complexity.* As shown in [111], access to PR-box correlations collapses the communication complexity of many distributed tasks, enabling the solution of classically hard problems with minimal communication.

- *Nonlocal computation.* Certain computational problems that are intractable even for quantum systems become trivial when super-quantum correlations are allowed [67].

These results suggest that super-quantum correlations may lead to unphysical computational consequences, further motivating the idea that quantum mechanics occupies a uniquely constrained region in the broader theory space.

In this context, it has been argued that quantum theory may represent "an island in theory space", balancing physical and computational consistency. Aaronson [26], drawing from complexity theory and quantum information, proposes that the particular structure of quantum mechanics—such as its interference pattern and bounded computational power—may arise from deeper, as-yet-undiscovered principles.

Understanding why quantum mechanics is precisely as nonlocal as it is—neither more nor less—remains a defining challenge at the intersection of physics, computation, and logic. Progress in this direction may not only illuminate the foundations of quantum theory but also guide the development of future post-quantum technologies.

[12] In the GPT (Generalized Probabilistic Theories) framework, a theory is said to be *classical* if its state space is a *simplex*—a convex set in which every mixed state can be written uniquely as a convex combination of pure states. This reflects the classical intuition that mixed states represent mere *ignorance* about the true underlying pure state: there is a single, well-defined probability distribution over a set of mutually distinguishable alternatives. For example, a classical bit is either 0 or 1, and a probabilistic mixture like 50% 0 and 50% 1 simply reflects our uncertainty about which is the case.

In contrast, *non-classical* theories—such as quantum theory—have *non-simplicial* state spaces: the same mixed state can be decomposed into different sets of pure states. For instance, the maximally mixed qubit state can be written as an equal mixture of $|0\rangle$, $|1\rangle$, or as an equal mixture of $|+\rangle$, $|-\rangle$, among infinitely many others. These decompositions are not just different labels for the same epistemic uncertainty—they correspond to physically distinct preparations. This ambiguity implies that mixed states in non-classical theories cannot be interpreted as mere ignorance about a pre-existing reality. As a result, such theories forbid universal cloning and broadcasting, since attempting to clone a non-orthogonal pure state (as required by the different decompositions) would violate basic operational constraints such as no-signaling.

Chapter 2

Dual perspective on the quantum set

As we have seen in Chapter 1, understanding quantum correlations requires a detailed analysis of the geometry of the quantum set $\mathcal{Q}$. Although a direct characterization of this set is notoriously difficult, valuable insight can be obtained by studying its dual counterpart, $\mathcal{Q}^*$, the space of all Bell expressions that are bounded over $\mathcal{Q}$. This dual viewpoint provides an alternative and powerful way to probe the boundary of $\mathcal{Q}$ through the Bell inequalities it satisfies.

In this chapter, we develop and apply the dual approach to charactere the quantum correlations. We begin with a brief introduction to some elements of convex duality and the definition of orthogonal faces: subsets of the dual space that identify extremal directions of $\mathcal{Q}$ and expose its boundary structure. We then address the problem of characterizing such faces, by identifying all Bell expressions that are maximized by a given quantum behavior.

The second half of the chapter is devoted to the *Tsirelson point*, that is the quantum behavior that maximally violates the CHSH inequality. Following [52], we identify the complete family of Bell expressions that are maximized by this point, determine their local bounds, and characterize the resulting orthogonal face of $\mathcal{Q}^*$. This analysis provides an interpretation of the exposed nature of the Tsirelson point through the geometry of its dual face.

2.1 The dual geometry of the quantum set

In this section, we explore the structure of convex sets through the lens of their facial geometry and its dual counterpart. This dual perspective is particularly useful in the study of quantum correlations, where one is often interested in identifying the set of Bell expressions that are maximized by specific families of probability distributions.

2.1.1 Elements of convex duality

Let $\mathcal{K} \subset \mathbb{R}^d$ be a convex set. We begin by extending the notion of extremal points to higher-dimensional subsets.

Definition 2.1. A subset $F \subset \mathcal{K}$ is called a *face* of $\mathcal{K}$, denoted $F \triangleleft \mathcal{K}$, if

$$\forall \boldsymbol{y}, \boldsymbol{z} \in \mathcal{K},\ \lambda \in (0,1),\ \lambda \boldsymbol{y} + (1-\lambda)\boldsymbol{z} \in F \implies \boldsymbol{y}, \boldsymbol{z} \in F.$$

The *face dimension* d_F of a face F is the affine dimension of F.

Every face is itself a convex set. Intuitively, a face is a part of the boundary of $\mathcal{K}$ such that if any line segment lies on this part of the boundary, its endpoints must also be on this part of the boundary. Extremal points correspond to the faces of dimension zero.

Definition 2.2. Given a convex set $\mathcal{K} \subset \mathbb{R}^d$, we define its *dual set* as:

$$\mathcal{K}^* := \left\{ \boldsymbol{f} \in \mathbb{R}^d \colon \sup_{\boldsymbol{x} \in \mathcal{K}} \boldsymbol{f} \cdot \boldsymbol{x} \leq 1 \right\}. \tag{2.1}$$

This set captures the linear functionals (vectors $\boldsymbol{f}$) that are bounded by 1 on $\mathcal{K}$. The dual set is always convex and reflects the linear constraints satisfied by elements of $\mathcal{K}$.

A related notion is that of *supporting hyperplane*:

Definition 2.3. A hyperplane $H \subset \mathbb{R}^d$ with normal vector $\boldsymbol{f} \in \mathbb{R}^d$ is said to *support* $\mathcal{K}$ at a point $\boldsymbol{x} \in \mathcal{K}$ if

$$\boldsymbol{f} \cdot \boldsymbol{x} = \sup_{\boldsymbol{y} \in \mathcal{K}} \boldsymbol{f} \cdot \boldsymbol{y}. \tag{2.2}$$

In this case, the set $\{\boldsymbol{y} \in \mathcal{K} \colon \boldsymbol{f} \cdot \boldsymbol{y} = \boldsymbol{f} \cdot \boldsymbol{x}\}$ forms a face of $\mathcal{K}$. This face consists of all points in $\mathcal{K}$ that lie on the supporting hyperplane. The normal vector $\boldsymbol{f}$ to a supporting hyperplane defines a valid Bell functional when $\mathcal{K}$ represents a correlation set. In fact, it identifies a linear combination of correlations ($\boldsymbol{f} \cdot \boldsymbol{y}$) that is maximized by points on that specific face of $\mathcal{K}$.

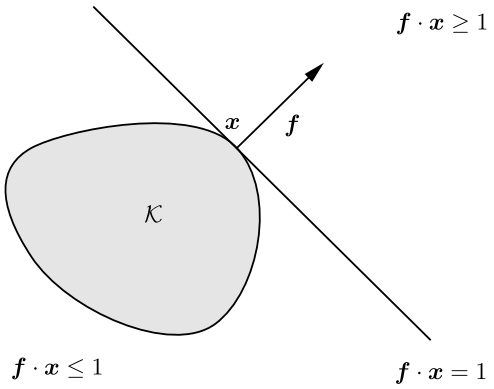

Figure 2.1: Illustration of a supporting hyperplane to the convex set $\mathcal{K}$ at the boundary point $\boldsymbol{x}$.

Remark 2.1. The relationship between supporting hyperplanes and the dual set is fundamental: if H is a supporting hyperplane for $\mathcal{K}$ with normal vector $\boldsymbol{f}$, then $\boldsymbol{f}$ is an element of $\mathcal{K}^*$ if we normalize $\boldsymbol{f}$ such that $\sup_{\boldsymbol{y} \in \mathcal{K}} \boldsymbol{f} \cdot \boldsymbol{y} = 1$. Conversely, any $\boldsymbol{f} \in \mathcal{K}^*$ with $\sup_{\boldsymbol{x} \in \mathcal{K}} \boldsymbol{f} \cdot \boldsymbol{x} = 1$ defines a supporting hyperplane. Thus, the boundary of $\mathcal{K}^*$ is defined by the normal vectors of the supporting hyperplanes of $\mathcal{K}$.

Definition 2.4. To any face $F \triangleleft \mathcal{K}$, it is associated an *orthogonal face* in the dual set $\mathcal{K}^*$, defined by

$$F^\perp := \{\boldsymbol{f} \in \mathcal{K}^* \colon \forall \boldsymbol{x} \in F,\ \boldsymbol{f} \cdot \boldsymbol{x} = 1\}\,. \tag{2.3}$$

The *orthogonal dimension* $d_F^\perp$ is the dimension of $F^\perp$.

In the context of quantum correlations, if F represents a set of observed correlations, $F^\perp$ represents all Bell expressions that achieve their maximum possible value for all correlations in F.

Definition 2.5. We say that a face $F \triangleleft \mathcal{K}$ is *exposed* if there exists $\boldsymbol{f} \in \mathcal{K}^*$ such that

$$\forall \boldsymbol{x} \in \mathcal{K} \colon \boldsymbol{f} \cdot \boldsymbol{x} = 1 \implies \boldsymbol{x} \in F. \tag{2.4}$$

An exposed face is one that can be defined as the set of points in $\mathcal{K}$ that maximize a specific linear functional. In this case, one can show that $d_F = d_{F^\perp}^\perp$. This relationship, often referred to as the *face-polar duality*, provides a powerful link between the geometry of a face and the geometry of its dual counterpart. The dimension pair $(d_F, d_F^\perp)$ thus provides a useful invariant that characterizes the geometry of exposed faces and the degeneracy of the Bell expressions they saturate. This degeneracy indicates how many independent Bell expressions are saturated by the face.

Moreover, since a face is a convex set, it can be expressed as the *convex hull*[1] of the extremal points it contains, namely

$$F = \operatorname{conv}\left(\mathcal{K}_{\text{ext}} \cap F\right). \tag{2.5}$$

As a result, the orthogonal face can be decomposed as the intersection of the orthogonal faces of the extremal points it contains:

$$F^\perp = \bigcap_{\boldsymbol{P} \in \mathcal{K}_{\text{ext}} \cap F} \{\boldsymbol{P}\}^\perp. \tag{2.6}$$

This observation allows one to reduce the analysis of arbitrary orthogonal faces to the simpler problem of characterizing the orthogonal faces of extremal points.

Let us conclude with one more definition, concerning the relation between two convex sets $\mathcal{K}, \mathcal{K}' \subset \mathbb{R}^d$:

Definition 2.6. We say that two convex sets $\mathcal{K}, \mathcal{K}' \subset \mathbb{R}^d$ are *isomorphic* if there exists a linear isomorphism

$$U \colon \mathbb{R}^d \to \mathbb{R}^d$$

such that

$$U(\mathcal{K}) = \mathcal{K}'.$$

This notion preserves all the relevant convex-geometric properties:

[1] The *convex hull* of a set is the collection of all convex combinations of its points. For a convex set $\mathcal{K}$, it holds $\mathcal{K} = \operatorname{conv} \mathcal{K}_{\text{ext}}$, see Appendix B for details.

Proposition 2.1. *Let $\mathcal{K}, \mathcal{K}' \subset \mathbb{R}^d$ be convex sets and suppose there exists a linear isomorphism $U\colon \mathbb{R}^d \to \mathbb{R}^d$ such that $\mathcal{K}' = U(\mathcal{K})$. Then:*

1. *For every face $F \triangleleft \mathcal{K}$, the image $U(F)$ is a face of $\mathcal{K}'$ with the same dimension.*
2. *The dual sets $\mathcal{K}^*$ and $\mathcal{K}'^*$ are isomorphic, with*

$$\mathcal{K}'^* = (U(\mathcal{K}))^* = (U^T)^{-1}(\mathcal{K}^*),$$

where U^T is the transpose of U.

The relationship between $\mathcal{K}$ and its dual $\mathcal{K}^*$ is knonw as *polarity*. For a closed convex set $\mathcal{K}$ containing the origin, its dual $\mathcal{K}^*$ is also a closed convex set containing the origin. Furthermore, if $\mathcal{K}$ is a compact set, then $\mathcal{K}^*$ is also compact. A key property is that the bidual of $\mathcal{K}$ is equal to the closure of $\mathcal{K}$, that is

$$\mathcal{K}^{**} \equiv \bar{\mathcal{K}}, \tag{2.7}$$

provided $\mathcal{K}$ contains the origin.[2] If $\mathcal{K}$ is closed and contains the origin, then $\mathcal{K}^{**} = \mathcal{K}$. This duality maps properties of one set directly to properties of the other one.

2.1.2 Dual of the quantum set

Following Definition 2.2, we introduce the dual of the quantum set in the CHSH scenario[3]

$$\mathcal{Q}^* \coloneqq \{\boldsymbol{\beta} \in \mathbb{R}^8 : \boldsymbol{\beta}\cdot\boldsymbol{P} \leq 1 \text{ for all } \boldsymbol{P} \in \mathcal{Q}\}. \tag{2.8}$$

Every vector $\boldsymbol{\beta} \in \mathcal{Q}^*$ defines a Bell functional whose value never exceeds 1 on quantum behaviours. Because the bidual of a closed convex cone coincides with the cone itself, a complete description of $\mathcal{Q}^*$ is equivalent to a complete description of $\mathcal{Q}$.

Duality has already proved powerful for the local and no-signalling sets. In binary-input/binary-output scenarios the local polytope is dual to the no-signalling polytope [112], yielding a one-to-one correspondence between tight Bell inequalities and extremal no-signalling behaviours.

By contrast, the quantum set is not a polytope and much less is known about its dual. A notable exception is the four-dimensional subspace $\mathcal{Q}_c$ of behaviours with vanishing marginals in the CHSH scenario: this subspace is *self-dual* [48, 112],

$$\mathcal{Q}_c \cong \mathcal{Q}_c^*. \tag{2.9}$$

[2] Consider the closed convex set $\mathcal{K} \coloneqq \{x \in \mathbb{R}\colon x \geq 1\}$, which does not contain the origin. The corresponding dual set is

$$\mathcal{K}^* \coloneqq \{y \in \mathbb{R}\colon yx \leq 1 \quad \forall x \geq 1\} = \{y \leq 0\},$$

since any positive y would violate the inequality for sufficiently large x. The bidual is then

$$\mathcal{K}^{**} \coloneqq \{x \in \mathbb{R}\colon xy \leq 1 \quad \forall y \leq 0\} = \{x \geq 0\},$$

as required for the inequality to hold when $y \to -\infty$. Thus, $\mathcal{K}^{**} = \{x \geq 0\} \supsetneq \mathcal{K}$, showing that the bidual strictly contains $\mathcal{K}$. The classical identity $\mathcal{K}^{**} = \overline{\mathcal{K}}$ fails in this case because $0 \notin \mathcal{K}$.

[3] Throughout this chapter we work in $\mathbb{R}^8$, the dimension of the behaviour space with binary inputs and outputs.

Both $\mathcal{Q}_c$ and its dual admit explicit analytic descriptions [44, 45, 48, 113–115]. Outside this subspace, however, the structure of the full eight-dimensional dual set $\mathcal{Q}^*$ remains largely unexplored.

It is important to highlight that elements of the dual set $\mathcal{Q}^*$ satisfying the condition $\boldsymbol{\beta}\cdot\boldsymbol{P}=1$ for some $\boldsymbol{P}\in\mathcal{Q}$ define supporting hyperplanes of the quantum set at the point $\boldsymbol{P}$. These vectors $\boldsymbol{\beta}$ provide a first-order, local approximation of the geometry of $\mathcal{Q}$ in the neighborhood of $\boldsymbol{P}$.

More precisely, consider a boundary point $\bar{\boldsymbol{P}}\in\mathcal{Q}$. Since $\mathcal{Q}$ is convex, its local structure around $\bar{\boldsymbol{P}}$ can be characterized by the set of all supporting hyperplanes at that point. These are given by Bell expressions $\boldsymbol{\beta}\in\mathbb{R}^8$ satisfying the conditions:

$$\boldsymbol{\beta}\cdot\boldsymbol{P}\leq 1 \quad \text{for all } \boldsymbol{P}\in\mathcal{Q}, \quad \text{and} \quad \boldsymbol{\beta}\cdot\bar{\boldsymbol{P}}=1. \tag{2.10}$$

The set of such $\boldsymbol{\beta}$ vectors defines the face of the dual set $\mathcal{Q}^*$ that is orthogonal to $\bar{\boldsymbol{P}}$, and encapsulates all Bell inequalities that are maximally violated by $\bar{\boldsymbol{P}}$. This characterizes the local boundary structure of $\mathcal{Q}$ at $\bar{\boldsymbol{P}}$ from a dual perspective.

The next two chapters develop a systematic study of $\mathcal{Q}^*$ in the full behaviour space. Our strategy is to fix a quantum point and to determine *all* Bell functionals that it maximises, thereby mapping the local boundary geometry of $\mathcal{Q}$ and revealing new extremal elements of $\mathcal{Q}^*$.

2.1.3 The Tsirelson behavior

In this chapter, following the approach of [52], we analytically characterize all elements of the dual quantum set $\mathcal{Q}^*$ that are maximized by the *Tsirelson point*, denoted $\boldsymbol{P}_T$. The Tsirelson behavior has already been introduced in Sec. 1.4.2, where we established that it achieves the quantum bound of the CHSH inequality (cf. Eq. (1.61)):

$$\boldsymbol{\beta}_{\text{CHSH}}\cdot\boldsymbol{P}_T=2\sqrt{2}.$$

Our goal is to provide a complete analytic description of the corresponding face of $\mathcal{Q}^*$, thereby obtaining a tight, first-order approximation of the quantum set $\mathcal{Q}$ in the neighborhood of this maximally nonlocal point.

Recall that the Tsirelson point corresponds to the following table of correlators:

$$\boldsymbol{P}_T \coloneqq \begin{array}{c|c|c} 1 & 0 & 0 \\ \hline 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \hline 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{array}. \tag{2.11}$$

Moreover, in Sec. 1.5.5, we discussed how this behavior self-tests the specific quantum realization

$$\begin{aligned} \left|\phi^+\right\rangle &\coloneqq \frac{1}{\sqrt{2}}(|00\rangle+|11\rangle), \\ A_x &\coloneqq \frac{Z_A+(-1)^x X_A}{\sqrt{2}}, \quad B_0\coloneqq Z_B, \quad B_1\coloneqq X_B. \end{aligned} \tag{2.12}$$

From the correlator table (2.11), we observe that $\boldsymbol{P}_T$ has vanishing marginal expectations, i.e., $\langle A_x\rangle=\langle B_y\rangle=0$. Within the subspace of quantum correlations with vanishing

marginals—denoted $\mathcal{Q}_{\text{c}}$—it is known that $\boldsymbol{P}_T$ is an *exposed* point, uniquely identified by the CHSH inequality. That is, the hyperplane defined by $\boldsymbol{\beta}_{\text{CHSH}} \cdot \boldsymbol{P} = 2\sqrt{2}$ intersects $\mathcal{Q}_{\text{c}}$ only at the Tsirelson point.

However, while this uniqueness result holds within $\mathcal{Q}_{\text{c}}$, recent numerical studies suggest that in the full 8-dimensional space of quantum behaviors (i.e., without assuming vanishing marginals), $\boldsymbol{P}_T$ may in fact maximize other Bell expressions as well [46]. Our objective in this chapter is to analytically identify all such expressions by characterizing the corresponding face of $\mathcal{Q}^*$.

2.2 Conditions for the maximal violation of a Bell inequality

In the following, we are going to identify all Bell expressions that are maximized by the Tsirelson point. To do so, we first introduce two essential tools for characterizing such Bell functionals: the *variational method*, which exploits optimality conditions for quantum behaviours, and the *Sum-of-Squares (SOS) decomposition method*, which provides constructive certificates of quantum bounds. These complementary techniques will allow us to fully describe the face of the dual quantum set $\mathcal{Q}^*$ associated with the Tsirelson point.

2.2.1 The variational method

We start by presenting the *variational method*, which provides a necessary condition for a quantum realization to maximize a given Bell expression. This method is particularly useful when attempting to characterize all Bell inequalities that are maximally violated by a specific quantum behavior, such as the Tsirelson point. As demonstrated in [79, 116], the variational method allows one to construct insightful Bell expressions tailored to various states.

Let $|\psi\rangle$ be a target quantum state, and consider a Bell inequality defined by a Bell expression $\boldsymbol{\beta}$. In order for $\boldsymbol{\beta}$ to be maximally violated by a quantum realization of the form $\{|\psi\rangle, \hat{\Pi}^A_{a|x}, \hat{\Pi}^B_{b|y}\}$, it is necessary that the corresponding Bell operator

$$\hat{S} \coloneqq \sum_{a,x,b,y} \beta_{a,b,x,y}\, \hat{\Pi}^A_{a|x} \otimes \hat{\Pi}^B_{b|y}$$

has $|\psi\rangle$ as an eigenstate associated with its maximal eigenvalue.[4]

In this context, the variational method seeks to identify Bell expressions $\boldsymbol{\beta}$ such that the associated Bell operator $\hat{S}$ is maximized by $|\psi\rangle$ for some choice of measurement operators. The

[4] Let $\hat{S}$ be a Hermitian operator on a finite-dimensional Hilbert space $\mathcal{H}$, and let $|\psi\rangle \in \mathcal{H}$ be a normalized state. Then the expectation value $\langle\psi|\hat{S}|\psi\rangle$ is maximized if and only if $|\psi\rangle$ belongs to the eigenspace of $\hat{S}$ associated with its largest eigenvalue.

Proof. Since $\hat{S}$ is Hermitian, it admits a spectral decomposition. Let $\{|v_i\rangle\}_{i=1}^n$ be an orthonormal basis of eigenvectors of $\hat{S}$ with corresponding real eigenvalues λ_i. Without loss of generality, arrange the eigenvalues so that $\lambda_1 = \cdots = \lambda_k > \lambda_{k+1} \geq \cdots \geq \lambda_n$, where λ_1 is the maximal eigenvalue with degeneracy k.

Any normalized state $|\psi\rangle \in \mathcal{H}$ can be written as

$$|\psi\rangle = \sum_{i=1}^{n} c_i |v_i\rangle, \quad \text{with } \sum_i |c_i|^2 = 1.$$

idea is to impose a first-order stationarity condition on $\hat{S}$ with respect to small variations of the measurement observables, thereby deriving constraints that must be satisfied by $\boldsymbol{\beta}$.

More concretely, the variational method proceeds through the following steps:

1. **Fix a target state.** Choose a quantum state $|\psi\rangle$ that is hypothesized to maximize a class of Bell expressions.

2. **Define a Bell operator.** Construct a Bell operator $\hat{S}$ that depends on a set of measurement operators $\hat{M}_x^{(i)}$, and which is such that $|\psi\rangle$ is an eigenstate with maximal eigenvalue.

3. **Parametrize the measurements.** Introduce a parametric form for the measurement observables (e.g., projective measurements defined by Bloch sphere angles) such that the Bell operator $\hat{S}$ becomes an explicit function of the measurement parameters.

4. **Induce a family of Bell expressions.** Express the Bell operator $\hat{S}$ in terms of its decomposition over the standard correlators (or probabilities), thereby identifying a corresponding Bell expression $\boldsymbol{\beta}$ that depends on the measurement parameters.

5. **Introduce infinitesimal variations.** Consider small perturbations[5] of the measurement operators:
$$\hat{M}_x^{(i)} \to \hat{M}_x^{(i)} + \delta_x^{(i)}\, \hat{M}_x^{(i)\perp}, \tag{2.13}$$
where $\hat{M}_x^{(i)\perp}$ is a Hermitian operator orthogonal (in the operator sense) to $\hat{M}_x^{(i)}$.

6. **Derive first-order stationarity conditions.** Impose the condition that $|\psi\rangle$ remains a stationary point of the Bell operator under these variations, i.e.,
$$\langle\psi|\, \frac{\partial \hat{S}}{\partial \delta_x^{(i)}}\, |\psi\rangle = 0 \qquad \forall\, x, i. \tag{2.14}$$

These conditions yield a system of linear equations that the coefficients of the Bell expression $\boldsymbol{\beta}$ must satisfy. Any solution to this system corresponds to a Bell expression which is at least

Then the expectation value is

$$\langle\psi|\, \hat{S}\, |\psi\rangle = \sum_{i=1}^{n} \lambda_i |c_i|^2 = \sum_{i=1}^{k} \lambda_1 |c_i|^2 + \sum_{i=k+1}^{n} \lambda_i |c_i|^2.$$

Since $\lambda_i \leq \lambda_1$ for all i, we have

$$\langle\psi|\, \hat{S}\, |\psi\rangle \leq \lambda_1 \sum_{i=1}^{n} |c_i|^2 = \lambda_1,$$

with equality if and only if $|c_i|^2 = 0$ for all $i > k$, i.e., if and only if $|\psi\rangle$ lies entirely within the eigenspace corresponding to λ_1. Therefore, the maximum of the expectation value is λ_1, and it is achieved if and only if $|\psi\rangle$ belongs to the eigenspace of $\hat{S}$ associated with its largest eigenvalue. □

[5] In general, a perturbation of the ideal measurements operators can always be written as

$$\hat{\Pi}_{a|x}^{(i)} \mapsto e^{ih_x\delta} \hat{\Pi}_{a|x}^{(i)} e^{-ih_x\delta} \approx \hat{\Pi}_{a|x}^{(i)} + \delta\, i[h_x, \hat{\Pi}_{a|x}^{(i)}]$$

for $h_x = h_x^\dagger$.

locally maximized by the quantum realization defined by $|\psi\rangle$ and the original unperturbed measurements.

In practice, the variational method can be used to explore the space of Bell expressions that are maximized by a fixed state and measurement strategy, and to determine whether any nontrivial family of such expressions exists. We will apply this method to the Tsirelson realization in the next section.

Remark 2.2. Note that:

- In Step 2 the measurements should be chosen such that
$$\mathrm{span}\{\hat{\Pi}^{(1)}_{a_1|x_1} \otimes \cdots \otimes \hat{\Pi}^{(n)}_{a_n|x_n}\}$$
contains the Bell operator $\hat{S}$. Furthermore, when the measurement operators for at least one party define an overcomplete operator basis (the identity excluded), several choices of β could be made in Step 3. In this case, the derivative in (2.14) is to be understood accordingly.

- Eigenvalue perturbation implies that (2.14) must be satisfied when $\beta(\boldsymbol{P})$ is maximized by the considered state and settings. Importantly, the method gives a *necessary condition* for the Bell expression candidate β obtained by the choice of operator $\hat{S}$ and of settings $\hat{M}^{(i)}_x$ that was made in Step 2.

- If a choice of measurements never verifies (2.14) for all possible $\hat{S}$, then these settings cannot be used in any expression maximized by the considered state. In particular, these settings cannot be used to self-test the considered state or, in other words, the realisation corresponding to the considered state and those measurements cannot be self-tested.

- Expression $\boldsymbol{\beta}$ may still be maximally violated by the considered state even when condition (2.14) is not verified. This is however only possible with other measurement settings and thus a different corresponding operator $\hat{S}'$. It may thus be helpful to consider several Bell operators $\hat{S}$.

- The operator $\hat{S}$ acts on a finite-dimensional product Hilbert space and can therefore be fully parametrized. The constraints are that the state $|\psi\rangle$ lies in the support[6] of $\hat{S}$ and is one of its maximal eigenstates. Furthermore, finite dimensional measurements can also be expressed with a finite number of parameters. In all generality, the whole variational method can therefore be parametrized with finitely many parameters. When considering a complete parametrization rather than a particular choice of $\hat{S}$ and measurements, the implications of (2.14) become stronger.

Example 2.1 (Application of the variational method). We illustrate the variational method by considering the maximally entangled state $|\psi\rangle = |\phi^+\rangle$ and the Bell operator

$$\hat{S} = \frac{1}{2}\left(\hat{X}_A\hat{X}_B + \hat{Z}_A\hat{Z}_B\right). \tag{2.15}$$

[6]The *support* of an operator $\hat{O}\colon \mathcal{H} \to \mathcal{H}$ is the subspace of the Hilbert space on which $\hat{O}$ acts non-trivially. Formally, it is the orthogonal complement of the kernel of $\hat{O}$, i.e., $\mathrm{supp}(\hat{O}) \coloneqq \ker(\hat{O})^\perp$.

This operator is maximized by $|\phi^+\rangle$ and corresponds to different Bell expressions depending on the actual choice of measurement observables.

We parametrize Alice's and Bob's local measurements as qubit observables in the $\hat{X}$–$\hat{Z}$ plane:

$$\hat{A}_x = \cos(a_x)\hat{Z}_A + \sin(a_x)\hat{X}_A, \tag{2.16a}$$
$$\hat{B}_y = \cos(b_y)\hat{Z}_B + \sin(b_y)\hat{X}_B, \tag{2.16b}$$

with angles $a_x, b_y \in \mathbb{R}$, $x, y \in \{0, 1\}$. Rewriting the Bell operator $\hat{S}$ in terms of the measurement observables $\hat{A}_x$, $\hat{B}_y$ yields:

$$\begin{aligned}\hat{S} = \frac{1}{2\sin(a_0 - a_1)\sin(b_0 - b_1)}\Big[&\cos(a_1 - b_1)\hat{A}_0\hat{B}_0 - \cos(a_1 - b_0)\hat{A}_0\hat{B}_1 \\ &- \cos(a_0 - b_1)\hat{A}_1\hat{B}_0 + \cos(a_1 - b_1)\hat{A}_1\hat{B}_1\Big].\end{aligned} \tag{2.17}$$

This in turn defines a family of Bell expressions of the form:

$$\begin{aligned}\beta(\boldsymbol{P}) = \frac{1}{2\sin(a_0 - a_1)\sin(b_0 - b_1)}\Big[&\cos(a_1 - b_1)\langle A_0B_0\rangle - \cos(a_1 - b_0)\langle A_0B_1\rangle \\ &- \cos(a_0 - b_1)\langle A_1B_0\rangle + \cos(a_0 - b_0)\langle A_1B_1\rangle\Big].\end{aligned} \tag{2.18}$$

However, not all such expressions are necessarily maximized by $|\phi^+\rangle$. As a counterexample, consider the angles $a_0 = 0$, $a_1 = \pi/2$, $b_y = -(-1)^{y+1}\pi/6$, which yield the Bell expression:

$$\beta(\boldsymbol{P}) = \frac{1}{2\sqrt{3}}\left[\langle A_0B_0\rangle + \langle A_0B_1\rangle + \sqrt{3}\langle A_1B_0\rangle - \sqrt{3}\langle A_1B_1\rangle\right]. \tag{2.19}$$

This expression evaluates to 1 for $|\phi^+\rangle$ with the specified measurements. However, this value can be exceeded by a different choice of measurements, such as $B_y = \cos(\pi/4)\hat{Z}_B - (-1)^{y+1}\hat{X}_B$, which gives:

$$\beta(\boldsymbol{P}) = \frac{\sqrt{2}}{2\sqrt{3}}(1 + \sqrt{3}) \approx 1.12 > 1.$$

Moreover, for these settings, the value $\beta(\boldsymbol{P}) = 1$ can be attained using a partially entangled state. This demonstrates that the Bell expression in Eq. (2.19) is not uniquely maximized by $|\phi^+\rangle$.

To rule out such cases, we apply the variational method to check whether $|\phi^+\rangle$ is a local extremum. We perturb the measurement operators as:

$$\hat{A}_x \to \hat{A}_x + \delta_{A_x}\left(-\sin(a_x)\hat{Z}_A + \cos(a_x)\hat{X}_A\right), \tag{2.20a}$$
$$\hat{B}_y \to \hat{B}_y + \delta_{B_y}\left(-\sin(b_y)\hat{Z}_B + \cos(b_y)\hat{X}_B\right), \tag{2.20b}$$

and compute the derivatives $\partial\hat{S}/\partial\delta_{A_x}$ and $\partial\hat{S}/\partial\delta_{B_y}$.

We then evaluate the conditions:

$$\begin{aligned}\left\langle\phi^+\right|\frac{\partial\hat{S}}{\partial\delta_{A_x}}\left|\phi^+\right\rangle &\propto \cos(a_0 - a_1) = 0, \\ \left\langle\phi^+\right|\frac{\partial\hat{S}}{\partial\delta_{B_y}}\left|\phi^+\right\rangle &\propto \cos(b_0 - b_1) = 0.\end{aligned} \tag{2.21}$$

These imply that $|\phi^+\rangle$ is a stationary point if and only if Alice's and Bob's measurements are mutually unbiased (i.e., complementary):

$$a_1 = a_0 \pm \frac{\pi}{2}, \qquad b_1 = b_0 \pm \frac{\pi}{2}.$$

Fixing $a_1 = a_0 + \frac{\pi}{2}$ and $b_1 = b_0 - \frac{\pi}{2}$, we introduce a parameter $c \coloneqq b_0 - a_0$ and write the corresponding Bell expression as:

$$\beta(\boldsymbol{P}) = \cos(c)\langle A_0 B_0\rangle + \sin(c)\langle A_0 B_1\rangle + \sin(c)\langle A_1 B_0\rangle - \cos(c)\langle A_1 B_0\rangle. \tag{2.22}$$

This one-parameter family of Bell expressions is locally maximized by $|\phi^+\rangle$ for all c.

From this example, we observe that the local optimality condition significantly constrains the set of candidate Bell expressions that could be maximally violated by the state $|\phi^+\rangle$. The resulting admissible family includes the CHSH inequality as a special case, but also encompasses a broader class of Bell expressions. It can be verified that, for all $c \in (0, \pi/4]$, the corresponding Bell expressions not only are locally maximized by $|\phi^+\rangle$, but also self-test the maximally entangled state.

Finally, note that a natural strategy to construct an initial Bell operator $\hat{S}$ that is tailored to a target state $|\psi\rangle$ is to consider its stabilizers. If a set of Hermitian operators $\{\hat{S}_i\}$ satisfies $\hat{S}_i |\psi\rangle = |\psi\rangle$ and has spectrum contained in $[-1, 1]$, then any convex combination of them, i.e.,

$$\hat{S} = \sum_i p_i \hat{S}_i, \qquad \text{with } p_i \geq 0,\ \sum_i p_i = 1,$$

yields a valid Bell operator with $|\psi\rangle$ as a maximal eigenstate. This method provides a constructive way to engineer Bell operators suited for a given state.

2.2.1.1 Second-order condition

In variational analysis, identifying a local maximum requires more than just the vanishing of the first derivative: one must also verify that the second derivative is negative in the relevant directions. Accordingly, the variational method can be extended to second order. This step not only strengthens the condition for maximality but also provides geometric insight into the structure of the Bell operator around the target state. In the case of multiple measurement parameters, the second-order behavior is encoded in a Hessian matrix that captures the sensitivity of the Bell value to infinitesimal perturbations in different directions.

For a given Bell operator $\hat{S}$, consider a perturbation of the measurement operators as in Eq. (2.13), inducing a corresponding perturbation of $\hat{S}$. Then, assuming the first-order condition

$$\langle\psi| \frac{\partial \hat{S}}{\partial \delta_i} |\psi\rangle = 0$$

is satisfied for all relevant perturbations δ_i, the second-order condition for $|\psi\rangle$ to be a local maximum is that the Hessian matrix γ is negative semidefinite:

$$\gamma \preceq 0, \quad \text{where} \quad \gamma = \mu + \nu, \tag{2.23a}$$

$$\mu_{ij} \coloneqq \langle\psi| \frac{\partial^2 \hat{S}}{\partial\delta_i \partial\delta_j} |\psi\rangle\,, \tag{2.23b}$$

$$\nu_{ij} \coloneqq 2 \sum_l \frac{\langle\psi| \frac{\partial \hat{S}}{\partial\delta_i} |\psi_l\rangle \langle\psi_l| \frac{\partial \hat{S}}{\partial\delta_j} |\psi\rangle}{1-\lambda_l}, \tag{2.23c}$$

where $\{|\psi_l\rangle\}_l$ and $\{\lambda_l\}_l$ denote the eigenstates and eigenvalues of $\hat{S}$ different from $|\psi\rangle$ and its maximal eigenvalue (normalized to 1). The matrix μ captures the direct second-order variation of the Bell operator expectation, while ν accounts for the variation in the eigenstate of $\hat{S}$ due to perturbations.

Example 2.2. Let us revisit the Bell operator defined in Eq. (2.15) with measurements of the form (2.16). We consider second-order perturbations of the measurements as follows:

$$\hat{A}_x \mapsto \hat{A}_x + \delta_{a_x}(-\sin(a_x)\hat{Z}_A + \cos(a_x)\hat{X}_A) - \tfrac{1}{2}\delta_{a_x}^2 \hat{A}_x,$$
$$\hat{B}_y \mapsto \hat{B}_y + \delta_{b_y}(-\sin(b_y)\hat{Z}_B + \cos(b_y)\hat{X}_B) - \tfrac{1}{2}\delta_{b_y}^2 \hat{B}_y.$$

By the symmetry under local unitaries, it is sufficient to consider variations in only one of the measurement settings per party. Imposing the first-order optimality conditions, we compute the second-order Hessian matrix for the variations of A_1 and B_1:

$$\gamma = \begin{pmatrix} -1 & \cos(2c) \\ \cos(2c) & -1 \end{pmatrix},$$

where c is the only relevant parameter. The eigenvalues of this matrix are

$$\lambda_\pm = -2\sin^2(c), \quad -2\cos^2(c),$$

which are both strictly negative for $c \in (0, \pi/4]$. In particular, the eigenvalues are equal when $c = \pi/4$, which corresponds to the standard CHSH expression.

These eigenvalues quantify the curvature of the Bell operator's expectation value in the principal directions of measurement perturbations. The fact that the CHSH case yields equal eigenvalues indicates that it is the most robust—i.e., isotropic—Bell expression in the family (2.22), in terms of second-order stability under small deviations in the measurement settings.

2.2.2 Sum-of-Squares decomposition method

The variational method provides a way to construct Bell expressions potentially maximized by a target quantum state. However, it lacks the ability to certify optimality: there is no guarantee that other quantum realizations cannot achieve a higher value. To address this limitation, we now present an alternative approach based on the algebraic method of *sum-of-squares (SOS) decompositions* [115]. This method not only constructs Bell expressions tailored to a given state

but also certifies that the constructed expression is globally maximized by that state within the quantum set.

The SOS approach is grounded in the representation of Bell expressions as formal polynomials over a noncommutative algebra of measurement operators. We begin by introducing the relevant algebraic framework.

2.2.2.1 Formal polynomials

Definition 2.7 ([117]). Let $\{X_i\}_i$ be a set of indeterminates in an associative algebra $\mathcal{A}$ over a field $\mathbb{K}$. A *formal multivariate polynomial* is a finite linear combination

$$S = \sum_i \alpha_i M_i, \tag{2.24}$$

where each M_i is a monomial, i.e., a finite product of the *indeterminates* (such as X_1, X_1X_2, or $X_1^2X_2X_1$), and $\alpha_i \in \mathbb{K}$ are scalar coefficients.

In the context of generalized Bell scenarios, we consider the associative algebra of polynomials $\mathbb{K}[\boldsymbol{X}, \boldsymbol{X}^\dagger]$, generated by indeterminates $X^{(j)}_{a_j|x_j}$ and their adjoints $X^{(j)\dagger}_{a_j|x_j}$ [7]. Each $X^{(j)}_{a_j|x_j}$ is associated with the outcome a_j of the measurement x_j performed by party j. These indeterminates are treated as formal variables subject to the following algebraic relations:

$$\text{Hermiticity:} \quad \left(X^{(j)}_{a_j|x_j}\right)^\dagger = X^{(j)}_{a_j|x_j}, \tag{2.25a}$$

$$\text{Orthogonality:} \quad X^{(j)}_{a_j|x_j} X^{(j)}_{a'_j|x_j} = \delta_{a_j,a'_j} X^{(j)}_{a_j|x_j}, \tag{2.25b}$$

$$\text{Normalization:} \quad \sum_{a_j} X^{(j)}_{a_j|x_j} = 1, \tag{2.25c}$$

$$\text{Commutation between parties:} \quad \left[X^{(j)}_{a_j|x_j}, X^{(j')}_{a_{j'}|x_{j'}}\right] = 0 \quad \text{for } j \neq j'. \tag{2.25d}$$

Each monomial M_i can thus be written as a product of local components:

$$M_i[\{X^{(j)}_{a_j|x_j}\}] = M_i^{(1)}[\{X^{(1)}_{a_1|x_1}\}] \cdots M_i^{(n)}[\{X^{(n)}_{a_n|x_n}\}], \tag{2.26}$$

where $M_i^{(j)}$ involves only the indeterminates associated with party j.

A polynomial S is said to have *local degree* 1 if all monomials $M_i^{(j)}$ are of degree one in the local variables. Such polynomials correspond to Bell expressions.

To associate a Bell expression to a formal polynomial, we substitute each indeterminate $X^{(j)}_{a_j|x_j}$ with a projection operator $\hat{\Pi}^{(j)}_{a_j|x_j}$ representing the outcome a_j of measurement x_j for party j, and take the expectation value with respect to a quantum state $|\psi\rangle$. That is, we define:

$$\beta(\boldsymbol{P}) \coloneqq \sum_{\boldsymbol{a},\boldsymbol{x}} \beta_{\boldsymbol{a}|\boldsymbol{x}} P(\boldsymbol{a}|\boldsymbol{x}), \qquad \text{where } P(\boldsymbol{a}|\boldsymbol{x}) \coloneqq \langle\psi| \bigotimes_j \hat{\Pi}^{(j)}_{a_j|x_j} |\psi\rangle . \tag{2.27}$$

[7] This is a *free associative $*$-algebra* over the field $\mathbb{K}$. A $*$-algebra is an associative algebra equipped with an involution, an anti-automorphic map (i.e., a map that reverses the order of multiplication) denoted by $\dagger$ (or $*$). In this context, the involution satisfies $(kA)^\dagger = \bar{k}A^\dagger$, $(A+B)^\dagger = A^\dagger + B^\dagger$, and $(AB)^\dagger = B^\dagger A^\dagger$ for $k \in \mathbb{K}$ and elements A, B in the algebra.

Conversely, every Bell expression of this form corresponds to a unique formal polynomial. We refer to such polynomials as *formal Bell polynomials*. We define the expectation of a monomial by

$$\langle M_i \rangle \coloneqq \langle \psi | \, \hat{M}_i \, | \psi \rangle \, , \tag{2.28}$$

where $\hat{M}_i$ is the quantum implementation of the monomial M_i in terms of projectors and $|\psi\rangle$ is a normalized state. When the local degree of M_i is one, the expectation value admits a direct probabilistic interpretation in terms of $P(a, b|x, y)$, as in eqs. (1.26).

Remark 2.3. The mapping between Bell expressions, formal polynomials, and Bell operators extends beyond the local degree 1 case. Substituting

$$X^{(j)}_{a_j|x_j} \mapsto \hat{\Pi}^{(j)}_{a_j|x_j} \tag{2.29}$$

allows one to define operator-valued versions of arbitrary polynomials and analyze their expectation values.

Example 2.3. The CHSH Bell expression

$$\beta_{\text{CHSH}}(\boldsymbol{P}) \coloneqq \sum_{a,b,x,y} (-1)^{xy} ab P(\boldsymbol{a}|\boldsymbol{x}) \tag{2.30}$$

with $a, b \in \{\pm 1\}$, and $x, y \in \{0, 1\}$ corresponds to the formal polynomial

$$\beta_{\text{CHSH}} = \sum_{a,b,x,y} (-1)^{xy} ab X^A_{a|x} X^B_{b|y} \tag{2.31}$$

and viceversa.

Lastly, just as measurement projectors can be parametrized using different bases (e.g., Pauli operators), it is sometimes convenient to represent formal polynomials using alternative sets of indeterminates. These may correspond to operator combinations better suited for symmetry analysis or optimization, and play a role in constructing SOS decompositions.

One such representation is obtained by applying the discrete Fourier transform to the outcome variables of a measurement. Specifically, for each measurement setting x_j of party j, we define the corresponding *Fourier-transformed measurement operator* as

$$Y^{(j)}_x \coloneqq \sum_{a_j=0}^{d-1} \omega_d^{a_j} \, X^{(j)}_{a_j|x_j}, \tag{2.32}$$

where $\omega_d \coloneqq \exp(2\pi i/d)$ is the primitive d-th complex root of unity. These new indeterminates encode the same information as the projective measurement variables $X^{(j)}_{a_j|x_j}$, but in a basis that diagonalizes the cyclic structure of outcomes.

The transformation is invertible: each $X^{(j)}_{a|x}$ can be recovered as a linear combination of the powers $\left(Y^{(j)}_x\right)^m$ with $m = 0, \ldots, d-1$. As a result, all polynomial expressions over the original indeterminates can be equivalently expressed in terms of the $Y^{(j)}_x$ operators. Note that this formulation requires working over the complex field $\mathbb{K} = \mathbb{C}$ when $d > 2$.

The indeterminates $Y_x^{(j)}$ are unitary by construction:

$$\left(Y_x^{(j)}\right)^\dagger Y_x^{(j)} = Y_x^{(j)} \left(Y_x^{(j)}\right)^\dagger = 1, \tag{2.33}$$

and in the special case $d = 2$ (binary outcomes), they are also Hermitian: $\left(Y_x^{(j)}\right)^\dagger = Y_x^{(j)}$.

Adopting the Fourier-transformed notation, one can rewrite familiar Bell expressions in a more compact algebraic form. For instance, the CHSH polynomial becomes:

$$\beta_{\text{CHSH}} = A_0B_0 + A_0B_1 + A_1B_0 - A_1B_1, \tag{2.34}$$

where we have denoted $A_x \coloneqq Y_x^{(0)}$ and $B_y \coloneqq Y_y^{(1)}$ for the two parties.

2.2.2.2 The SOS method

The polynomial framework introduced above enables a *purely algebraic* certification of quantum bounds through *sum-of-squares* (SOS) decompositions. Given a formal Bell polynomial β, we say that it satisfies the inequality

$$\beta \preceq \beta_{\mathcal{Q}} \tag{2.35}$$

if there exists a decomposition of the form

$$\beta_{\mathcal{Q}} - \beta = \sum_s O_s^\dagger O_s, \tag{2.36}$$

where each O_s is a formal polynomial over noncommuting indeterminates in $\mathcal{A}$. This expression guarantees that for any quantum implementation (i.e., any state $|\psi\rangle$ and set of projective measurements), the corresponding Bell operator $\hat{S}$ satisfies

$$\langle\psi|\,\hat{S}\,|\psi\rangle \leq \beta_{\mathcal{Q}}, \tag{2.37}$$

establishing $\beta_{\mathcal{Q}}$ as a Tsirelson-type bound for the expression β.

Remark 2.4. The SOS structure provides a *dimensional-independent certificate* of the bound, as it depends only on the algebraic relations among the operators and not on a specific Hilbert space representation.

Moreover, if $\langle\phi|\,\hat{S}\,|\phi\rangle = \beta_{\mathcal{Q}}$ for some implementation, then

$$\sum_s \|\hat{O}_s\,|\phi\rangle\,\|^2 = 0 \quad \Longrightarrow \quad \hat{O}_s\,|\phi\rangle = 0 \quad \forall s, \tag{2.38}$$

meaning that the state $|\phi\rangle$ is in the kernel of all $\hat{O}_s$, and the operators $\hat{O}_s$ *nullify* the state. For this reason, the $\hat{O}_i$ are also called *nullifiers* of $|\phi\rangle$. The converse also holds:

$$\langle\phi|\,\hat{S}\,|\phi\rangle = \beta_{\mathcal{Q}} \quad \Longleftrightarrow \quad \hat{O}_s\,|\phi\rangle = 0 \quad \forall s. \tag{2.39}$$

We then consider the subspace

$$\mathcal{N}_{\mathcal{A}} \coloneqq \left\{O \in \mathcal{A}\colon \hat{O}\,|\psi\rangle = 0\right\}, \tag{2.40}$$

and a generating set $\{N_i\}$ for the nullifier subspace $\mathcal{N}_{\mathcal{A}}$. We denote by $\boldsymbol{N}$ the vector whose components are the generators N_s. Any formal polynomial $O \in \mathcal{N}_{\mathcal{A}}$ can then be written as a linear combination:

$$O_s = \sum_i c_{si} N_i, \qquad c_{si} \in \mathbb{R} \quad \forall i.$$

An SOS decomposition with nullifiers can thus be written as:

$$\begin{aligned} \beta_{\mathcal{Q}} - \beta = \sum_s O_s^\dagger O_s = \sum_s \sum_{i,j} c_{si}^\dagger c_{sj} N_i^\dagger N_j \\ = \sum_{i,j} \left(\sum_s c_{si}^\dagger c_{sj} \right) N_i^\dagger N_j = \boldsymbol{N}^\dagger W \boldsymbol{N}, \end{aligned} \tag{2.41}$$

where the matrix $W \coloneqq \sum_s c_{si}^\dagger c_{sj}$ is positive semidefinite, i.e., $W \succeq 0$.

We therefore conclude that the problem of verifying whether a Bell expression β (in formal polynomial representation) satisfies $\beta \preceq \beta_{\mathcal{Q}}$, where the inequality is saturated for a given quantum implementation $(|\phi\rangle, \hat{S})$ reduces to checking the existence of a positive semidefinite matrix W and a set of formal nullifiers for the state $|\phi\rangle$ such that:

$$\beta_{\mathcal{Q}} - \beta = \boldsymbol{N}^\dagger W \boldsymbol{N}. \tag{2.42}$$

If (for a set of nullifiers) such a matrix W exists, we say that it provides a *certificate* for the inequality $\beta(\boldsymbol{P}) \leq \beta_{\mathcal{Q}}$.

In practice, the SOS decomposition is often constructed with a *residual term*:

$$\boldsymbol{N}^\dagger W \boldsymbol{N} = \beta_{\mathcal{Q}} - \beta + R, \tag{2.43}$$

where R is a formal polynomial containing only monomials of local degree strictly greater than one. A valid SOS decomposition corresponds to enforcing the *consistency condition*

$$R = 0. \tag{2.44}$$

Building on the variational method, one can enforce (2.44) by appropriately choosing the measurement parameters. This leads to the following SOS-based construction procedure for Bell expressions:

1. Choose a target quantum state $|\psi\rangle$ and construct a set of nullifiers $\{\hat{N}_i\}_{i \in I}$ such that

$$\hat{N}_i |\psi\rangle = 0 \quad \forall i \in I.$$

2. Parametrize the measurement operators $\hat{M}^{(j)}_{a_j|x_j}$ for each party.

3. Express the nullifier operators $\hat{N}_i$s in terms of these measurement operators, and define the associated formal polynomials N_is by replacing $\hat{M}^{(j)}_{a_j|x_j}$ with symbolic noncommuting indeterminates.

4. Fixed $W \in \mathbb{R}^{I,I}$, s.t. $W \succeq 0$, compute the polynomial $\boldsymbol{N}^\dagger W \boldsymbol{N}$ and identify the Bell polynomial β and scalar $\beta_{\mathcal{Q}}$ such that (2.43) holds.

5. Impose the condition $R = 0$ to eliminate terms of local degree greater than one.

When successful, this method yields both a Bell expression and an exact quantum bound $\beta_{\mathcal{Q}}$, algebraically certified by the SOS decomposition.

Example 2.4. We illustrate the SOS method using the CHSH inequality evaluated on the singlet state $|\psi\rangle = |\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ [37, 70, 118]. This state is maximally entangled and satisfies

$$\hat{Z}_A \left|\phi^+\right\rangle = \hat{Z}_B \left|\phi^+\right\rangle, \qquad \hat{X}_A \left|\phi^+\right\rangle = \hat{X}_B \left|\phi^+\right\rangle.$$

Step 1: Nullifiers. Choose the following nullifiers, which annihilate $|\phi^+\rangle$:

$$\hat{N}_0 \coloneqq \hat{Z}_A - \hat{Z}_B, \qquad \hat{N}_1 \coloneqq \hat{X}_A - \hat{X}_B.$$

These reflect the symmetry of the state under exchange of local observables.

Step 2: Measurement parametrization. Use the standard CHSH-optimal observables:

$$\hat{A}_0 \coloneqq \frac{1}{\sqrt{2}}(\hat{Z}_A + \hat{X}_A), \qquad \hat{A}_1 \coloneqq \frac{1}{\sqrt{2}}(\hat{Z}_A - \hat{X}_A),$$
$$\hat{B}_0 \coloneqq \hat{Z}_B, \qquad \hat{B}_1 \coloneqq \hat{X}_B.$$

To express the nullifiers in terms of these observables, write:

$$\hat{N}_0 = \frac{1}{\sqrt{2}}(\hat{A}_0 + \hat{A}_1) - \hat{B}_0, \qquad \hat{N}_1 = \frac{1}{\sqrt{2}}(\hat{B}_0 - \hat{B}_1) - \hat{B}_1.$$

Step 3: Promotion to formal polynomials. Promote operators to noncommuting indeterminates:

$$\hat{A}_x \mapsto A_x, \qquad \hat{B}_y \mapsto B_y.$$

The associated formal polynomials N_0, N_1 are:

$$N_0 = \frac{A_0 + A_1}{\sqrt{2}} - B_0, \qquad N_1 = \frac{A_0 - A_1}{\sqrt{2}} - B_1.$$

Step 4: SOS decomposition. Then, choosing $W = \mathbb{1}_2$, the SOS decomposition reads

$$N_0^\dagger N_0 + N_1^\dagger N_1 = \left(\frac{A_0 + A_1}{\sqrt{2}} - B_0\right)^2 + \left(\frac{A_0 - A_1}{\sqrt{2}} - B_1\right)^2 = 4\left(1 - \frac{\beta_{\mathrm{CHSH}}}{2\sqrt{2}}\right),$$

where we used the facts that $A_x^2 = B_y^2 = 1$, $[A_0, A_1] \neq 0$, $[B_0, B_1] \neq 0$ and

$$\beta_{\mathrm{CHSH}} = A_0B_0 + A_0B_1 + A_1B_0 - A_1B_1.$$

This is, up to normalization of the nullifiers, exactly of the form

$$\sum_i N_i^\dagger N_i = \beta_{\mathcal{Q}} - \beta,$$

with $\beta_{\mathcal{Q}} = 1$ and $\beta = \beta_{\mathrm{CHSH}}/2\sqrt{2}$.

Step 5: Verifying the SOS condition. All higher-order terms cancel exactly, so the condition $R = 0$ is satisfied. Hence, this is a valid SOS decomposition certifying the quantum bound.

Conclusion. The method confirms that

$$\beta_{\text{CHSH}} \preceq 2\sqrt{2},$$

and the nullifiers N_i vanish on $|\psi\rangle$ under the given measurements. Thus, any implementation reaching the quantum value must also nullify $|\psi\rangle$ with respect to these operators. This confirms that $|\phi^+\rangle$ achieves the maximal violation and is self-tested by this expression.

Remark 2.5. We highlight several important features of the SOS method:

- **Formal nullifiers.** Step 3 involves promoting operator-valued nullifiers $\hat{N}_i$ to formal polynomials N_i by replacing the measurement operators with symbolic noncommuting indeterminates. While the choice of $\hat{N}_i$ depends only on the target state, this mapping to N_i depends on the measurement parametrization. The resulting polynomials are referred to as *formal nullifiers* and, when evaluated on the chosen measurement operators, yield operators that nullify the target state.

- **Sufficiency of the SOS condition.** The SOS condition $R = 0$ is a *sufficient* criterion for certifying that a Bell expression is maximized by the target state. However, failure to satisfy the condition for a specific choice of $\hat{N}_i$ and measurements does not rule out the existence of a valid SOS decomposition for other choices. This flexibility allows one to use incomplete measurement parametrizations when full generality is impractical, especially in high-dimensional Hilbert spaces.

- **Structure of the formal nullifier space.** The set of nullifiers acting on a finite-dimensional Hilbert space is finite-dimensional, but their formal polynomial representations are not: each operator can be expressed in infinitely many ways in terms of measurement operators, and formal polynomials can have arbitrary monomial length. Thus, the space of formal nullifiers is infinite-dimensional unless one restricts to monomials of bounded length, in which case the parameter space becomes finite.

- **Asymptotic completeness.** The SOS method is known to be dual to the NPA hierarchy [70], which converges in the asymptotic limit. When one considers all formal nullifiers of monomial length up to n and fully parametrizes the measurements, the SOS condition becomes *necessary and sufficient* in the limit $n \to \infty$. In this sense, the method is asymptotically complete: a Bell expression is maximized by a target state if and only if there exists an SOS decomposition with $R = 0$ in this limit.

- **Relation to the variational method.** Unlike the variational method, which does not guarantee global optimality, the SOS method does. In fact, any solution of the SOS method satisfies the variational conditions as well: if $\beta_{\mathcal{Q}} - \beta = \boldsymbol{N}\dagger W \boldsymbol{N}$, then in the target implementation one obtains $\beta_{\mathcal{Q}}\mathbb{1} - \hat{S} = \hat{\boldsymbol{N}}^\dagger W \hat{\boldsymbol{N}}$, and

$$\hat{S}\,|\psi\rangle = \beta_{\mathcal{Q}}\,|\psi\rangle\,.$$

Thus, the state $|\psi\rangle$ is an eigenvector of $\hat{S}$ with maximal eigenvalue, and all first-order variations vanish. This observation can be used to restrict the search space for nullifiers and measurement settings by first applying the variational method.

2.3 Bell expressions maximized by the Tsirelson point

We are now ready to characterize all Bell expressions that attain their quantum maximum at the Tsirelson point [52].

In general, any two-party, two-setting, two-outcome Bell expression can be parametrized by eight real coefficients α_x, β_y, and γ_{xy} for $x, y \in \{0, 1\}$, and written as a formal polynomial in the observables:

$$\beta = \alpha_0 A_0 + \alpha_1 A_1 + \beta_0 B_0 + \beta_1 B_1 + \gamma_{00} A_0 B_0 + \gamma_{10} A_1 B_0 + \gamma_{01} A_0 B_1 + \gamma_{11} A_1 B_1. \tag{2.45}$$

Our goal is to identify constraints on the coefficients that ensure that

$$\boldsymbol{\beta} \cdot \boldsymbol{P}_T = 1 \tag{2.46}$$

is satisfied. To this end, we adopt the variational method and construct the corresponding Bell operator $\hat{S}$ using the standard CHSH-optimal observables:

$$\begin{aligned} \hat{A}_0 &\coloneqq \frac{1}{\sqrt{2}}(\hat{Z}_A + \hat{X}_A), & \hat{A}_1 &\coloneqq \frac{1}{\sqrt{2}}(\hat{Z}_A - \hat{X}_A), \\ \hat{B}_0 &\coloneqq \hat{Z}_B, & \hat{B}_1 &\coloneqq \hat{X}_B. \end{aligned} \tag{2.47}$$

In terms of these observables, the Bell operator associated with Eq. (2.45) can be written as:

$$\hat{S} = p_1 \hat{Z}_A + p_2 \hat{X}_A + p_3 \hat{Z}_B + p_4 \hat{X}_B + p_5 \hat{Z}_A \hat{Z}_B + p_6 \hat{X}_A \hat{X}_B + p_7 \hat{Z}_A \hat{X}_B + p_8 \hat{X}_A \hat{Z}_B, \tag{2.48}$$

where the coefficients p_r are linear combinations of the original parameters α_x, β_y, and γ_{xy}.

Since $\hat{S}$ is a real Hermitian operator, all of its eigenvalues are real. The condition

$$\boldsymbol{\beta} \cdot \boldsymbol{P}_T = \left\langle \phi^+ \right| \hat{S} \left| \phi^+ \right\rangle = 1 \tag{2.49}$$

implies that the singlet state $\left|\phi^+\right\rangle$ is an eigenvector of $\hat{S}$ with eigenvalue 1:

$$\hat{S} \left|\phi^+\right\rangle = \left|\phi^+\right\rangle. \tag{2.50}$$

This requirement leads to the following linear constraints on the parameters p_r:

$$\begin{cases} p_1 + p_3 = 0, \\ p_2 + p_4 = 0, \\ p_5 + p_6 = 1, \\ p_7 - p_8 = 0. \end{cases} \tag{2.51}$$

Using these constraints, we can rewrite the Bell expression (in the formal polynomial representation) in a more structured form:

$$\begin{aligned}\beta \coloneqq\ & r_0\left(\frac{A_0+A_1}{\sqrt{2}}-B_0\right)+r_1\left(\frac{A_0-A_1}{\sqrt{2}}-B_1\right)\\ & +r_2\left(\frac{A_0+A_1}{\sqrt{2}}B_1+\frac{A_0-A_1}{\sqrt{2}}B_0\right)+\lambda\frac{A_0+A_1}{\sqrt{2}}B_0+(1-\lambda)\frac{A_0-A_1}{\sqrt{2}}B_1,\end{aligned} \tag{2.52}$$

where $r_0, r_1, r_2, \lambda \in \mathbb{R}$ are free real parameters.

2.3.1 Variational conditions

In order to ensure that $\boldsymbol{P}_T$ is a local maximum of the function $\boldsymbol{\beta}\cdot\boldsymbol{P}_{\theta,a_x,b_y}$ over the set of quantum correlations, we consider the parametrized quantum realization:

$$|\phi_\theta\rangle \coloneqq c_\theta\,|00\rangle + s_\theta\,|11\rangle\,, \qquad \theta\in[0,\pi/4], \tag{2.53a}$$

$$\hat{A}_x \coloneqq c_{a_x}\hat{Z}_A + s_{a_x}\hat{X}_A, \qquad a_x\in[0,2\pi), \tag{2.53b}$$

$$\hat{B}_y \coloneqq c_{b_y}\hat{Z}_B + s_{b_y}\hat{X}_B, \qquad b_y\in[0,2\pi), \tag{2.53c}$$

where $c_\varphi \coloneqq \cos(\varphi)$, $s_\varphi \coloneqq \sin(\varphi)$. Note that this implementation only gives local probabilities whenever $\theta = 0$. Therefore, in all the following, we assume $\theta > 0$. This implementation gives the following probability distribution:

$$\boldsymbol{P}_{\theta,a_x,b_y} = \begin{array}{c|c|c} 1 & c_{2\theta}c_{b_1} & c_{2\theta}c_{b_2} \\ \hline c_{2\theta}c_{a_1} & \multicolumn{2}{c}{\multirow{2}{*}{$c_{a_x}c_{b_y} + s_{2\theta}s_{a_x}s_{b_y}$}} \\ \cline{1-1} c_{2\theta}c_{a_2} & \multicolumn{2}{c}{} \end{array}, \quad \theta, a_x, b_y \in \mathbb{R}. \tag{2.54}$$

The requirement that $\boldsymbol{P}_T$ is a stationary point under small variations implies the vanishing of the gradient:

$$\boldsymbol{\beta}\cdot\frac{\partial\boldsymbol{P}_{\theta,a_x,b_y}}{\partial\theta}\Big|_{\boldsymbol{P}=\boldsymbol{P}_T} = \boldsymbol{\beta}\cdot\frac{\partial\boldsymbol{P}_{\theta,a_x,b_y}}{\partial a_x}\Big|_{\boldsymbol{P}=\boldsymbol{P}_T} = \boldsymbol{\beta}\cdot\frac{\partial\boldsymbol{P}_{\theta,a_x,b_y}}{\partial b_y}\Big|_{\boldsymbol{P}=\boldsymbol{P}_T} = 0. \tag{2.55}$$

This condition yields the following constraints:

$$\lambda = \frac{1}{2}, \qquad r_2 = 0. \tag{2.56}$$

2.3.2 Characterization of the Bell family

The previous analysis restricts the Bell expressions to a two-parameter family:

$$\beta_{r_0,r_1} \coloneqq r_0\left(\frac{A_0+A_1}{\sqrt{2}}-B_0\right)+r_1\left(\frac{A_0-A_1}{\sqrt{2}}-B_1\right)+\frac{1}{2\sqrt{2}}\beta_{\mathrm{CHSH}}, \tag{2.57}$$

where $r_0, r_1 \in \mathbb{R}$ and $\beta_{\mathrm{CHSH}} \coloneqq (A_0+A_1)B_0+(A_0-A_1)B_1$ is the standard CHSH Bell polynomial.

As shown in [52], further refinement through higher-order perturbations can be employed to reduce the admissible region in the (r_0, r_1)-plane. However, a more immediate and robust constraint is obtained by enforcing that the local bound of the expression is less than or equal to 1. Indeed, any Bell expression β_{r_0,r_1} whose local bound exceeds 1 necessarily admits a quantum value larger than 1, and hence cannot be maximized by the Tsirelson point.

2.4 Local bounds

We then proceed to characterize the dual of the local set, denoted $\mathcal{L}^*$. As we have seen in Sec. 1.4.6, in the CHSH scenario $\mathcal{L}$ is a compact convex polytope whose extremal points correspond to the 16 deterministic strategies. That is,

$$\mathcal{L} = \mathrm{conv}\{\boldsymbol{D}_{ijkl} \mid i,j,k,l \in \{-1,1\}\}, \tag{2.58}$$

where the deterministic points $\boldsymbol{D}_{ijkl}$ are defined by[8]

$$\boldsymbol{D}_{ijkl} \coloneqq \begin{array}{c|c|c} 1 & i & j \\ \hline k & ik & jk \\ \hline l & il & jl \end{array}, \qquad i,j,k,l \in \{-1,1\}. \tag{2.59}$$

The dual local set $\mathcal{L}^*$ is then defined as the set of all Bell functionals $\boldsymbol{\beta}$ that are non-violated by any local distribution. Formally:

$$\mathcal{L}^* \coloneqq \left\{\boldsymbol{\beta} \in \mathbb{R}^8 \colon \boldsymbol{\beta} \cdot \boldsymbol{P} \leq 1 \quad \forall \boldsymbol{P} \in \mathcal{L}\right\}. \tag{2.60}$$

Equivalently, since $\mathcal{L}$ is a polytope determined by its finite set of extremal points, $\mathcal{L}^*$ is the intersection of the half-spaces defined by the inequalities:

$$\boldsymbol{\beta} \cdot \boldsymbol{D}_{ijkl} \leq 1, \qquad \forall i,j,k,l \in \{-1,1\}. \tag{2.61}$$

In our setting, we restrict attention to the two-parameter family β_{r_0,r_1} defined in Eq. (2.57), which forms a linear slice of the full space of Bell functionals. The intersection of this slice with $\mathcal{L}^*$ then defines a convex region in the (r_0, r_1) plane —namely, the set of Bell expressions that maximize the Tsirelson point whose local bound is less than or equal to 1. That region is defined as the intersection of the half-spaces described by

$$\boldsymbol{\beta}_{r_0,r_1} \cdot \boldsymbol{D}_{ijkl} \leq 1, \qquad \forall; i,j,k,l \in \{-1,1\}. \tag{2.62}$$

This region is a regular octagon, whose eight vertices are given by [52]

$$\left\{\left(1 - \frac{1}{\sqrt{2}}\right) R_{\frac{\pi}{4}}^k (1,0) \qquad k = 0,\ldots,7\right\}, \tag{2.63}$$

where $R_{\frac{\pi}{4}}$ denotes a counterclockwise rotation by $\pi/4$ in the (r_0, r_1) plane.

Any point (r_0, r_1) lying outside this region corresponds to a Bell inequality whose local bound exceeds 1, and hence cannot be saturated by $\boldsymbol{P}_T$. Therefore, the region of Bell expressions β_{r_0,r_1} that are maximized by the Tsirelson point is precisely the interior of this octagon.

Remark 2.6. We remark that the inclusion

$$\mathcal{Q}^* \subseteq \mathcal{L}^*$$

holds, since every quantum correlation is, in particular, a nonlocal correlation that also respects all local constraints.

[8] In the CHSH scenario, outcomes are taken in $\{-1,1\}$ (and not in $\{0,1\}$) to match the spectrum of the associated unitary observables $\hat{A}_x$, $\hat{B}_y$. In fact, deterministic strategies assign fixed values in $\{-1,1\}$ to each local observable: $A_0 = i$, $A_1 = j$, $B_0 = k$, $B_1 = l$, with $i,j,k,l \in \{-1,1\}$.

Proof. . Since $\mathcal{L} \subseteq \mathcal{Q}$, it follows that every $\boldsymbol{P} \in \mathcal{L}$ is also an element of $\mathcal{Q}$. Therefore, any Bell expression $\boldsymbol{\beta}$ that satisfies $\boldsymbol{\beta} \cdot \boldsymbol{P} \leq 1$ for all $\boldsymbol{P} \in \mathcal{Q}$ in particular satisfies the same for all $\boldsymbol{P} \in \mathcal{L}$. Hence:

$$\boldsymbol{\beta} \in \mathcal{Q}^* \quad \iff \quad \boldsymbol{\beta} \cdot \boldsymbol{P} \leq 1 \ \forall \boldsymbol{P} \in \mathcal{Q} \quad \implies \quad \boldsymbol{\beta} \cdot \boldsymbol{P} \leq 1 \ \forall \boldsymbol{P} \in \mathcal{L} \quad \implies \quad \boldsymbol{\beta} \in \mathcal{L}^*.$$

This proves that $\mathcal{Q}^* \subseteq \mathcal{L}^*$.

□

2.5 Quantum bounds

Having identified the region of Bell expressions β_{r_0,r_1} that are not violated by any local strategy (i.e., those lying inside the local dual polytope $\mathcal{L}^*$), we now aim to determine which of these expressions are also quantum-valid, i.e., upper bounded by 1 over all quantum correlations. Specifically, we seek to characterize those Bell expressions for which the Tsirelson bound $\beta_{\mathcal{Q}}$ is equal to 1, and saturated by the Tsirelson realization.

To formalize this, let us consider the algebra of formal indeterminates $\mathcal{R}$ generated by arbitrary products of the measurement operators A_x and B_y for $x, y \in \{0, 1\}$, subject to the relations

$$A_x^2 = B_y^2 = 1, \qquad [A_x, B_y] = 0. \tag{2.64}$$

We interpret the representation of the Bell expression in indeterminates space β as an element of $\mathcal{R}$.

As previously seen, a general method to upper bound the quantum value of β is to search for a sum-of-squares (SOS) decomposition of the operator $1 - \beta$. That is, a sufficient condition for $\beta \preceq 1$ is the existence of a decomposition of the form

$$1 - \beta = \sum_s O_s^\dagger O_s, \qquad O_s \in \mathcal{R}. \tag{2.65}$$

If such a decomposition exists, then for any quantum state $|\psi\rangle$ and observables implementing A_x and B_y in accordance with $\mathcal{R}$, it follows that

$$\boldsymbol{\beta} \cdot \boldsymbol{P} = \langle\psi| \hat{S} |\psi\rangle \leq 1.$$

However, the algebra $\mathcal{R}$ is infinite-dimensional, making the full SOS search intractable in practice. To address this, one typically considers a finite-dimensional relaxation $\mathcal{T} \subset \mathcal{R}$, such as the subspace of all polynomials in A_x and B_y of degree at most d. Any SOS decomposition with operators $O_s \in \mathcal{T}$ still provides a valid upper bound on the quantum value of β.

Since we are looking for Bell expressions that are maximized by the Tsirelson behavior, we can look at nullifiers of the state $\left|\phi^+\right\rangle$ and consider the subspace

$$\mathcal{N}_{\mathcal{T}} \coloneqq \left\{ O \in \mathcal{T} \colon \hat{O} \left|\phi^+\right\rangle = 0 \right\}. \tag{2.66}$$

Let us now fix a finite relaxation $\mathcal{T}$ and consider a generating set $\{N_i\}$ for the nullifier subspace $\mathcal{N}_{\mathcal{T}}$. We denote by $\boldsymbol{N}$ the vector whose components are the generators N_i. From what

we have seen in Sec. 2.2.2, the problem of verifying whether our Bell expression β satisfies $\beta \preceq 1$ reduces to checking the existence of a positive semidefinite matrix W such that

$$1 - \beta = \boldsymbol{N}^\dagger W \boldsymbol{N}. \tag{2.67}$$

For a finite relaxation $\mathcal{T}$, we consider a generating sequence $\boldsymbol{N}_\mathcal{T}$ of $\mathcal{N}_\mathcal{T}$. Given any symmetric matrix $W = (w_{kl})_{k,l}$, the condition (2.67) imposes linear conditions on the coefficients w_{kl} and the positivity constraint $W \succeq 0$ remains to be checked. Different relaxations $\mathcal{T}$ could be considered here:

1. The first order relaxation $\mathcal{T}_{1+A+B} \coloneqq \{1, A_0, A_1, B_0, B_1\}$ only gives a certificate for the CHSH inequality.

2. The relaxation at the almost quantum level [119],

$$\mathcal{T}_{1+A+B+AB} \coloneqq \mathcal{T}_{1+A+B} \cup \{A_x B_y,\ x, y \in \{0,1\}\} \tag{2.68}$$

   can also be computed numerically and gives a certificate for a disk in the (r_0, r_1) plane of center $(0,0)$ and radius $\frac{1}{4\sqrt{2}}$.

3. The next relaxation is given by

$$\mathcal{T}_{1+A+B+AB+ABB'} \coloneqq \mathcal{T}_{1+AB} \cup \{A_x B_y B_{y'}, y \neq y'\}. \tag{2.69}$$

   We can show analytically (see Sec. 2.5.1) that a certificate can be found for the inequality $\beta_{1-1/\sqrt{2},0}$ at this level of relaxation. This ensures that the Bell expression

$$\beta_T \coloneqq \beta_{1-1/\sqrt{2},0} = \left(1 - \frac{1}{\sqrt{2}}\right)\left(\frac{A_0 + A_1}{\sqrt{2}} - B_0\right) + \frac{\beta_{\text{CHSH}}}{2\sqrt{2}} \tag{2.70}$$

   is maximized by the Tsirelson point.

One can check that this inequality is an extremal point of $\mathcal{Q}^*$, that it is exposed, in the sense that the quantum set admits a point which only saturates this Tsirelson inequality, and that it is only maximized by 3 extremal points of $\mathcal{Q}$: $\boldsymbol{P}_T$ and two deterministic realizations.

2.5.1 Certificate for the T_{ABB} relaxation

Now we focus on the finite relaxation $T_{ABB} \coloneqq \mathcal{T}_{1+A+B+AB+ABB'}$. The polynomials that nullify the state $|\phi^+\rangle$ in the ideal implementation (2.12) at this level of relaxation are linear combinations

of nine operators and the generating sequence can be chosen to be

$$\boldsymbol{N}_{T_{ABB}} \coloneqq \begin{pmatrix} N_1 \\ N_2 \\ N_3 \\ N_4 \\ N_5 \\ N_6 \\ N_7 \\ N_8 \\ N_9 \end{pmatrix} = \begin{pmatrix} \frac{A_0+A_1}{\sqrt{2}} - B_0 \\ \frac{A_0-A_1}{\sqrt{2}} - B_1 \\ 1 - \frac{A_0+A_1}{\sqrt{2}} B_0 \\ 1 - \frac{A_0-A_1}{\sqrt{2}} B_1 \\ \frac{A_0+A_1}{\sqrt{2}} B_1 + \frac{A_0-A_1}{\sqrt{2}} B_0 \\ B_1 \left(1 - \frac{A_0+A_1}{\sqrt{2}} B_0\right) \\ B_0 \left(1 - \frac{A_0-A_1}{\sqrt{2}} B_1\right) \\ \left(1 + \frac{A_0+A_1}{\sqrt{2}} B_0\right) B_1 \\ \left(1 + \frac{A_0-A_1}{\sqrt{2}} B_1\right) B_0 \end{pmatrix}. \tag{2.71}$$

One verifies that the symmetric matrix

$$W_6 \coloneqq \frac{1}{16} \begin{pmatrix} \sqrt{2} & -2s & -s & & & \\ -2s & 2s & 0 & & & \\ -s & 0 & \sqrt{2} & & & \\ & & & 2 & 0 & s \\ & & & 0 & \sqrt{2}s & s \\ & & & s & s & s \end{pmatrix}, \tag{2.72}$$

expressed in the basis $\{N_1, N_3, N_7, N_2, N_6, N_5\}$, with $s \coloneqq 2 - \sqrt{2}$, and defining $\boldsymbol{N}_6 \coloneqq (N_1, N_3, N_7, N_2, N_6, N_5)^T$, fulfills the condition

$$1 - \beta_T = \boldsymbol{N}_6^\dagger W_6 \boldsymbol{N}_6. \tag{2.73}$$

Furthermore, this matrix has 4 non-zero eigenvalues:

$$\sigma(W_6) = \frac{1}{8} \left\{ 1 - \sqrt{10 - 7\sqrt{2}}, 1 + \sqrt{10 - 7\sqrt{2}}, 2 - \sqrt{2}, \frac{3\sqrt{2}}{2} - 1, 0, 0 \right\}. \tag{2.74}$$

Since all 4 are positive, $W_6 \succeq 0$ and provides a *certificate* for the inequality β_T.

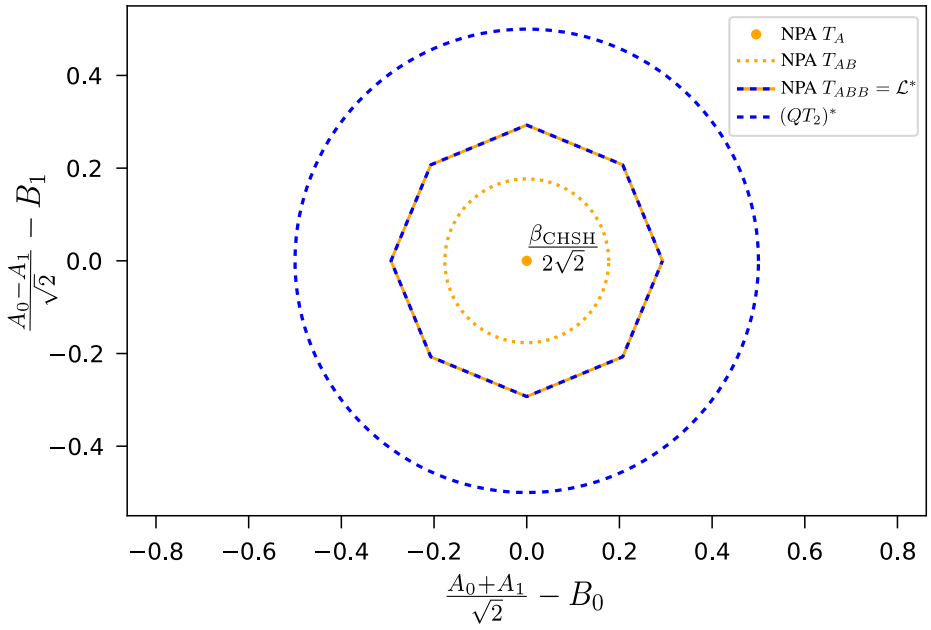

Figure 2.2: Summary of the bounds on the dual quantum set $\mathcal{Q}^*$ in the (r_0, r_1) slice of Bell expressions β_{r_0,r_1} considered in the main text. Blue lines represent outer approximations of $\mathcal{Q}^*$: the dashed blue octagon corresponds to the region of inequalities with local bound less than or equal to 1 (i.e., the local dual set $\mathcal{L}^*$), while the dotted blue circle marks the maximal radius around the Tsirelson point $\boldsymbol{P}_T$ for which second-order variations in two-dimensional quantum realizations do not increase the Bell value—any inequality outside this circle lies outside $\mathcal{Q}^*$. Orange lines show inner certificates of membership in $\mathcal{Q}^*$, obtained via SOS decompositions using the NPA hierarchy: the orange dot, dotted orange circle, and dashed orange octagon correspond to relaxations $T_A \coloneqq \mathcal{T}_{1+A+B}$, $T_{AB} \coloneqq \mathcal{T}_{1+A+B+AB}$, and $T_{ABB} \coloneqq \mathcal{T}_{1+A+B+AB+ABB'+AA'B}$, respectively. The certification for T_{AB} was obtained numerically. The coincidence of the blue and orange octagons implies that, in this slice, the quantum dual set $\mathcal{Q}^*$ exactly coincides with the local dual set $\mathcal{L}^*$. From [52].

2.5.2 Exploiting symmetries

To fully characterize the dual quantum set $\mathcal{Q}^*$ within the slice defined by the Bell expressions β_{r_0,r_1}, it is useful to exploit the discrete symmetries preserved by both the family of expressions β_{r_0,r_1} and the quantum set $\mathcal{Q}$. One such symmetry is described by the transformation

$$S\colon (r_0, r_1) \mapsto R_{\frac{\pi}{4}}(r_0, r_1), \quad \begin{cases} A_0 \mapsto -B_1, & A_1 \mapsto -B_0, \\ B_0 \mapsto -A_0, & B_1 \mapsto A_1, \end{cases} \tag{2.75}$$

where $R_{\frac{\pi}{4}}$ denotes a rotation by an angle $\frac{\pi}{4}$ in the (r_0, r_1) plane. This symmetry maps Bell expressions within the family β_{r_0,r_1} to others in the same family, while preserving the structure of the quantum set.

As a consequence, the quantum bound of a Bell expression with parameters (r_0, r_1) is identical to that of its rotated image under S. In particular, the eight Bell expressions of the form

$$S^k \cdot \beta_T \quad (\text{ for } k = 0, \ldots, 7), \tag{2.76}$$

corresponding to the vertices of the regular octagon with polar coordinates $(1 - \frac{1}{\sqrt{2}}, k\frac{\pi}{4})$, all achieve a quantum value of 1. This shows that the full octagon constitutes the set of Bell expressions within the slice that are maximized by the Tsirelson point $\boldsymbol{P}_T$ and have quantum value equal to 1. Thus, the interior of the octagon precisely corresponds to the dual face

$$\mathcal{Q}^* \cap \mathrm{aff}(\beta_{r_0,r_1})$$

(see fig. 2.3).

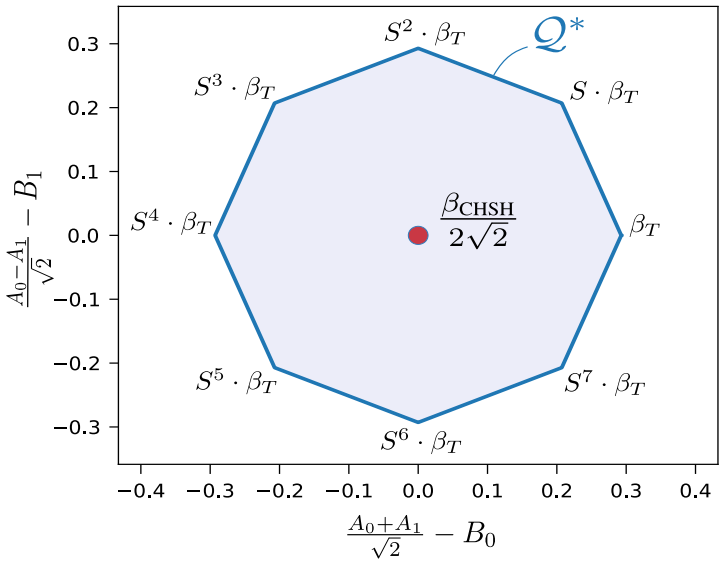

Figure 2.3: Dual face of the quantum set $\mathcal{Q}^*$ in the two-dimensional affine slice spanned by the Bell expressions β_{r_0,r_1} with real coefficients r_0, r_1. The eight vertices of the octagon correspond to the Tsirelson inequalities $S^k \cdot \boldsymbol{\beta}_T$, which are extremal in $\mathcal{Q}^*$. The red point at the center represents the normalized CHSH expression, which is not extremal and can be written as a convex combination of opposite vertices of the octagon. From [52].

2.6 Conclusions

2.6.1 Dual quantum set

An interesting feature of this dual face is that the CHSH expression lies at its center, indicating that it is not an extremal point of $\mathcal{Q}^*$. In fact, it admits multiple convex decompositions in terms of extremal Tsirelson inequalities. One such decomposition is given by

$$\beta_{\mathrm{CHSH}} = \frac{1}{2}\left(\beta_T + S^4 \cdot \beta_T\right). \tag{2.77}$$

This decomposition is not unique, as β_{CHSH} lies in the relative interior of a two-dimensional face of $\mathcal{Q}^*$. Geometrically, this means that the CHSH inequality lies on a line segment joining any pair of opposite extremal points of the octagon.

Moreover, any inequality inside the octagon can be expressed as a convex combination of the CHSH expression and a boundary inequality β_b lying on the octagon

$$\beta = p\frac{\beta_{\text{CHSH}}}{2\sqrt{2}} + (1-p)\beta_b, \qquad p \in (0,1]. \tag{2.78}$$

Remark 2.7. This implies that if a quantum correlation $\boldsymbol{P}$ achieves the maximal value $\boldsymbol{\beta}\cdot\boldsymbol{P} = 1$, then it must also satisfy $\boldsymbol{\beta}_{\text{CHSH}} \cdot \boldsymbol{P} = 2\sqrt{2}$. By the known self-testing properties of the CHSH inequality, this uniquely identifies $\boldsymbol{P} = \boldsymbol{P}_T$. Therefore, any Bell expression in the interior of the octagon self-tests the Tsirelson realization.

Proof. We start from:

$$\boldsymbol{\beta} = p\frac{\boldsymbol{\beta}_{\text{CHSH}}}{2\sqrt{2}} + (1-p)\boldsymbol{\beta}_b, \quad p \in (0,1].$$

Now let $\bar{\boldsymbol{P}} \in \mathcal{Q}$ be such that $\boldsymbol{\beta} \cdot \bar{\boldsymbol{P}} = 1$. Compute the scalar product

$$1 = \boldsymbol{\beta} \cdot \bar{\boldsymbol{P}} = p\frac{\boldsymbol{\beta}_{\text{CHSH}}}{2\sqrt{2}} \cdot \bar{\boldsymbol{P}} + (1-p) \cdot \boldsymbol{\beta}_b \cdot \bar{\boldsymbol{P}}. \tag{2.79}$$

Since $\boldsymbol{\beta}_b$ lies on the boundary of the octagon, we know that

$$\boldsymbol{\beta}_b \cdot \boldsymbol{P} \leq 1, \quad \text{with equality only if } \boldsymbol{P} = \boldsymbol{P}_T.$$

Substituting this into Eq. (2.79):

$$1 = \boldsymbol{\beta} \cdot \bar{\boldsymbol{P}} \leq p\frac{\boldsymbol{\beta}_{\text{CHSH}}}{2\sqrt{2}} \cdot \bar{\boldsymbol{P}} + (1-p) \cdot 1.$$

Subtracting $(1-p)$ on both sides,

$$p\left(\frac{\boldsymbol{\beta}_{\text{CHSH}}}{2\sqrt{2}} \cdot \bar{\boldsymbol{P}} - 1\right) \geq 0.$$

Since $p > 0$, this implies

$$\frac{\boldsymbol{\beta}_{\text{CHSH}}}{2\sqrt{2}} \cdot \bar{\boldsymbol{P}} \geq 1 \quad \Longrightarrow \quad \boldsymbol{\beta}_{\text{CHSH}} \cdot \bar{\boldsymbol{P}} \geq 2\sqrt{2}.$$

But $\boldsymbol{P}$ is a quantum correlation, and Tsirelson's bound tells us that

$$\boldsymbol{\beta}_{\text{CHSH}} \cdot \bar{\boldsymbol{P}} \leq 2\sqrt{2}.$$

Hence,

$$\boldsymbol{\beta}_{\text{CHSH}} \cdot \bar{\boldsymbol{P}} = 2\sqrt{2}.$$

By the self-testing property of the CHSH inequality, the only quantum point that achieves this is the Tsirelson point. Thus,

$$\bar{\boldsymbol{P}} = \boldsymbol{P}_T.$$

This means that any Bell expression $\boldsymbol{\beta}$ strictly inside the octagon self-tests the Tsirelson realization. □

In contrast, the Bell expressions on the boundary of the octagon (i.e., the *extremal Tsirelson inequalities*) are also maximized by local deterministic strategies and thus cannot uniquely identify the quantum realization. As a result, they do not provide self-testing of the Tsirelson state, despite being extremal in the dual set.

2.6.2 Primal quantum set

The geometric picture obtained for the dual quantum set $\mathcal{Q}^*$ translates, via convex duality, into this local description of the quantum set $\mathcal{Q}$ near $\boldsymbol{P}_T$. Specifically, each Bell expression $\boldsymbol{\beta}$ satisfying

$$\boldsymbol{\beta} \cdot \boldsymbol{P}_T = 1 \quad \text{and} \quad \boldsymbol{\beta} \in \mathcal{Q}^*$$

defines a supporting hyperplane of $\mathcal{Q}$ at $\boldsymbol{P}_T$. As such, the collection of these inequalities provides a precise characterization of the boundary of the quantum set around $\boldsymbol{P}_T$.

The analysis of the dual structure reveals that the Tsirelson point $\boldsymbol{P}_T$ is an *exposed extremal point* of the quantum set $\mathcal{Q}$. In fact, it has a *dimension pair* $(0, 2)$: it is a 0-dimensional extremal point of $\mathcal{Q}$, supported by a 2-dimensional face of the dual set $\mathcal{Q}^*$. This is demonstrated by the fact that the Bell expressions β_{r_0,r_1} satisfying $\boldsymbol{\beta} \cdot \boldsymbol{P}_T = 1$ form a two-dimensional octagon.

The existence of a higher-dimensional supporting face in the dual set implies that $\boldsymbol{P}_T$ is exposed—that is, it is uniquely singled out by a family of Bell inequalities. In this case, every Bell expression in the interior of the octagon defines a supporting hyperplane that touches $\mathcal{Q}$ only at $\boldsymbol{P}_T$, as shown in fig. 2.4.

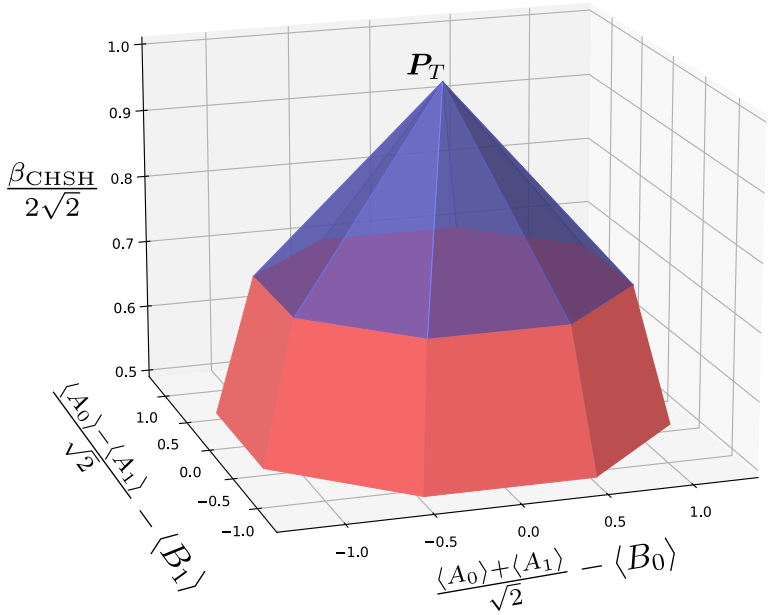

Figure 2.4: Three-dimensional projection of the local polytope $\mathcal{L}$ (shown in red) and of the quantum set of correlations $\mathcal{Q}$ (shown in red and blue), as introduced in [52]. The figure highlights a slice of the correlation space in which both sets are embedded. The only point attaining the maximal value of 1 along the vertical (z-) axis is the Tsirelson point $\boldsymbol{P}_T$, corresponding to the quantum realization that maximally violates the CHSH inequality. This point lies at the apex of a pyramid with octagonal base, whose eight facets represent tight Bell inequalities of the form $S^k \cdot \boldsymbol{\beta}_T = 1$. Each facet corresponds to a supporting hyperplane of the quantum set $\mathcal{Q}$ at the point $\boldsymbol{P}_T$, characterizing the tangent structure of the set at this extremal point.

Chapter 3

Orthogonal faces in the CHSH scenario

In the previous chapter, we reviewed the results of Barizien and Bancal [52], where the orthogonal face of the Tsirelson point—a paradigmatic nonlocal extremal point of the quantum set—was fully characterized. There, the analysis focused on a single point, exploiting its algebraic symmetries to completely identify the supporting hyperplanes.

In this chapter, we extend and generalize that approach to a broader family of nonlocal extremal points in the CHSH scenario. Our goal is to characterize their associated orthogonal faces: that is, the sets of Bell expressions which are maximized by a given extremal point and orthogonal to its local tangent directions. This leads to a refined geometric understanding of the quantum set's boundary and provides a systematic method to construct tight Bell inequalities tailored to specific quantum correlations.

Sec. 3.2 presents a general method for identifying these orthogonal faces, based on nullspace constraints and variational considerations, following [120]. Sec. 3.3 develops this framework by analyzing the Bell expressions that support the extremal points and reduce the problem to a constrained, parameterized form suitable for both symbolic and numerical analysis.

The second part of the chapter contains the original contributions of this thesis. In Sec. 3.5, we compute the local bounds associated with each Bell expression by analyzing the projected dual local polytope. We extract its vertices, provide explicit formulas, and offer visual representations. Sec. 3.6 then addresses the quantum bounds, computed via the NPA hierarchy. We further investigate the dual solutions returned by the semidefinite program and search for SOS decompositions that certify quantum optimality.

Altogether, this chapter advances the characterization of quantum extremality in the CHSH scenario. It generalizes known results and contributes new analytic and numerical techniques for identifying and certifying the boundaries of quantum correlations.

3.1 Characterizing $\mathcal{Q}_{\text{ext}}$ in the CHSH scenario

As discussed in Chapter 1, a key geometric feature of the quantum set $\mathcal{Q}$ is its convexity. This implies that $\mathcal{Q}$ is fully determined by its extremal points:

$$\mathcal{Q} = \text{conv}(\mathcal{Q}_{\text{ext}}).$$

Accordingly, the structure and characterization of the extremal set $\mathcal{Q}_{\text{ext}}$ play a central role in understanding the boundary of quantum correlations. Over the years, various techniques have been developed to constrain and classify the quantum realizations that give rise to these extremal points, providing insight into the fine geometry of $\mathcal{Q}$.

First, since the dimension of the local Hilbert spaces is not fixed a priori, Naimark's dilation theorem (see Appendix A) ensures that it suffices to consider realizations involving pure states and projective (or unitary) measurements. Second, by Jordan's lemma, the problem further reduces to analyzing qubit systems [121]:

Theorem 3.1 (Jordan's Lemma for Extremal Correlations in the CHSH Scenario). *Every extremal quantum correlation in the CHSH scenario can be realized by performing real unitary measurements on a two-qubit pure state.*

More recently, it has been shown that real Hilbert spaces are sufficient to describe all extremal quantum correlations in this scenario [122]. Altogether, these results imply that

$$\mathcal{Q}_{\text{ext}} \subset \mathcal{Q}_2, \tag{3.1}$$

where $\mathcal{Q}_2$ denotes the subset of $\mathcal{Q}$ that can be generated by measuring a partially entangled two-qubit real state with real unitary observables. Concretely, any point in $\mathcal{Q}_2$ admits a realization of the form

$$|\phi_\theta\rangle \coloneqq c_\theta |00\rangle + s_\theta |11\rangle, \tag{3.2a}$$

$$\hat{A}_x \coloneqq c_{a_x} \hat{Z}_A + s_{a_x} \hat{X}_A, \qquad a_x \in [0, 2\pi), \tag{3.2b}$$

$$\hat{B}_y \coloneqq c_{b_y} \hat{Z}_B + s_{b_y} \hat{X}_B, \qquad b_y \in [0, 2\pi), \tag{3.2c}$$

where $c_\alpha \coloneqq \cos(\alpha)$, $s_\alpha \coloneqq \sin(\alpha)$, and the measurement directions lie in the XZ-plane of the Bloch sphere.

The quantum set $\mathcal{Q}$ exhibits intrinsic symmetries [51], such as relabeling of inputs and outputs and exchange of the two parties. Importantly, these symmetries preserve the extremality of points in $\mathcal{Q}$, and act on the measurement parameters (a_0, a_1, b_0, b_1) as a group of involutions. In particular, the quantum set is invariant under the following involutions:

$$\text{Relabeling of parties:} \quad S_{\text{part}} \colon \hat{A}_x \longleftrightarrow \hat{B}_x, \tag{3.3a}$$

$$\text{Relabeling of inputs:} \quad S_{\text{in}} \colon \hat{A}_0 \longleftrightarrow \hat{A}_1, \tag{3.3b}$$

$$\text{Relabeling of outputs:} \quad S_{\text{out}} \colon \hat{A}_0 \longrightarrow -\hat{A}_0. \tag{3.3c}$$

These symmetries induce transformations on the measurement angles a_x, b_y appearing in the real qubit realizations (3.2). In particular:

1. Relabeling of the parties corresponds to swapping the roles of Alice and Bob:
$$a_x \longleftrightarrow b_x.$$
2. Relabeling of the inputs for Alice corresponds to exchanging her measurement angles:
$$a_0 \longleftrightarrow a_1.$$
3. Relabeling of the outputs corresponds to flipping the sign of the measurement outcomes, which, in the real qubit parametrization, amounts to shifting the measurement direction by π:
$$a_0 \longrightarrow a_0 + \pi \mod 2\pi.$$

As a consequence, any realization of the form (3.2) can be transformed via these symmetries into a canonical form within the following parameter range [51]:

$$\theta \in [0, \pi), \qquad 0 \leq a_0 \leq b_0 \leq b_1 < \pi, \qquad a_0 \leq a_1 < \pi. \tag{3.4}$$

This canonical form is unique up to symmetry, and the extremality of a point remains unchanged under such transformations.

A recent result [51] provides a full characterization of the nonlocal extremal points in $\mathcal{Q}$ in terms of the above canonical representation:

Theorem 3.2 (Characterization of $\mathcal{Q}_{\text{ext}}$). *A quantum realization yields a nonlocal extremal point of the set $\mathcal{Q}$ if and only if it can be mapped, via local unitaries and relabelings, to a realization of the form (3.2), where the parameters satisfy the* full alternation condition

$$\forall (s,t) \in \{\pm 1\}^2, \qquad 0 \leq [\tilde{a}_0^s]_\pi \leq b_0 \leq [\tilde{a}_1^t]_\pi \leq b_1 < \pi, \tag{3.5}$$

where

$$\tilde{a}_x^j := 2 \operatorname{atan}\left(\tan\left(\frac{a_x}{2}\right) \tan^j \theta\right),$$

and $[\alpha]_\pi$ denotes reduction modulo π, i.e., $[\alpha]_\pi := \alpha \mod \pi$, bringing α into the interval $[0, \pi)$.

This condition defines a compact, fully characterizable subset of the two-qubit quantum set whose points correspond to the nonlocal extremal correlations in the CHSH scenario. Together with the 16 deterministic strategies [19] which are local extremal points of $\mathcal{Q}$, Theorem 3.2 gives a complete characterization of the extremal point of the quantum set in the minimal scenario. Note that it can be shown that the nonlocal extremal points identified by Theorem 3.2 all exhibit the self-test property, and thus identify a unique underlying qubit realization [51].

Example 3.1. Using Theorem 3.2, one can verify that the Tsirelson point corresponds to a nonlocal extremal point of the quantum set $\mathcal{Q}$. Recall that the Tsirelson behavior arises from a realization of the form (3.2), with the following measurement angles:

$$a_0 = \frac{\pi}{4}, \qquad a_1 = -\frac{\pi}{4}, \qquad b_0 = 0, \qquad b_1 = \frac{\pi}{2}.$$

However, this configuration does not satisfy the canonical ordering defined in Eq. (3.4). To bring the realization into canonical form, we apply a sequence of relabelings that preserve the extremality of the point:

1. Relabeling of inputs (Alice): exchange $a_0 \longleftrightarrow a_1$

$$a_0 = -\frac{\pi}{4}, \qquad a_1 = \frac{\pi}{4}, \qquad b_0 = 0, \qquad b_1 = \frac{\pi}{2}.$$

2. Relabeling of outputs (Alice): flip the outcome of setting $x = 0$, i.e., $a_0 \longrightarrow a_0 + \pi$

$$a_0 = \frac{3\pi}{4}, \qquad a_1 = \frac{\pi}{4}, \qquad b_0 = 0, \qquad b_1 = \frac{\pi}{2}.$$

3. Relabeling of inputs (Alice): exchange $a_0 \longleftrightarrow a_1$ again

$$a_0 = \frac{\pi}{4}, \qquad a_1 = \frac{3\pi}{4}, \qquad b_0 = 0, \qquad b_1 = \frac{\pi}{2}.$$

4. Relabeling of parties: exchange the roles of Alice and Bob, i.e., $a_x \longleftrightarrow b_x$ for all x

$$a_0 = 0, \qquad a_1 = \frac{\pi}{2}, \qquad b_0 = \frac{\pi}{4}, \qquad b_1 = \frac{3\pi}{4}.$$

The resulting configuration now satisfies the canonical ordering

$$0 \leq a_0 \leq b_0 \leq b_1 < \pi, \qquad a_0 \leq a_1 < \pi.$$

It is straightforward to verify that, for $\theta = \pi/4$, the full alternation condition in Theorem 3.2 is satisfied. Thus, the Tsirelson point indeed corresponds to a nonlocal extremal point of $\mathcal{Q}$.

3.2 Orthogonal faces in the CHSH scenario

Given a quantum realization of the form (3.2), the resulting correlation point $\boldsymbol{P}_{\theta,a_x,b_y}$ can be expressed as:

$$\boldsymbol{P}_{\theta,a_x,b_y} = \left(\begin{array}{c|c|c} 1 & c_{2\theta}c_{b_1} & c_{2\theta}c_{b_2} \\ \hline c_{2\theta}c_{a_1} & \multicolumn{2}{c}{\multirow{2}{*}{$c_{a_x}c_{b_y} + s_{2\theta}s_{a_x}s_{b_y}$}} \\ \cline{1-1} c_{2\theta}c_{a_2} & & \end{array}\right), \quad \theta, a_x, b_y \in \mathbb{R}. \tag{3.6}$$

As established in the previous section, this family of behaviors provides a complete parametrization of all nonlocal extremal points of the quantum set $\mathcal{Q}$ in the CHSH scenario.

Our goal is to characterize the orthogonal face of such a point $\boldsymbol{P}_{\theta,a_x,b_y}$, i. e., the set of all Bell expressions $\boldsymbol{\beta}$ that are maximized at $\boldsymbol{P}$, or equivalently, for which $\boldsymbol{P}$ achieves the quantum bound

$$\max_{\boldsymbol{P}\in\mathcal{Q}} \boldsymbol{\beta}\cdot\boldsymbol{P} = 1.$$

This analysis is especially significant because, as previously explained in (2.6) the orthogonal face of any point in $\mathcal{Q}$—whether extremal or not—coincides with the intersection of the orthogonal faces of the extremal points lying on the same face. Therefore, a complete characterization of the orthogonal faces of extremal points provides, in principle, a full description of the facial structure of the quantum set.

3.2.1 General method

We now present the general method, following [120], for deriving necessary conditions that a Bell expression must satisfy to be maximized by an extremal point of the quantum set.

Let $(|\psi\rangle, \{\hat{A}_x\}_x, \{\hat{B}_y\}_y)$ be a fixed quantum realization that produces the correlation point $\boldsymbol{P}$. Define the *measurement vector* $\hat{\boldsymbol{M}}$ as the collection of all observables involved in computing $\boldsymbol{P}$:

$$\hat{\boldsymbol{M}} \coloneqq (\hat{A}_x, \hat{B}_y, \hat{A}_x\hat{B}_y)_{x,y}. \tag{3.7}$$

Then, the correlations produced by this realization can be written as

$$\boldsymbol{P} = \langle\psi|\, \hat{\boldsymbol{M}}\, |\psi\rangle\,. \tag{3.8}$$

and the value of the Bell expression on this point is

$$\boldsymbol{\beta}\cdot\boldsymbol{P} = \boldsymbol{\beta}\cdot\langle\psi|\, \hat{\boldsymbol{M}}\, |\psi\rangle = \langle\psi|\, \hat{S}\, |\psi\rangle\,.$$

For any given Bell expression $\boldsymbol{\beta} \in \mathcal{Q}^*$, one can then construct the Bell operator associated to this measurement choice as

$$\hat{S} = \boldsymbol{\beta}\cdot\hat{\boldsymbol{M}}, \tag{3.9}$$

which is Hermitian.

As we have seen, if the point $\boldsymbol{P}$ is to give the maximal quantum value of $\boldsymbol{\beta}$, then the state $|\psi\rangle$ must be an eigenstate of $\hat{S}$ of maximal eigenvalue. This implies that for all vector $|\xi_i\rangle$ orthogonal to $|\psi\rangle$, we have

$$\langle\xi_i|\, \hat{S}\, |\psi\rangle = \lambda\, \langle\xi_i|\psi\rangle = 0. \tag{3.10}$$

Since $\langle\xi_i|\, \hat{S}\, |\psi\rangle = \boldsymbol{\beta}\cdot\langle\xi_i|\, \hat{\boldsymbol{M}}\, |\psi\rangle$, we obtain

$$\boldsymbol{\beta}\cdot\boldsymbol{T}_i = 0 \tag{3.11}$$

where we denoted

$$\boldsymbol{T}_i \coloneqq \langle\xi_i|\, \hat{\boldsymbol{M}}\, |\psi\rangle \tag{3.12}$$

for any $|\xi_i\rangle$ orthogonal to $|\psi\rangle$. The collection $\{\boldsymbol{T}_i\}_i$, corresponding to a basis of the orthogonal complement of $|\psi\rangle$, defines a set of linear constraints that any Bell expression maximized by $\boldsymbol{P}$ must satisfy.

Suppose now that the measurement observables $\hat{A}_x$, $\hat{B}_y$ depend smoothly on a set of real parameters $\boldsymbol{r} = (r_1, \ldots, r_m)$, such that $\boldsymbol{r} = 0$ corresponds to our realization of interest. Then the correlators $\boldsymbol{P}$ can also be viewed as a smooth function $\boldsymbol{P}(\boldsymbol{r})$. As we have seen in Sec. 2.2.1, a necessary condition for $\boldsymbol{P}$ to locally maximize the Bell expression $\boldsymbol{\beta}$ is that its first-order variation must vanish:

$$\forall j, \qquad \boldsymbol{\beta}\cdot\left.\frac{\partial\boldsymbol{P}}{\partial r_j}\right|_{\boldsymbol{r}=0} = 0. \tag{3.13}$$

Collecting all such linear constraints, we define the subspace

$$V \coloneqq \mathrm{Vect}\left\langle \{\boldsymbol{T}_i\colon\ \langle\xi_i|\psi\rangle = 0\} \cup \left\{ \left.\frac{\partial\boldsymbol{P}}{\partial r_j}\right|_{\boldsymbol{r}=0} \right\}_j \right\rangle. \tag{3.14}$$

and conclude that any Bell functional $\boldsymbol{\beta}$ maximized by $\boldsymbol{P}$ must lie in the orthogonal complement

$$\boldsymbol{\beta} \in V^{\perp}. \tag{3.15}$$

Remark 3.1. This result shows that the orthogonal face $\{\boldsymbol{P}\}^{\perp}$ is contained in the orthogonal complement of a finite-dimensional vector space V, which depends only on the state and the measurement observables. Therefore, the dimension of the face of $\mathcal{Q}$ supported at $\boldsymbol{P}$ is upper bounded by $\dim V$.

Moreover, these conditions provide a powerful tool in applications involving facial structure, such as dimension bounds and semidefinite programming techniques for Bell inequalities [46, 47].

3.3 Bell expressions maximized by the extremal points

We are now ready to derive necessary conditions that any Bell expression must satisfy in order to be maximized by extremal points of the quantum set in the CHSH scenario.[1]

In general, all possible Bell expressions can be parameterized by eight real coefficients α_x, β_y, γ_{xy} for $x, y \in \{0,1\}$ and be written as

$$\beta(\boldsymbol{P}) = \alpha_0\langle A_0\rangle + \alpha_1\langle A_1\rangle + \beta_0\langle B_0\rangle + \beta_1\langle B_1\rangle + \gamma_{00}\langle A_0B_0\rangle + \gamma_{01}\langle A_0B_1\rangle + \gamma_{10}\langle A_1B_0\rangle + \gamma_{11}\langle A_1B_1\rangle. \tag{3.16}$$

In this specific case the Bell operator induced by such a Bell expression and the considered realization (3.2) is given by

$$\hat{S} = \lambda_1\hat{Z}_A + \lambda_2\hat{X}_A + \lambda_3\hat{Z}_B + \lambda_4\hat{X}_B + \lambda_5\hat{Z}_A\hat{Z}_B + \lambda_6\hat{X}_A\hat{X}_B + \lambda_7\hat{Z}_A\hat{X}_B + \lambda_8\hat{X}_A\hat{Z}_B, \tag{3.17}$$

replacing all unknown measurements $\hat{A}_x$, $\hat{B}_y$ by their quantum implementation, and where parameters λ_j are linear combinations of parameters α_x, β_y, γ_{xy}.

3.3.1 Nullspace conditions

Now, following the method depicted in Sec. 3.2.1, one finds three vectors which form a basis of the orthogonal of $|\phi_\theta\rangle$. For instance,

$$|\xi_1\rangle \coloneqq -s_\theta\,|00\rangle + c_\theta\,|11\rangle\,, \tag{3.18a}$$
$$|\xi_2\rangle \coloneqq |01\rangle \tag{3.18b}$$
$$|\xi_3\rangle \coloneqq |10\rangle\,. \tag{3.18c}$$

Then, we compute the vectors

$$\boldsymbol{T}_i \coloneqq \langle\xi_i|\,\hat{S}\,|\phi_\theta\rangle \tag{3.19}$$

in terms of the parameters λ_j that define the Bell operator $\hat{S}$. The orthogonality conditions

$$\boldsymbol{\beta}\cdot\boldsymbol{T}_i = 0 \tag{3.20}$$

[1] All Mathematica and Python notebooks used to obtain the results presented in this thesis are available in a public GitHub repository: https://github.com/andreazingarofalo/OrthogonalCHSH.

restrict the admissible Bell expressions to those for which the associated Bell operator admits a decomposition of the form

$$\hat{S} = \mathbb{1} + \sum_{i=1}^{5} p_i \hat{N}_i, \tag{3.21}$$

where the operators $\hat{N}_i$ form a basis of nullifiers for the state $|\phi_\theta\rangle$, that is, they satisfy $\hat{N}_i |\phi_\theta\rangle = 0$ for all $i = 1, \ldots, 5$. The nullifiers are given by

$$\begin{aligned}
\hat{N}_1 &\coloneqq \hat{Z}_A - \hat{Z}_B, \\
\hat{N}_2 &\coloneqq \mathbb{1} - \hat{Z}_A \hat{Z}_B, \\
\hat{N}_3 &\coloneqq \hat{X}_A \hat{Z}_B + \hat{Z}_A \hat{X}_B - \mathbb{1} - s_{2\theta} c_{2\theta} (\hat{X}_A + \hat{X}_B), \\
\hat{N}_4 &\coloneqq (1 + s_{2\theta})(\hat{X}_A - \hat{X}_B) - c_{2\theta} (\hat{X}_A \hat{Z}_B - \hat{Z}_A \hat{X}_B), \\
\hat{N}_5 &\coloneqq \mathbb{1} - s_{2\theta} \hat{X}_A \hat{X}_B - c_{2\theta} \hat{Z}_B.
\end{aligned} \tag{3.22}$$

Remark 3.2. The parametrization in Eq. (3.21) automatically satisfies the orthogonality condition (3.20), since each term $\hat{N}_i$ annihilates the state $|\phi_\theta\rangle$. Therefore, for any vector $|\xi_i\rangle \perp |\phi_\theta\rangle$, we have

$$\langle \xi_i | \hat{S} | \phi_\theta \rangle = \langle \xi_i | \left(\mathbb{1} + \sum_{i=1}^{5} p_i \hat{N}_i \right) | \phi_\theta \rangle = \langle \xi_i | \phi_\theta \rangle = 0,$$

which confirms Eq. (3.20). Furthermore, this implies

$$\hat{S} |\phi_\theta\rangle = \mathbb{1} |\phi_\theta\rangle = |\phi_\theta\rangle ,$$

so the Bell operator $\hat{S}$ has $|\phi_\theta\rangle$ as an eigenvector of eigenvalue 1.

We have thus reduced the problem of characterizing the Bell operators compatible with the extremal point $\boldsymbol{P}_{\theta,a_x,b_y}$ to a five-dimensional linear subspace.

Remark 3.3. This reduction can be justified by a simple degree-of-freedom argument: in the CHSH scenario, a general Bell operator is a linear combination of 8 linearly independent expectation values (one for each correlator), giving rise to an 8-dimensional parameter space. The orthogonality conditions impose 3 linearly independent constraints on the Bell expression $\boldsymbol{\beta}$, corresponding to the orthogonal complement of $|\phi_\theta\rangle$. Hence, the space of Bell expressions maximized by $\boldsymbol{P}$ is a 5-dimensional subspace. Since we can explicitly construct 5 linearly independent nullifiers $\hat{N}_i$ composed of the observables $\hat{X}_A, \hat{Z}_A, \hat{X}_B, \hat{Z}_B$ and their products (e.g., $\hat{X}_A \hat{X}_B$, $\hat{Z}_A \hat{Z}_B$, etc.), and since Eq. (3.21) realizes a Bell operator that (i) satisfies all orthogonality constraints and (ii) has $|\phi_\theta\rangle$ as an eigenstate of eigenvalue 1.

We rewrite our candidate Bell expression $\beta(\boldsymbol{P})$ by inverting the relation (3.2), replacing the operators $\hat{Z}_{A(B)}$, $\hat{X}_{A(B)}$ by the measurement operators $\hat{A}_x$, $\hat{B}_y$,

$$\begin{aligned}
\hat{X}_A &= \frac{c_{a_1} \hat{A}_0 - c_{a_0} \hat{A}_1}{s_{a_0 - a_1}}, \qquad & \hat{X}_B &= \frac{c_{b_1} \hat{B}_0 - c_{b_0} \hat{B}_1}{s_{b_0 - b_1}}, \\
\hat{Z}_A &= \frac{s_{a_0} \hat{A}_1 - s_{a_1} \hat{A}_0}{s_{a_0 - a_1}}, \qquad & \hat{Z}_B &= \frac{s_{b_0} \hat{B}_1 - s_{b_1} \hat{B}_0}{s_{b_0 - b_1}}.
\end{aligned} \tag{3.23}$$

and one-to-one lifting to formal polynomials. We then get a reparametrization of the Bell expression (in vector representation):

$$\boldsymbol{\beta} = \begin{pmatrix} {s_{b_0-b_1}}^{-1}((p_1 + c_{2\theta}p_5)s_{b_1} - c_{b_1}(p_4(1+s_{2\theta}) + p_3(c_{2\theta}^{-1} - t_{2\theta}))) \\ {s_{b_0-b_1}}^{-1}(-((p_1 + c_{2\theta}p_5)s_{b_0}) + c_{b_0}(p_4(1+s_{2\theta}) + p_3(c_{2\theta}^{-1} - t_{2\theta}))) \\ {s_{a_0-a_1}}^{-1}(-p_1 s_{a_1} + c_{a_1}(p_4(1+s_{2\theta}) + p_3(-c_{2\theta}^{-1} + t_{2\theta}))) \\ -{s_{a_0-a_1}}^{-1}{s_{b_0-b_1}}^{-1}(c_{2\theta}p_4 s_{a_1-b_1} + s_{a_1}(c_{b_1}p_3 + p_2 s_{b_1}) + c_{a_1}(c_{b_1}p_5 s_{2\theta} + p_3 s_{b_1})) \\ {s_{a_0-a_1}}^{-1}{s_{b_0-b_1}}^{-1}(c_{2\theta}p_4 s_{a_1-b_0} + s_{a_1}(c_{b_0}p_3 + p_2 s_{b_0}) + c_{a_1}(c_{b_0}p_5 s_{2\theta} + p_3 s_{b_0})) \\ {s_{a_0-a_1}}^{-1}(p_1 s_{a_0} - c_{a_0}(p_4(1+s_{2\theta}) + p_3(-c_{2\theta}^{-1} + t_{2\theta}))) \\ {s_{a_0-a_1}}^{-1}{s_{b_0-b_1}}^{-1}(c_{2\theta}p_4 s_{a_0-b_1} + s_{a_0}(c_{b_1}p_3 + p_2 s_{b_1}) + c_{a_0}(c_{b_1}p_5 s_{2\theta} + p_3 s_{b_1})) \\ -{s_{a_0-a_1}}^{-1}{s_{b_0-b_1}}^{-1}(c_{2\theta}p_4 s_{a_0-b_0} + s_{a_0}(c_{b_0}p_3 + p_2 s_{b_0}) + c_{a_0}(c_{b_0}p_5 s_{2\theta} + p_3 s_{b_0})) \end{pmatrix} \tag{3.24}$$

From the fact that in the general Bell expression (3.16) the coefficient of the identity operator is zero, we get the following linear constraint:

$$1 + p_2 + p_5 = 0. \tag{3.25}$$

Notice that we get the same constraint also by imposing

$$1 = \boldsymbol{\beta} \cdot \boldsymbol{P}_{\theta,a_x,b_y} = \langle \phi_\theta | \, \hat{S} \, | \phi_\theta \rangle \, .$$

3.3.2 Variational constraints

We can now use the fact that small variations around the point of interest should not increase the value of the Bell expression. This condition gives a set of linear equations in the p_is:

$$0 = \boldsymbol{\beta} \cdot \frac{\partial \boldsymbol{P}_{\theta,a_x,b_y}}{\partial a_x} = \boldsymbol{\beta} \cdot \frac{\partial \boldsymbol{P}_{\theta,a_x,b_y}}{\partial b_y} \tag{3.26}$$

which reduce to

$$s_{a_0}(-c_{a_1}(p_3 + c_{2\theta}p_4)s_{2\theta} + (p_1 c_{2\theta} - p_2)s_{a_1}) - c_{a_0}s_{2\theta}(c_{a_1}s_{2\theta}p_5 + (p_3 + c_{2\theta}p_4)s_{a_1}) = 0 \tag{3.27a}$$

$$-c_{b_0-b_1}(p_2 + p_5) + (c_{b_0+b_1}(p_2 + c_{4\theta}p_5) - 2c_{2\theta}s_{b_0}s_{b_1}p_1 + (-2p_3 s_{2\theta} + p_4 s_{4\theta}s_{b_0+b_1})) = 0 \tag{3.27b}$$

$$2s_{a_0}(c_{a_1}(p_3 + c_{2\theta}p_4)s_{2\theta} + (-c_{2\theta}p_1 + p_2)s_{a_1}) + 2c_{a_0}s_{2\theta}(c_{a_1}p_5 s_{2\theta} + (p_3 + c_{2\theta}p_4)s_{a_1}) = 0 \tag{3.27c}$$

$$t_{b_0-b_1}^{-1}(p_2 + p_5) - s_{b_0-b_1}^{-1}(c_{b_0+b_1}(p_2 + c_{4\theta}p_5) - 2c_{2\theta}p_1 s_{b_0}s_{b_1} + (-2p_3 s_{2\theta} + p_4 s_{4\theta})s_{b_0+b_1}) = 0. \tag{3.27d}$$

Upon examination, it is evident that these four equations are not linearly independent. Specifically:

- Eq. (3.27c) can be shown to be linearly dependent on Eq. (3.27a). By letting $X \coloneqq c_{a_1}(p_3 + c_{2\theta}p_4)s_{2\theta}$ and $Y \coloneqq (p_1 c_{2\theta} - p_2)s_{a_1}$, and similarly $Z \coloneqq c_{a_1}s_{2\theta}p_5$ and $W \coloneqq (p_3 + c_{2\theta}p_4)s_{a_1}$, Eq. (3.27a) can be written as $s_{a_0}(-X + Y) - c_{a_0}s_{2\theta}(Z + W) = 0$. In contrast, Eq. (3.27c) is $2s_{a_0}(X - Y) + 2c_{a_0}s_{2\theta}(Z + W) = 0$. Hence, (3.27c) is equivalent to $-2 \times$ (3.27a).

- Eq. (3.27d) is directly related to Eq. (3.27b). Let $K_1 \coloneqq (p_2 + p_5)$ and $K_2 \coloneqq (c_{b_0+b_1}(p_2 + c_{4\theta}p_5) - 2c_{2\theta}s_{b_0}s_{b_1}p_1 + (-2p_3 s_{2\theta} + p_4 s_{4\theta} s_{b_0+b_1}))$. Eq. (3.27b) becomes $-c_{b_0-b_1}K_1 + K_2 = 0$, while Eq. (3.27d) can be written as $\frac{c_{b_0-b_1}}{s_{b_0-b_1}}K_1 - \frac{1}{s_{b_0-b_1}}K_2 = 0$. Assuming $s_{b_0-b_1} \neq 0$, Eq. (3.27d) provides no new information beyond that contained in (3.27b).

Consequently, the system of four equations effectively reduces to two independent linear equations in terms of the five parameters $(p_1, p_2, p_3, p_4, p_5)$. Taking into account Eq. (3.25), this implies that the solution space is a subspace of dimension $5 - 3 = 2$. Thus, the search for Bell expressions maximized by bmP_{θ,a_x,b_y} can be reduced to an affine plane.

3.4 A Tsirelson-inspired family of extremal points

We now begin the original part of our analysis, where we apply the general framework developed in previous sections to a specific family of quantum realizations.

To reduce the problem and draw a direct connection to a well-known quantum realization, we choose the measurement angles corresponding to those that achieve the Tsirelson bound for the CHSH inequality:

$$a_0 = \frac{\pi}{4}, \qquad a_1 = -\frac{\pi}{4}, \qquad b_0 = 0, \qquad b_1 = \frac{\pi}{2}. \tag{3.28}$$

Since our interest lies in the characterization of extremal points of the quantum set $\mathcal{Q}$, we make systematic use of the full alternation condition introduced in Theorem 3.2 to determine which values of the entanglement parameter θ are admissible. Specifically, we consider quantum realizations of the form (3.2), brought into the canonical ordering specified by (3.4) via appropriate relabelings, as previously done in Example 3.1.

The full alternation condition imposes a set of inequalities on the steered measurement angles,

$$\tilde{a}_x^j \coloneqq 2\arctan\left(\tan\left(\frac{a_x}{2}\right)\tan^j(\theta)\right),$$

requiring that for all $s, t \in \{\pm 1\}$, the following ordered relation holds:

$$0 \leq [\tilde{a}_0^s]_\pi \leq b_0 \leq [\tilde{a}_1^t]_\pi \leq b_1 < \pi.$$

A careful analytical inspection shows that, once the realization is cast into canonical form, this condition is satisfied if and only if $\theta \in (\pi/8, \pi/4]$. For this reason, in the remainder of the thesis we restrict our analysis to this interval of θ, which fully captures all nonlocal extremal points arising from real two-qubit realizations given the measurement settings (3.28).

Substituting the values from (3.28) into (3.27) yields the following system of linear equations:

$$\begin{cases} c_{2\theta}p_1 - p_2 + s_{4\theta}^2 p_5 = 0 \\ 2s_{2\theta}p_3 - s_{4\theta}p_4 = 0 \end{cases} \tag{3.29}$$

Considering these relations, and also taking into account the global normalization or constraint (Eq. (3.25)), we can re-parameterize the entire problem in terms of p_1 and p_4.

Our final Bell expression, expressed as a formal polynomial in terms of the measurement outcomes, can thus be written as

$$\begin{aligned}\beta_{\theta,p_1,p_4} = \frac{1}{6-2c_{4\theta}}\Big[&p_4\Big(-4(3-c_{4\theta})B_1 + 2\sqrt{2}(3-c_{4\theta})s_{2\theta}A_0 - 2\sqrt{2}(3-c_{4\theta})s_{2\theta}A_1 \\ &+ \sqrt{2}(5c_{2\theta}-c_{6\theta})A_0B_1 + \sqrt{2}(5c_{2\theta}-c_{6\theta})A_1B_1\Big) \\ &+ p_1\Big(\sqrt{2}(3-c_{4\theta})A_0 + \sqrt{2}(3-c_{4\theta})A_1 - 4(1-c_{4\theta})B_0 \\ &- 2\sqrt{2}c_{2\theta}A_0B_0 - 2\sqrt{2}c_{2\theta}A_1B_0 + \sqrt{2}s_{4\theta}A_0B_1 - \sqrt{2}s_{4\theta}A_1B_1\Big) \\ &4c_{2\theta}B_0 + \sqrt{2}(1-c_{4\theta})A_0B_0 + \sqrt{2}(1-c_{4\theta})A_1B_0 + 2\sqrt{2}s_{2\theta}A_0B_1 - 2\sqrt{2}s_{2\theta}A_1B_1\Big].\end{aligned} \tag{3.30}$$

Note that, for $\theta = \pi/4$, we recover

$$\beta_{p_1,p_4} = p_1\left(\frac{A_0+A_1}{\sqrt{2}} - B_0\right) + 2p_4\left(\frac{A_0-A_1}{\sqrt{2}} - B_1\right) + \frac{1}{2\sqrt{2}}\beta_{\text{CHSH}}. \tag{3.31}$$

that is exactly (2.57) with the identification $p_1 = r_0$ and $p_4 = r_1/2$. For this reason, in the remainder of the thesis, we will adopt the reparametrization

$$\beta_{\theta,r_0,r_1} \coloneqq \beta_{\theta,p_1,2p_4}. \tag{3.32}$$

3.5 Local bounds

We now turn to the study of the dual of the local set, denoted $\mathcal{L}^*$ in the parametric plane of Bell expression defined by Eq. (3.30). This generalizes the analysis we did in Sec. 2.4 for the plane defined by Eq. (2.57), which corresponds to the special case $\theta = \pi/4$.

In this case, the relevant Bell functional $\boldsymbol{\beta}_{\theta,r_0,r_1}$ depends on a continuous parameter θ and two real coefficients r_0, r_1, and we study the subset of $\mathcal{L}^*$ defined by the inequality

$$\boldsymbol{\beta}_{\theta,r_0,r_1} \cdot \boldsymbol{D}_{ijkl} \leq 1, \qquad \forall i,j,k,l \in \{-1,1\}, \tag{3.33}$$

which yields a family of 16 linear inequalities. For fixed θ, these inequalities describe a 2D region in the (r_0, r_1) plane—specifically, a polygonal region given by the intersection of 16 half-planes.

3.5.1 Vertices of the projected dual local polytope

To identify the shape and boundary of this polygon, we proceeded as follows. Each of the 16 inequalities defines a line in the (r_0, r_1) plane, and the feasible region is the intersection of the corresponding half-planes. The vertices of the polygon are given by the intersection points of pairs of such lines which lie on the boundary of the feasible region.

There are

$$\binom{16}{2} = 120$$

pairs of inequalities, and thus at most 120 candidate intersection points. Using a symbolic computation software, we analytically computed the coordinates of all such intersections. In fact, we identified 114 distinct intersections. Each point was then tested against all 16 inequalities to determine whether it lies in the feasible region.

By repeating this procedure for values of θ in the interval $\theta \in (\pi/8, \pi/4]$, we analyzed how the shape of the projected dual local polytope changes as a function of θ. Our analysis revealed two qualitatively distinct regimes:

- For $\theta \leq \bar{\theta}$, the polygon has six vertices;
- For $\theta > \bar{\theta}$, the polygon has eight vertices.

The transition occurs at a critical value $\bar{\theta} \approx 0.644539533$.

At the endpoint $\theta = \pi/4$, the dual polytope projection recovers the regular octagon described in Sec. 2.4.

3.5.2 Visualizations and vertex formulas

In what follows, we provide the explicit analytic expressions for the vertices of the polygon in both regimes, together with plots of the polytope projection for representative values of θ. These results provide a complete characterization of the dual local bounds in the (r_0, r_1) plane as a function of θ.

For $\theta \leq \bar{\theta}$, the vertices as functions of θ are given by

$$v_1(\theta) = \left(\frac{(\sqrt{2}s_{2\theta} - 1)c_\theta s_{2\theta}^{-1}(2\sqrt{2}s_\theta - 3c_\theta + c_{3\theta})}{\sqrt{2}c_{2\theta} + 8s_\theta^2 c_\theta(s_\theta - \sqrt{2}c_\theta)}, 0 \right), \tag{3.34}$$

$$v_2(\theta) = \left(-\frac{s_\theta^{-1}((\sqrt{2}-2)s_{6\theta} + 2(\sqrt{2}+2)s_{4\theta} + (6-11\sqrt{2})s_{2\theta} - 4c_{4\theta} - 5\sqrt{2}c_{2\theta} + \sqrt{2}c_{6\theta} - 4(\sqrt{2}-3))}{4(-2c_{3\theta} - 3(\sqrt{2}-2)c_\theta + \sqrt{2}(-3s_\theta + s_{5\theta} + c_{5\theta}))}, \right.$$
$$\left. \frac{s_{2\theta}((\sqrt{2}-2)s_\theta + (3\sqrt{2}-4)c_\theta)}{2c_{3\theta} + 3(\sqrt{2}-2)c_\theta - \sqrt{2}(-3s_\theta + s_{5\theta} + c_{5\theta})} \right), \tag{3.35}$$

$$v_3(\theta) = \left(-\frac{s_\theta^{-1}((\sqrt{2}-2)s_{6\theta} + 2(\sqrt{2}+2)s_{4\theta} + (6-11\sqrt{2})s_{2\theta} - 4c_{4\theta} - 5\sqrt{2}c_{2\theta} + \sqrt{2}c_{6\theta} - 4(\sqrt{2}-3))}{4(-2c_{3\theta} - 3(\sqrt{2}-2)c_\theta + \sqrt{2}(-3s_\theta + s_{5\theta} + c_{5\theta}))}, \right.$$
$$\left. \frac{2s_\theta c_\theta((\sqrt{2}-2)s_\theta + (3\sqrt{2}-4)c_\theta)}{-2c_{3\theta} - 3(\sqrt{2}-2)c_\theta + \sqrt{2}(-3s_\theta + s_{5\theta} + c_{5\theta})} \right), \tag{3.36}$$

$$v_4(\theta) = \left(\frac{s_\theta(\sqrt{2} - s_{2\theta}^{-1})(-3s_\theta - s_{3\theta} + 2\sqrt{2}c_\theta)}{\sqrt{2}c_{2\theta} + 8s_\theta c_\theta^2(\sqrt{2}s_\theta - c_\theta)}, 0 \right), \tag{3.37}$$

$$v_5(\theta) = \left(\frac{12s_\theta - \sqrt{2}s_{5\theta} + (4-5\sqrt{2})s_{3\theta} + (\sqrt{2}-2)c_{5\theta} - 8(\sqrt{2}-1)c_\theta + (3\sqrt{2}+2)c_{3\theta}}{4(s_\theta + c_\theta)((\sqrt{2}-3)s_{2\theta} + \sqrt{2}s_{4\theta} - c_{2\theta} + \sqrt{2} - 1)}, \right.$$
$$\left. \frac{s_\theta^2(\sqrt{2}s_\theta + (\sqrt{2}-2)c_\theta)}{(s_\theta + c_\theta)((\sqrt{2}-3)s_{2\theta} + \sqrt{2}s_{4\theta} - c_{2\theta} + \sqrt{2} - 1)} \right), \tag{3.38}$$

$$v_6(\theta) = \left(\frac{12s_\theta - \sqrt{2}s_{5\theta} + (4-5\sqrt{2})s_{3\theta} + (\sqrt{2}-2)c_{5\theta} - 8(\sqrt{2}-1)c_\theta + (3\sqrt{2}+2)c_{3\theta}}{4(s_\theta + c_\theta)((\sqrt{2}-3)s_{2\theta} + \sqrt{2}s_{4\theta} - c_{2\theta} + \sqrt{2} - 1)}, \right.$$
$$\left. - \frac{s_\theta^2(\sqrt{2}s_\theta + (\sqrt{2}-2)c_\theta)}{(s_\theta + c_\theta)((\sqrt{2}-3)s_{2\theta} + \sqrt{2}s_{4\theta} - c_{2\theta} + \sqrt{2} - 1)} \right). \tag{3.39}$$

For $\theta > \bar{\theta}$, the vertices as functions of θ are given by

$$w_1(\theta) = \left(\frac{c_{2\theta}(4 - 3\sqrt{2} + (-2+\sqrt{2})c_{4\theta})}{4 - 3\sqrt{2} + \sqrt{2}c_{4\theta}}, \frac{(4-3\sqrt{2})s_{2\theta}^2}{4 - 3\sqrt{2} + \sqrt{2}c_{4\theta}} \right), \tag{3.40}$$

$$w_2(\theta) = \left(\frac{c_\theta(4 + \sqrt{2} + 2(-2+\sqrt{2})c_{2\theta} - \sqrt{2}c_{4\theta} - 2(-1+\sqrt{2})c_\theta^{-2} - 4\sqrt{2}s_{2\theta} + (-2+\sqrt{2})s_{4\theta} - 2(-2+\sqrt{2})t_\theta)}{-3\sqrt{2}c_\theta + \sqrt{2}c_{5\theta} - 3(-2+\sqrt{2})s_\theta + 2s_{3\theta} + \sqrt{2}s_{5\theta}}, \right.$$
$$\left. \frac{2c_\theta s_\theta((-2+\sqrt{2})c_\theta + (-4+3\sqrt{2})s_\theta)}{-3\sqrt{2}c_\theta + \sqrt{2}c_{5\theta} - 3(-2+\sqrt{2})s_\theta + 2s_{3\theta} + \sqrt{2}s_{5\theta}} \right), \tag{3.41}$$

$$w_3(\theta) = \left(\frac{c_\theta s_{2\theta}^{-1}(-3c_\theta + c_{3\theta} + 2\sqrt{2}s_\theta)(-1+\sqrt{2}s_{2\theta})}{\sqrt{2}c_{2\theta} + 8c_\theta s_\theta^2(-\sqrt{2}c_\theta + s_\theta)}, 0 \right), \tag{3.42}$$

$$w_4(\theta) = \left(- \frac{s_\theta^{-1}(-4(-3+\sqrt{2}) - 5\sqrt{2}c_{2\theta} - 4c_{4\theta} + \sqrt{2}c_{6\theta} + (6-11\sqrt{2})s_{2\theta} + 2(2+\sqrt{2})s_{4\theta} + (-2+\sqrt{2})s_{6\theta})}{4(-3(-2+\sqrt{2})c_\theta - 2c_{3\theta} + \sqrt{2}(c_{5\theta} - 3s_\theta + s_{5\theta}))}, \right.$$
$$\left. \frac{((-4+3\sqrt{2})c_\theta + (-2+\sqrt{2})s_\theta)s_{2\theta}}{3(-2+\sqrt{2})c_\theta + 2c_{3\theta} - \sqrt{2}(c_{5\theta} - 3s_\theta + s_{5\theta})} \right), \tag{3.43}$$

$$w_5(\theta) = \left(\frac{c_\theta(4 + \sqrt{2} + 2(-2+\sqrt{2})c_{2\theta} - \sqrt{2}c_{4\theta} - 2(-1+\sqrt{2})c_\theta^{-2} - 4\sqrt{2}s_{2\theta} + (-2+\sqrt{2})s_{4\theta} - 2(-2+\sqrt{2})t_\theta)}{-3\sqrt{2}c_\theta + \sqrt{2}c_{5\theta} - 3(-2+\sqrt{2})s_\theta + 2s_{3\theta} + \sqrt{2}s_{5\theta}}, \right.$$
$$\left. - \frac{2c_\theta s_\theta((-2+\sqrt{2})c_\theta + (-4+3\sqrt{2})s_\theta)}{-3\sqrt{2}c_\theta + \sqrt{2}c_{5\theta} - 3(-2+\sqrt{2})s_\theta + 2s_{3\theta} + \sqrt{2}s_{5\theta}} \right), \tag{3.44}$$

$$w_6(\theta) = \left(\frac{c_{2\theta}(4 - 3\sqrt{2} + (-2+\sqrt{2})c_{4\theta})}{4 - 3\sqrt{2} + \sqrt{2}c_{4\theta}}, \frac{(-4+3\sqrt{2})s_{2\theta}^2}{4 - 3\sqrt{2} + \sqrt{2}c_{4\theta}} \right), \tag{3.45}$$

$$w_7(\theta) = \left(- \frac{s_\theta^{-1}(-4(-3+\sqrt{2}) - 5\sqrt{2}c_{2\theta} - 4c_{4\theta} + \sqrt{2}c_{6\theta} + (6-11\sqrt{2})s_{2\theta} + 2(2+\sqrt{2})s_{4\theta} + (-2+\sqrt{2})s_{6\theta})}{4(-3(-2+\sqrt{2})c_\theta - 2c_{3\theta} + \sqrt{2}(c_{5\theta} - 3s_\theta + s_{5\theta}))}, \right.$$
$$\left. \frac{2c_\theta s_\theta((-4+3\sqrt{2})c_\theta + (-2+\sqrt{2})s_\theta)}{-3(-2+\sqrt{2})c_\theta - 2c_{3\theta} + \sqrt{2}(c_{5\theta} - 3s_\theta + s_{5\theta})} \right), \tag{3.46}$$

$$w_8(\theta) = \left(\frac{(\sqrt{2} - s_{2\theta}^{-1})s_\theta(2\sqrt{2}c_\theta - 3s_\theta - s_{3\theta})}{\sqrt{2}c_{2\theta} + 8c_\theta^2 s_\theta(-c_\theta + \sqrt{2}s_\theta)}, 0 \right). \tag{3.47}$$

The boundaries of the dual local set in the (r_0, r_1) plane for $\theta \in (\pi/8, \pi/4]$ are represented in Fig. 3.1.

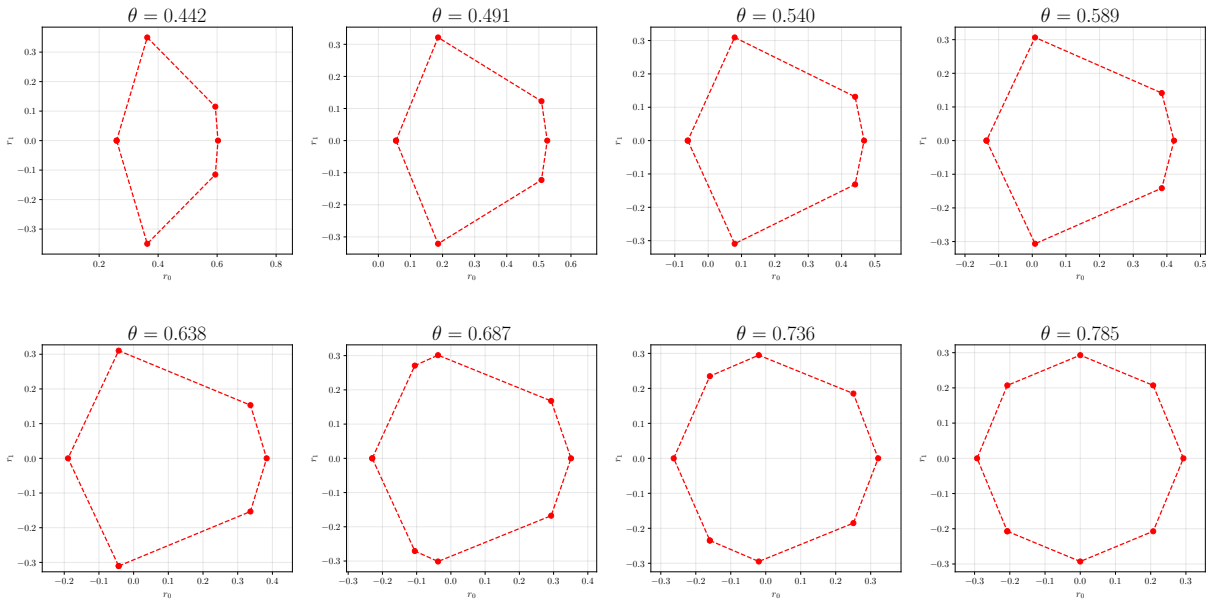

Figure 3.1: Dual local polytope $\mathcal{L}^*$ in the considered (r_0, r_1) plane, for various values of the parameter $\theta \in (\frac{\pi}{8}, \frac{\pi}{4}]$. The analysis reveals two distinct geometric regimes: for $\theta < \bar{\theta} \approx 0.645$, the dual polytope has six vertices, while for $\theta > \bar{\theta}$, the polygon has eight vertices. At the endpoint $\theta = \frac{\pi}{4}$, the projected polytope recovers the regular octagon described in Sec. 2.4. This behavior highlights a structural transition in the local dual geometry as a function of θ.

3.6 Quantum bounds

Finally, we turn to the task of characterizing the boundary of the dual quantum set $\mathcal{Q}$. Our goal is to understand which Bell functionals can be maximally violated by quantum correlations and how these violations depend on $\theta \in (\pi/8, \pi/4)$.

3.6.1 NPA hierarchy estimation

To characterize the boundary of the dual quantum set $\mathcal{Q}^*$ in the (r_0, r_1)-plane for various values of the parameter θ, we resort to numerical techniques. Recall that the dual set $\mathcal{Q}^*$ consists of the Bell expressions whose quantum value does not exceed a fixed bound—here normalized to one. Its boundary corresponds to the family of Bell functionals that are maximally violated by quantum correlations.

We employ the NPA hierarchy as described in eqs. (1.84)-(1.85), as implemented in the Python package `ncpol2sdpa` [123]. For a given point (r_0, r_1), we define a Bell functional $\beta_{\theta,r_0,r_1}(\boldsymbol{P})$, and use the dual NPA hierarchy to compute its maximal quantum value. The optimization was performed over functionals on moment matrices associated with a fixed level of the NPA hierarchy, subject to the commutation relations and normalization constraints corresponding to unitary measurements $\hat{A}_x$, $\hat{B}_y$ in a bipartite Bell experiment, as in (1.29).

To determine whether a point (r_0, r_1) lies within the set $\mathcal{Q}^*$, we verified whether the quantum value of $\beta_{\theta,r_0,r_1}(\boldsymbol{P})$ remains less than or equal to one, as computed via the dual SDP. We systematically explored the (r_0, r_1)-plane using a polar parametrization centered at the centroid

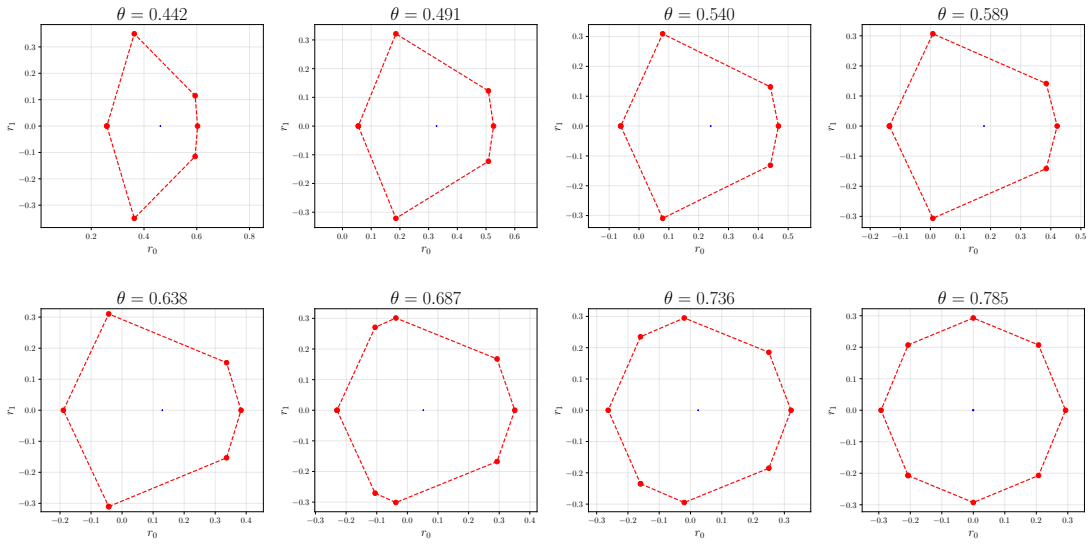

Figure 3.2: Numerical estimation of the dual quantum boundaries (shown in blue) at the level $T_A \coloneqq \mathcal{T}_{1+A+B}$, for values of $\theta \in (\pi/8, \pi/4]$ versus the dual local boundaries (in red). At this level of the NPA hierarchy, the dual quantum set $\mathcal{Q}^*$ reduces to a single point when $\theta = \pi/4$, reflecting the fact that the relaxation is too coarse to capture the structure of the boundary. Cf. Fig. 2.2.

of the local polytope. For each angle $\varphi \in [0, 2\pi]$, we minimized the radial distance r such that the quantum value of the corresponding Bell functional remains below the threshold. This procedure was repeated across a discretized set of directions, to construct a closed approximation of the boundary.

At each step, the SDP was solved numerically using `ncpol2sdpa`, interfacing with the MOSEK solver [124] for efficiency and accuracy.

The resulting estimation thus yields a numerical approximation of the boundary of $\mathcal{Q}^*$, with increasing levels of the NPA hierarchy providing progressively tighter inner approximations.

Coarse approximations. At the lowest tested level $T_A \coloneqq \mathcal{T}_{1+A+B}$ (see Fig. 3.2), the outer approximation collapses to a single point for each θ, indicating that the relaxation is too coarse to capture any meaningful geometric structure of the quantum set. This is expected, as no operator products are yet included in the relaxation. Consequently, the dual quantum set $\mathcal{Q}^*$ degenerates into a trivial set, incapable of distinguishing between different Bell functionals.

Emergence of structure. As we enrich the relaxation with the products AB, the dual quantum set begins to exhibit nontrivial geometry. At level $T_{AB} \coloneqq \mathcal{T}_{1+A+B+AB}$ (Fig. 3.3), the boundary of $\mathcal{Q}^*$ acquires a smoothly curved, approximately elliptical shape, which becomes increasingly symmetric as $\theta \to \pi/4$. At this maximally entangled point, the boundary is circular and centered at the origin, matching the known symmetry of the corresponding Tsirelson region.

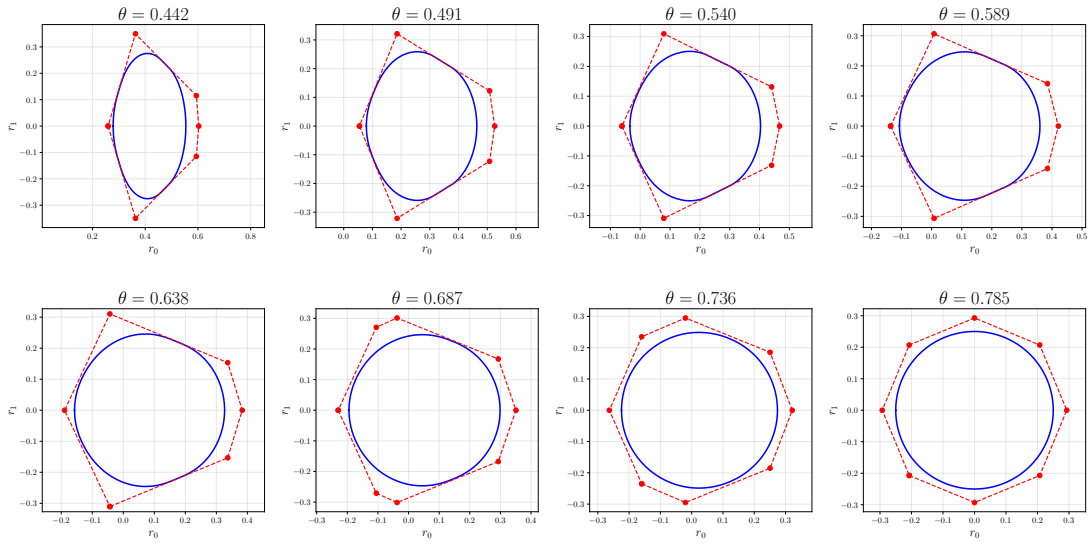

Figure 3.3: Numerical estimation of the dual quantum boundaries (shown in blue) at the level $T_{AB} \coloneqq \mathcal{T}_{1+A+B+AB}$, for values of $\theta \in (\pi/8, \pi/4]$ versus the dual local boundaries (in red). At this level of the NPA hierarchy, the dual quantum set $\mathcal{Q}^*$ acquires a nontrivial, smoothly curved shape. The boundaries exhibit an approximately elliptical profile, which becomes increasingly symmetric as θ approaches $\pi/4$. In the special case $\theta = \pi/4$, we recover the circular boundary centered at the origin in the (r_0, r_1)-plane. Cf. Fig. 2.2.

Near saturation and vertex rounding. With the inclusion of higher-order products such as AA', BB', and AB, we obtain the level $T'_{AB} \coloneqq \mathcal{T}_{1+A+B+AA'+BB'+AB}$ (Fig. 3.4). The boundary of $\mathcal{Q}^*$ at this level closely resembles that of the local dual set $\mathcal{L}^*$, yet the approximation fails to reach its extremal vertices. A more refined numerical analysis confirms this: Bell functionals associated with local vertices yield quantum values that deviate from the local bound by several digits (e.g., 0.9998, 1.001). These deviations are significant given the solver's precision and indicate that the relaxation does not fully capture the sharp corners of $\mathcal{L}^*$.

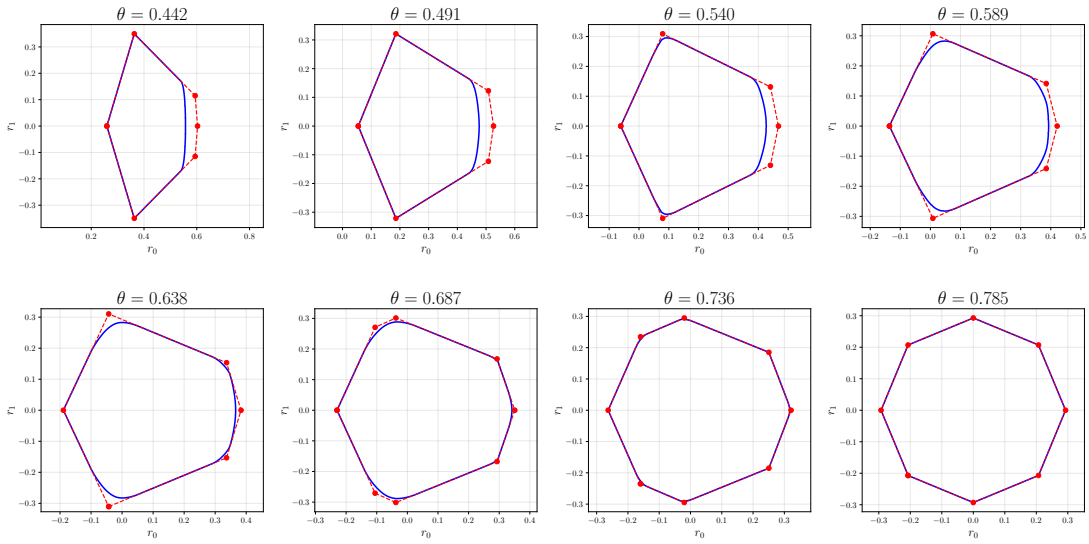

Figure 3.4: Numerical estimation of the dual quantum boundaries (shown in blue) at the level $T'_{AB} \coloneqq \mathcal{T}_{1+A+B+AA'+BB'+AB}$, for values of $\theta \in (\pi/8, \pi/4]$ versus the dual local boundaries (in red). At this level of the NPA hierarchy, the boundary resembles more closely that of the dual local polytope, exhibiting similar overall geometry. However, the quantum boundaries remain strictly rounded near the vertices, and most extremal points of the local dual set are not actually reached. The rounding effect observed at the corners is characteristic of the relaxation's inability to capture sharp extremal points at this level. Although some of the vertices may appear to lie on the boundary at visual resolution, a more refined numerical analysis reveals otherwise: the maximal quantum value of the corresponding Bell expressions differs from the local bound by several digits, typically yielding values such as 0.9998 or 1.001 (see e.g. Table 3.1). Given the solver's internal precision (up to at least six decimal places), this confirms that the relaxation does not fully saturate the local dual vertices.

Tighter relaxations. At level $T_{ABB} \coloneqq \mathcal{T}_{1+A+B+AB+ABB'}$ (Fig. 3.5), the approximation becomes markedly tighter: for most values of θ, the boundary of $\mathcal{Q}^*$ visually overlaps with that of $\mathcal{L}^*$. However, as with previous levels, careful inspection shows that the relaxation still fails to exactly saturate the local vertices. This reflects the inherent softness of semidefinite relaxations in capturing extremal combinatorial points of polytopes.

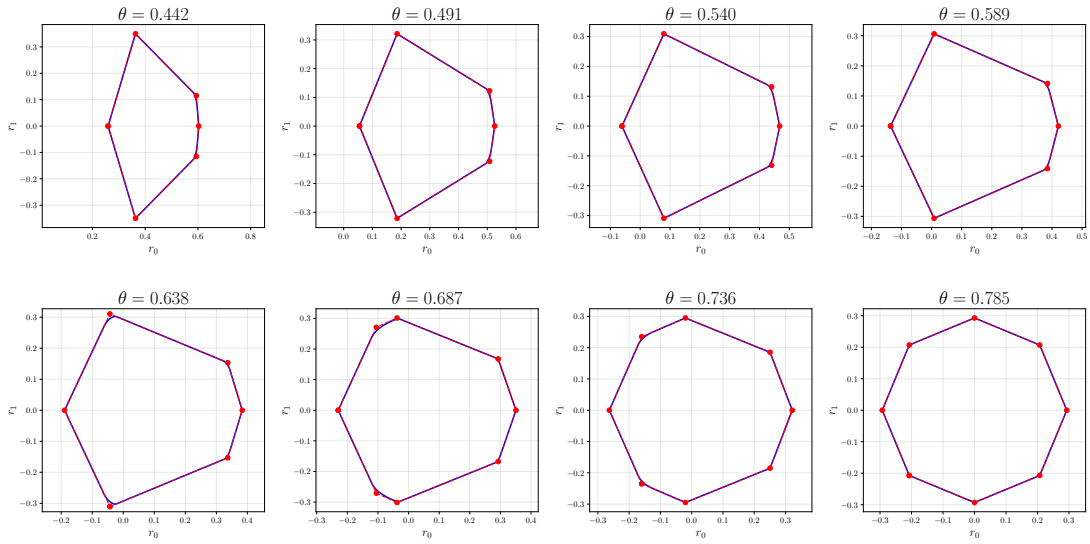

Figure 3.5: Numerical estimation of the dual quantum boundaries (shown in blue) at the level $T_{ABB} \coloneqq \mathcal{T}_{1+A+B+AB+ABB'}$, for values of $\theta \in (\pi/8, \pi/4]$ versus the dual local boundaries (in red). At this level of the NPA hierarchy, the approximation becomes remarkably tight: for almost all values of θ, the quantum boundary seems to coincide with the local boundary (recall that $\mathcal{Q}^* \subseteq \mathcal{L}^*$). However, although most of the vertices may appear to lie on the boundary at visual resolution, a more refined numerical analysis reveals otherwise: the maximal quantum value of the corresponding Bell expressions differs from the local bound by several digits, typically yielding values such as 0.9998 or 1.001 (see e.g. Table 3.1). Given the solver's internal precision (up to at least six decimal places), this confirms that the relaxation does not fully saturate the local dual vertices. The relaxation T_{ABB} is not sufficient to capture the extremal quantum violations of the Bell functionals under consideration.

Level	**Primal**	**Dual**	**Status**
$\mathcal{T}_{1+A+B}$	1.0857864377	1.0857864365	optimal
$\mathcal{T}_{1+A+B+AA'}$	1.0857864379	1.0857864367	optimal
$\mathcal{T}_{1+A+B+BB'}$	1.0857864383	1.0857864346	optimal
$\mathcal{T}_{1+A+B+AB}$	1.0091153569	1.0091153431	optimal
$\mathcal{T}_{1+A+B+AA'B}$	1.0828606998	1.0828606844	optimal
$\mathcal{T}_{1+A+B+ABB'}$	1.0857864420	1.0857864355	optimal
$\mathcal{T}_{1+A+B+AB+ABB'}$	1.0000000084	0.9999999940	optimal
$\mathcal{T}_{1+A+B+AA'+AB+ABB'}$	1.0000000012	0.9999999992	optimal
$\mathcal{T}_{1+A+B+BB'+AB+ABB'}$	1.0000000014	0.9999999990	optimal
$\mathcal{T}_{1+A+B+A'A+B'B+AB+ABB'}$	1.0000000006	0.9999999993	optimal
$\mathcal{T}_{1+A+B+AB+AA'B+ABB'}$	1.0000000015	0.9999999990	optimal

Table 3.1: Maximal quantum value of the Bell score estimated numerically for vertex $v_1(\pi/4)$ ($(r_0, r_1) = (0, 1 - 1/\sqrt{2})$) under various levels of the NPA hierarchy.

Saturation of local boundaries. Two higher levels, $T'_{ABB} \coloneqq \mathcal{T}_{1+A+B+BB'+AB+ABB'}$ (Fig. 3.6) and $T''_{ABB}\mathcal{T}_{1+A+B+AB+AA'B+ABB'}$ (Fig. 3.7), yield dual quantum boundaries that match the local dual boundaries up to numerical precision for all θ. This confirms that these levels of the

NPA hierarchy are sufficient to capture the extremal quantum violations of the Bell functionals under study. The boundary of $\mathcal{Q}^*$ is indistinguishable from that of $\mathcal{L}^*$ at the resolution of the SDP solver, indicating convergence of the hierarchy at these levels for the considered scenario.

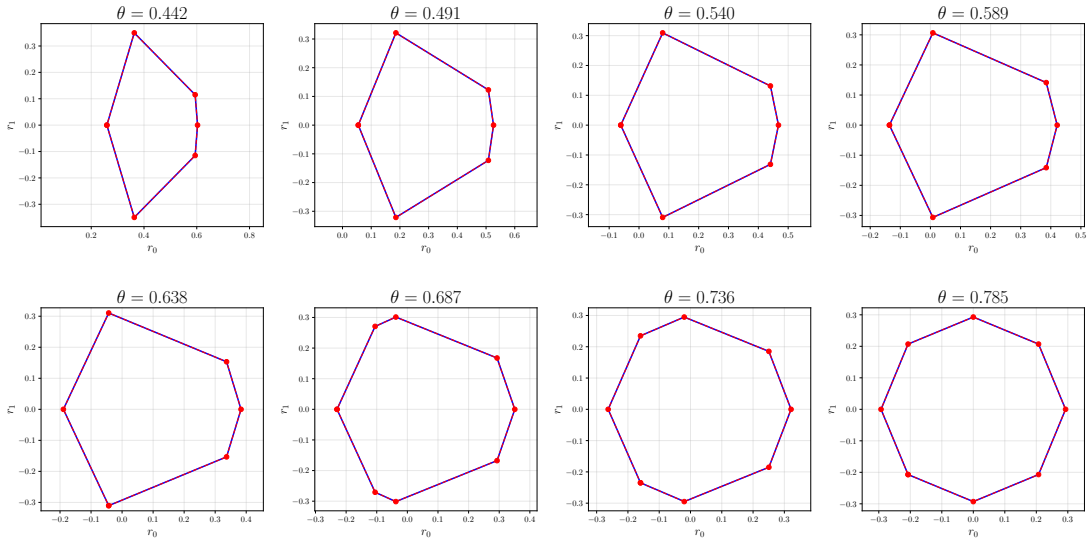

Figure 3.6: Numerical estimation of the dual quantum boundaries (shown in blue) at the level $T'_{ABB} \coloneqq \mathcal{T}_{1+A+B+BB'+AB+ABB'}$, for values of $\theta \in (\pi/8, \pi/4]$ versus the dual local boundaries (in red). At this level of the NPA hierarchy, the approximation becomes remarkably tight: for all values of θ, the quantum boundary coincides with the local boundary of $\mathcal{Q}^*$ up to numerical precision (recall that $\mathcal{Q}^* \subseteq \mathcal{L}^*$). This indicates that the relaxation T'_{ABB} is sufficient to capture the extremal quantum violations of the Bell functionals under consideration. The resulting boundary perfectly matches the expected shape derived from the exact local dual geometry.

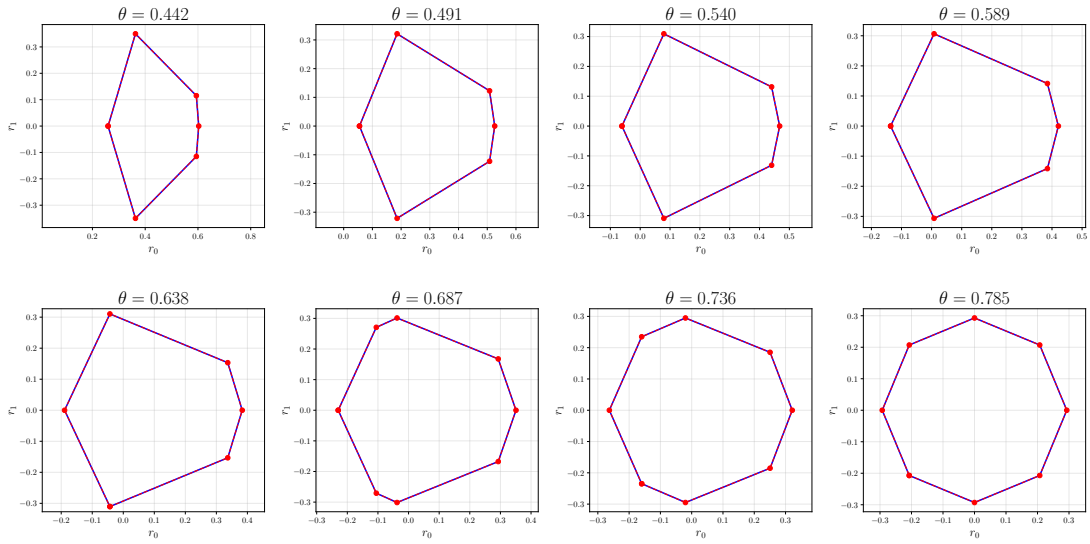

Figure 3.7: Numerical estimation of the dual quantum boundaries (shown in blue) at the level $T''_{ABB} \coloneqq \mathcal{T}_{1+A+B+AB+AA'B+ABB'}$, for values of $\theta \in (\pi/8, \pi/4]$ versus the dual local boundaries (in red). At this level of the NPA hierarchy, the approximation becomes remarkably tight: for all values of θ, the quantum boundary coincides with the local boundary of $\mathcal{Q}^*$ up to numerical precision (recall that $\mathcal{Q}^* \subseteq \mathcal{L}^*$). This indicates that the relaxation T''_{ABB} is sufficient to capture the extremal quantum violations of the Bell functionals under consideration. The resulting boundary perfectly matches the expected shape derived from the exact local dual geometry. Cf. Fig. 2.2.

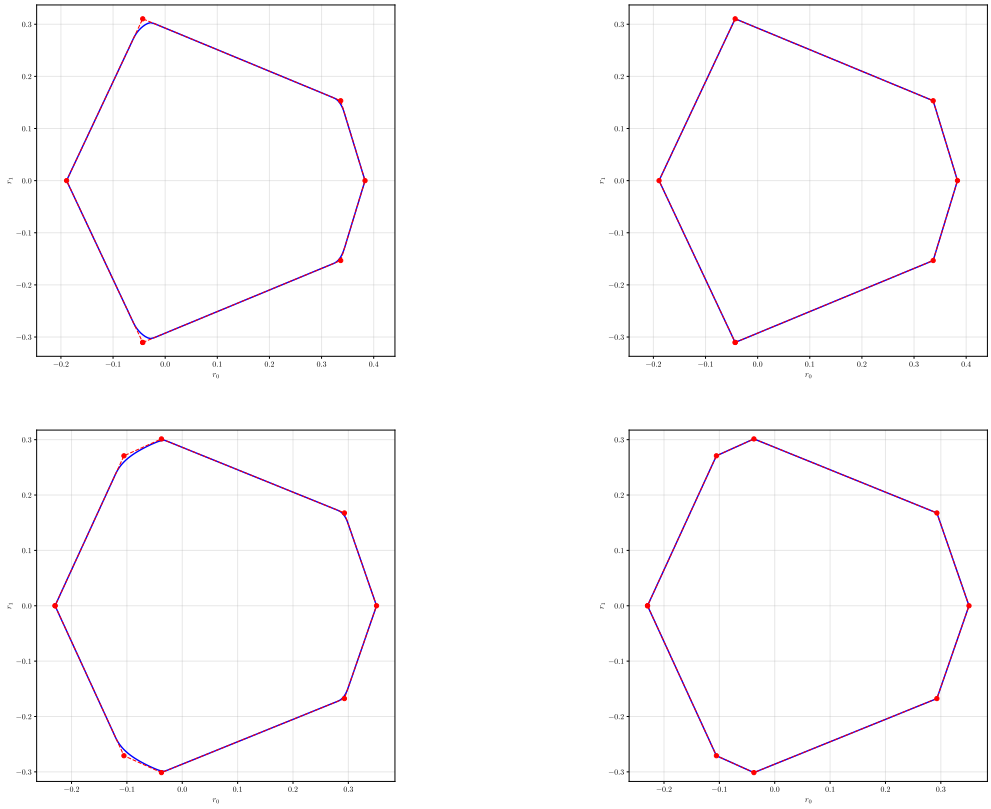

Figure 3.8: Numerical estimation of the dual quantum boundary (in blue) under two successive levels of the NPA hierarchy versus the dual local boundary (in red), for two different values of θ. The top row shows the results for $\theta = \pi/8 + 5\pi/64$, while the bottom row corresponds to $\theta = \pi/8 + 3\pi/32$. In both cases, the left panels use the relaxation $\mathcal{T}_{1+A+B+AB+ABB'}$, which provides a close but incomplete approximation, failing to saturate certain extremal vertices of the local dual set. The right panels use the enhanced relaxation $\mathcal{T}_{1+A+B+BB'+AB+ABB'}$, which successfully captures the full boundary of $\mathcal{Q}^*$ up to numerical precision.

3.6.2 Searching for an SOS decomposition

From the numerical results obtained in the previous sections, we observed that all the vertices of the dual local polytope (in the (r_0, r_1)-projection) are reached by the boundary of the dual quantum set $\mathcal{Q}^*$ up to numerical precision. In particular, for every vertex, the corresponding Bell expression β_{r_0,r_1} always achieves quantum value 1 at level $\mathcal{T}_{1+A+B+BB'+AB+ABB'}$, and at the same time at level $\mathcal{T}_{1+A+B+AB+AA'B+ABB'}$.

In light of these observations, a natural goal is to provide an *analytic proof* of this saturation. As discussed in Chapter 2, this can be done by constructing a sum-of-squares (SOS) decomposition of the form:

$$1 - \beta_{\theta,r_0,r_1} = \boldsymbol{M}^\dagger W_{\mathcal{B}} \boldsymbol{M}, \tag{3.48}$$

where each formal monomial M_i belongs to a suitable monomial basis $\mathcal{B}$. The existence of such a decomposition implies that the quantum value of β_{θ,r_0,r_1} is upper bounded by 1, and this bound is attained by the state $|\phi_\theta\rangle$.

Moreover, when the Bell expression saturates the Tsirelson bound at a given level of the NPA hierarchy $\mathcal{T}_{\mathcal{B}}$, defined by a monomial basis $\mathcal{B}$, it follows from the dual formulation of the SDP that such an SOS decomposition must exist using operators in $\mathcal{B}$ [70, 117]. For instance, one such basis at level $\mathcal{T}_{1+A+B+AB+ABB'}$ is given by

$$\mathcal{B} \coloneqq \{1, A_0, A_1, B_0, B_1, A_0B_0, A_0B_1, A_1B_0, A_1B_1, A_0B_0B_1, A_0B_1B_0, A_1B_0B_1, A_1B_1B_0\}. \quad (3.49)$$

As we have seen, a particularly convenient basis for constructing SOS decompositions is the one generated by the nullifiers of the target quantum state $|\phi_\theta\rangle$, since these operators vanish on the state and naturally span the kernel of the moment matrix. Expressing the shifted Bell expression $1 - \beta_{\theta,r_0,r_1}$ as a quadratic form in the nullifiers provides a structured and physically meaningful decomposition.

Unfortunately, despite extensive effort, we were not able to find a complete analytic SOS certificate for the set of vertices considered. The difficulty of the problem stems from the algebraic structure of the Bell expressions involved and the complexity of identifying a positive semidefinite matrix W such that Eq. (3.48) is fulfilled. Constructing such a matrix requires solving nontrivial semidefinite feasibility problems, often involving symbolic or exact arithmetic in operator algebras.

Nevertheless, we developed and implemented several computational techniques aimed at gaining partial information about the matrix W. These methods can help guide future attempts toward a full analytic decomposition and may assist in understanding the facial structure of the dual quantum set beyond numerical approximation.

3.6.2.1 Proposed techniques

When the NPA hierarchy at a given relaxation $\mathcal{T}_{\mathcal{B}}$ reaches a dual local polytope vertex, strong duality guarantees the existence of an SOS certificate for the corresponding Bell expression. Crucially, the monomial basis $\mathcal{B}$ for this SOS decomposition is precisely the same monomial basis established by the specific level of the NPA hierarchy necessary to reach that particular vertex. Moreover, the positive semidefinite matrix $W_{\mathcal{B}}$ of the SOS decomposition is exaclty the optimal solution to the dual SDP [70, 117]. This inherent link between the hierarchy level, the dual solution, and the SOS certificate provides a lot of information from numerics.

Symbolic approach. Our approach begins by fixing a monomial basis $\mathcal{B}$ for which the chosen vertex is reached. The cardinality of the basis provides an upper bound on the dimension of the problem—specifically, an upper bound on the dimension of the symmetric matrix W to be determined.

Working in the basis of nullifiers further reduces the complexity of the problem. Given a monomial basis $\mathcal{B}$, we search for linear combinations of the measurement operators (formally associated to monomials in $\mathcal{B}$) that vanish on the state $|\phi_\theta\rangle$. These combinations form the nullifiers. Parametrizing the general nullifier operator $\hat{N}_i$ in terms of the monomials in $\mathcal{B}$ and imposing the nullification condition

$$\hat{N}_i |\phi_\theta\rangle = 0 \quad (3.50)$$

yields a system of linear constraints on the coefficients of the nullifiers, thereby reducing the effective dimension of the operator space. Since $|\phi_\theta\rangle$ is a pure two-qubit state, there are at most four linearly independent nullification conditions.

For instance, in the space of indeterminates generated by the basis (3.49), a basis of nullifiers for $|\phi_\theta\rangle$ is given by

$$\boldsymbol{N}_\theta := \begin{pmatrix} \frac{A_0+A_1}{\sqrt{2}} - B_0 \\ \frac{A_0-A_1+\sqrt{2}(t_{2\theta}^{-1}+(1-s_{2\theta}^{-1})B_0-B_1)+c_{2\theta}s_\theta^{-1}c_\theta^{-1}A_1B_1+(2-s_\theta^{-1}c_\theta^{-1})A_1B_1B_0}{\sqrt{2}} \\ 1 - \frac{A_0B_0+A_1B_0}{\sqrt{2}} \\ \frac{s_\theta^{-1}c_\theta^{-1}+(1-t_\theta^{-1})B_0-\sqrt{2}A_0B_1+\sqrt{2}t_\theta A_1B_1}{2} \\ \frac{s_\theta^{-1}c_\theta^{-1}(-\sqrt{2}(s_{2\theta}-1)+\sqrt{2}c_{2\theta}B_0-(-2+s_{2\theta})A_1B_1+s_{2\theta}(A_0B_0+A_0B_1-A_1B_0)-2c_{2\theta}A_1B_1B_0)}{2\sqrt{2}} \\ B_1\left(\frac{1-A_0B_0-A_1B_0}{\sqrt{2}}\right) \\ \frac{t_\theta-t_\theta^{-1}+s_\theta^{-1}c_\theta^{-1}B_0-\sqrt{2}A_0B_0B_1+\sqrt{2}A_1B_0B_1}{2} \\ \left(\frac{1+A_0B_0+A_1B_0}{\sqrt{2}}\right)B_1 \\ \frac{+t_\theta-t_\theta^{-1}+2s_{2\theta}^{-1}B_0+\sqrt{2}A_0B_1B_0-\sqrt{2}A_1B_1B_0}{2} \end{pmatrix}. \tag{3.51}$$

Note that, for $\theta = \pi/4$, we recover exactly the nullifiers in (2.71).

Once a linearly independent set of nullifiers $\{N_i\}_i$ is constructed, we seek to express the shifted Bell operator as a quadratic form:

$$1 - \beta_{\theta,r_0,r_1} = \boldsymbol{N}^\dagger W \boldsymbol{N}, \tag{3.52}$$

where W is a real symmetric matrix.

We implement this procedure symbolically: for a fixed value of θ and a given vertex of the dual local polytope—corresponding to a fixed Bell expression β_{θ,r_0,r_1}- we consider a general real symmetric matrix $W \in \mathbb{R}^{d\times d}$, and impose the algebraic constraint (3.52) by matching both sides under operator algebra expansion.[2] This yields a system of linear equations in the entries of W, whose solutions define all symmetric matrices expressible as quadratic forms in the nullifier basis that reproduce the Bell expression $1 - \beta_{\theta,r_0,r_1}$. If among these solutions there exists one that is positive semidefinite (i.e., $W \succeq 0$), then this decomposition constitutes an exact analytic *certificate* of quantum saturation at the given vertex, verifying that the Bell expression achieves its Tsirelson bound on $|\phi_\theta\rangle$.

In practice, however, the resulting linear system typically does not uniquely determine all entries of W. Instead, it provides a set of necessary constraints: some entries are fixed by the decomposition, while others remain free. To complete such a partially determined matrix into

[2] A useful tool to perform this kind of computations with noncommutative variables is given by the Mathematica package `NCAlgebra` [125].

a valid SOS certificate, one must find a positive semidefinite (PSD) completion consistent with the algebraic constraints. This task is highly nontrivial for several reasons:

- Analytically parametrizing the cone of PSD matrices is notoriously difficult, especially in higher dimensions.
- Determining whether a partially specified symmetric matrix admits a PSD completion typically requires solving a semidefinite feasibility problem, which is computationally demanding and generally lacks closed-form criteria.
- Verifying positive semidefiniteness symbolically amounts to analyzing the eigenvalues of W, a process that becomes quickly intractable as the matrix size increases (e.g., for $n \geq 5$), due to the algebraic complexity of the characteristic polynomial.

Numerical aids. To gain further insight into the structure of W, we complemented our symbolic analysis with numerical methods. The NPA hierarchy algorithm (as implemented in the `ncpol2sdpa` library) provides not only the optimal value of the SDP, but also both the primal and dual solutions. The primal solution yields the moment matrix Γ, while the dual solution, as we've seen, gives a numerical approximation of the SOS matrix W, expressed in the monomial basis $\mathcal{B}$ used to define Γ. This introduces a subtle complication: the nullifier basis and the monomial basis differ in dimension. Specifically, the nullifier basis spans only a subspace of the full operator space associated with the moment matrix. As a result, while any nullifier can be expanded in terms of monomials—allowing symbolic expressions to be translated into the monomial basis—the converse is not generally possible. There is no one-to-one correspondence between the two. Consequently, the numerical matrix W obtained from the SDP solver cannot always be directly interpreted in the nullifier basis. However, it can still inform the symbolic search through a one-way change of basis.

By comparing the symbolic and numerical representations of W in the monomial basis $\mathcal{B}$, we can extract valuable structural insights: patterns of zeros, symmetries between entries, relative magnitudes, and potential rank deficiencies. These observations help refine the symbolic ansatz for W, reducing the number of degrees of freedom and guiding the construction of the SOS decomposition. Nevertheless, in our case, this information remains insufficient to obtain a complete analytic solution.

One final route exploits a subtle observation: since the Bell expressions we consider correspond to vertices of the dual local polytope, there exist primal solutions associated with local deterministic strategies. In this case, the monomials—built from unitary operators $\hat{A}_x$, $\hat{B}_y$—evaluate to either $+1$ or -1, and the resulting moment matrix Γ has entries in $\{\pm 1\}$. If the SDP solver returns such a matrix, we recover an analytic primal solution directly. Moreover, since the solver reports optimal convergence status, we can invoke *strong duality* [126], which guarantees a relationship between the primal and dual solutions.

Due to strong duality, optimal convergence of the SDP ensures that the *complementary slackness condition* [126] is satisfied:

$$\Gamma W = O_d, \tag{3.53}$$

where O_d is the zero matrix of dimension d. This condition yields further constraints on the structure of W, offering an additional tool for symbolic inference. Still, despite these insights, the full analytic construction of the SOS decomposition remains a non-trivial problem.

Thus, while our approach enables a partial characterization of the SOS matrix and reveals meaningful structural features, a full analytic derivation remains an open challenge. Future efforts may explore symbolic-numeric hybrid strategies or seek alternative operator bases better adapted to the monomial structure imposed by the SDP relaxation.

Conclusions

In this thesis, we explored the structure of quantum correlations in the CHSH scenario through the lens of convex geometry and duality theory. Our focus was on the local geometry of the quantum set $\mathcal{Q}$ in the CHSH scenario near its extremal points. Specifically, we characterized the supporting hyperplanes associated with a family of extremal quantum behaviors. This approach is motivated by the fact that $\mathcal{Q}$, being convex, is entirely determined by its extremal points:

$$\mathcal{Q} = \text{conv}(\mathcal{Q}_{\text{ext}}),$$

and every boundary-defining hyperplane supports the set at one or more of these points.

In Chapter 1, we introduced the formalism of *Bell scenarios* and reviewed the foundational aspects of *nonlocality*, emphasizing the *device-independent paradigm*. We presented the key classes of correlations—local, quantum, and no-signaling—and described their geometric relationships. Particular attention was given to the CHSH scenario, whose minimality allows for a detailed and insightful analysis of the geometry of correlations and the associated Bell inequalities. We also reviewed the concepts of *self-testing* and their foundational implications for quantum theory.

Chapter 2 provided an introduction to the *dual perspective* on the quantum set, developing a mathematical framework rooted in convex duality. We introduced tools such as the variational method and the Sum-of-Squares (SOS) decomposition to characterize Bell expressions maximized by a given quantum point. In particular, we reviewed the results of Barizien *et al.* [52] that achieved the complete study of the orthogonal face of the Tsirelson point. There, the analysis benefited from the high symmetry of the point, allowing a complete characterization of the Bell expressions that define its supporting hyperplane.

In Chapter 3, we extended this approach to a broader family of nonlocal extremal points in the CHSH scenario. This chapter contains the original contributions of this thesis. We introduced a general method to characterize orthogonal faces of the quantum set, applicable to arbitrary nonlocal extremal points, not just the Tsirelson point. By analyzing nullspace conditions and variational constraints, we derived a parametric description of the Bell expressions maximized by a given extremal point.

The second half of the chapter focused on computing both local and quantum bounds associated with these Bell expressions. For the local bounds, we studied the projection of the dual local polytope and extracted its vertices, obtaining explicit analytical formulas and visualizations. For the quantum bounds, we employed the NPA hierarchy to get numerical

estimation of the boundaries of the dual quantum set $\mathcal{Q}^*$. Our results are summarized in fig. 3.9. We also developed some hybrid (both numerical and analytic) techniques to find an SOS decomposition for the Bell expressions associated to the vertices, in order to confirm analytically our numerical results.

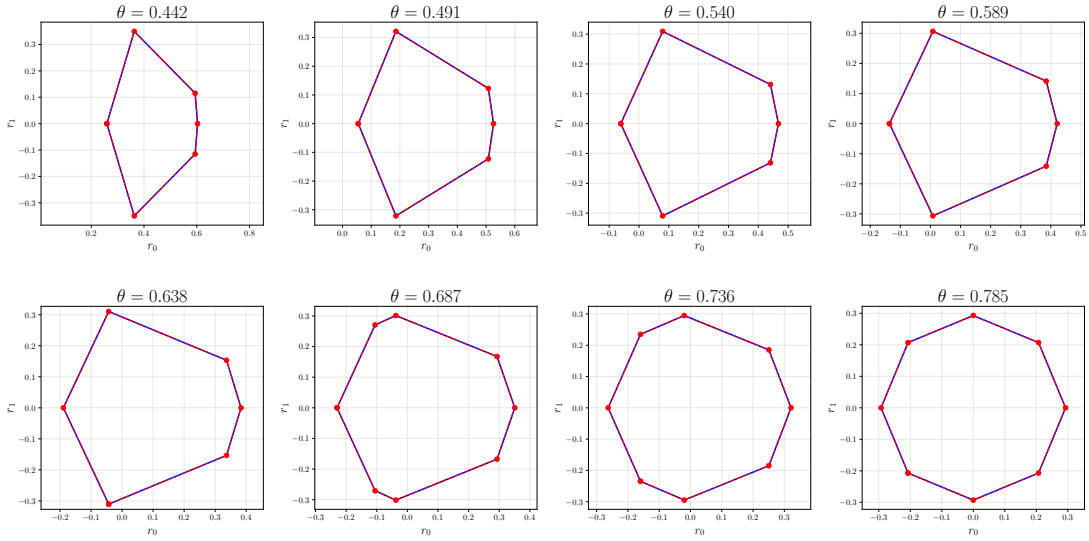

Figure 3.9: Summary visualization of the main results of this thesis. The plot shows the numerically estimated dual quantum boundary (in blue) at the level $\mathcal{T}_{1+A+B+AB+AA'B+ABB'}$ of the NPA hierarchy, for values of $\theta \in (\pi/8, \pi/4]$, alongside the dual local boundary (in red) in the (r_0, r_1)-plane. Since the NPA hierarchy approximates $\mathcal{Q}^*$ from the inside, and we know that $\mathcal{Q}^* \subseteq \mathcal{L}^*$, the fact that the numerical boundary of $\mathcal{Q}^*$ reaches that of $\mathcal{L}^*$ in this projection implies, up to numerical precision, that the two sets coincide in the (r_0, r_1) plane. This observation supports our central conjecture: for all $\theta \in (\pi/8, \pi/4]$, the dual quantum boundary $\mathcal{Q}^*$ saturates the local dual boundary $\mathcal{L}^*$ in this projection.

These results offer a refined understanding of the geometry of the quantum set and its boundary. They generalize previous work and demonstrate how symbolic and numerical methods can be combined to analyze the extremal structure of quantum correlations. The techniques developed here may also have implications for the classification of quantum behaviors in more complex Bell scenarios, and for the design of tailored Bell inequalities for self-testing and certification tasks.

In conclusion, this thesis contributes to the ongoing effort to understand the mathematical structure of quantum correlations and the principles that constrain them. While many questions remain open—such as the full characterization of $\mathcal{Q}^*$ or the identification of operational principles that single out $\mathcal{Q}$ among all no-signaling sets—the tools and results presented here constitute a step toward a deeper and more constructive description of quantum nonlocality.

Appendix A

Essentials of quantum information

This appendix collects several foundational results from quantum information theory that are used throughout the thesis. While these concepts are standard, we include them here for completeness and to provide a self-contained presentation of the formal tools underpinning our analysis.

A.1 Quantum measurements

We start by recalling the measurement postulate of quantum mechanics [127]:

Postulate A.1 (Measurement postulate). *A general quantum measurement is described by a collection of operators* $\{\hat{\Lambda}_x\}$*, called* measurement operators*, acting on the Hilbert space* $\mathcal{H}$ *of the system. The index* x *labels the possible outcomes of the measurement. If the system is in the pure state* $|\psi\rangle \in \mathcal{H}$ *immediately before the measurement, then:*

- *The probability of obtaining outcome* x *is given by* Born's rule*:*

$$P(x) = \langle\psi| \hat{\Lambda}_x^\dagger \hat{\Lambda}_x |\psi\rangle . \tag{A.1}$$

- *The* post-measurement state*, conditional on outcome* x *being observed, is:*

$$|\psi'\rangle \coloneqq \frac{\hat{\Lambda}_x |\psi\rangle}{\sqrt{\langle\psi| \hat{\Lambda}_x^\dagger \hat{\Lambda}_x |\psi\rangle}}. \tag{A.2}$$

The measurement operators must satisfy the completeness relation:

$$\sum_x \hat{\Lambda}_x^\dagger \hat{\Lambda}_x = \mathbb{1}, \tag{A.3}$$

which ensures that the probabilities $P(x)$ sum to 1.

A.1.1 PVM measurements

A special case of quantum measurements is that of *Projective-Valued Measurements (PVM)* or simply *projective measurements*, which are associated with self-adjoint observables. Let $\hat{O} \in L(\mathcal{H})$ be an observable with spectral decomposition:

$$\hat{O} = \sum_x x\hat{\Pi}_x, \tag{A.4}$$

where x are the eigenvalues of $\hat{O}$, and $\hat{\Pi}_x$ is the orthogonal projector onto the eigenspace corresponding to x.

If the system is prepared in a pure state $|\psi\rangle \in \mathcal{H}$, the projective measurement defined by $\hat{O}$ has the following properties:

- The probability of obtaining outcome x is:

$$P_O(x) = \langle\psi|\,\hat{\Pi}_x\,|\psi\rangle\,.$$

- The post-measurement state, given that outcome x was obtained, is:

$$|\psi'\rangle \coloneqq \frac{\hat{\Pi}_x\,|\psi\rangle}{\sqrt{\langle\psi|\,\hat{\Pi}_x\,|\psi\rangle}}.$$

Projective measurements satisfy the following properties:

$$\text{(Orthogonality)} \quad \hat{\Pi}_x\hat{\Pi}_{x'} = \delta_{xx'}\hat{\Pi}_x, \tag{A.5a}$$

$$\text{(Completeness)} \quad \sum_x \hat{\Pi}_x = \mathbb{1}, \tag{A.5b}$$

$$\text{(Hermiticity)} \quad \hat{\Pi}_x = \hat{\Pi}_x^\dagger. \tag{A.5c}$$

Hence, projective measurements constitute a special case of general measurements where the measurement operators $\hat{\Lambda}_x$ are orthogonal projectors.

Projective measurements have many nice properties. In particular, it is very easy to calculate average values for projective measurements. By definition, the average value of the measurement is

$$\langle\hat{O}\rangle \coloneqq \sum_x xP_O(x) = \sum_x x\,\langle\psi|\,\hat{\Pi}_x\,|\psi\rangle = \langle\psi|\left(\sum_x x\hat{\Pi}_x\right)|\psi\rangle = \langle\psi|\,\hat{O}\,|\psi\rangle\,.$$

A.1.2 POVM measurements

In some experimental scenarios, the post-measurement state is not of interest—only the statistics of outcomes matter. In such cases, the measurement can be described using the *POVM (Positive Operator-Valued Measurement) formalism*.

Given a general measurement defined by operators $\{\Lambda_x\}$, we define the associated POVM elements as:

$$\hat{M}_x \coloneqq \hat{\Lambda}_x^\dagger\hat{\Lambda}_x. \tag{A.6}$$

Each $\hat{M}_x$ is a positive semidefinite operator, and the completeness relation becomes:

$$\sum_x \hat{M}_x = \mathbb{1}. \tag{A.7}$$

5 The probability of obtaining outcome x when measuring a state $|\psi\rangle$ is:

$$P(x) = \langle\psi| \hat{M}_x |\psi\rangle . \tag{A.8}$$

The collection $\{\hat{M}_x\}$ is called a *POVM*. It encodes all the probabilistic information about the measurement, without reference to the post-measurement state.

As a particular case, a projective measurement $\{\hat{\Pi}_x\}$ satisfies:

$$\hat{\Pi}_x^2 = \hat{\Pi}_x, \quad \hat{\Pi}_x^\dagger = \hat{\Pi}_x, \quad \hat{\Pi}_x\hat{\Pi}_{x'} = \delta_{xx'}\hat{\Pi}_x, \quad \sum_x \hat{\Pi}_x = \mathbb{1}, \tag{A.9}$$

In this instance (and only this instance) all the POVM elements are the same as the measurement operators themselves, since $\hat{M}_x \equiv \hat{\Pi}_x^\dagger\hat{\Pi}_x = \hat{\Pi}_x$.

A.2 Elementary entanglement theory

Quantum entanglement, first discussed by Schrödinger [128] and by Einstein, Podolsky, and Rosen [4], has long been at the heart of foundational debates in quantum mechanics. From a modern perspective, entanglement is not only a striking nonclassical feature but also a central resource in quantum information science. The structure of the set of entangled states is particularly rich and complex, especially in the multipartite setting. In this section, we review basic notions from entanglement theory that are relevant for the rest of the thesis. A comprehensive and authoritative treatment can be found in [129].

Definition A.1. Let $\mathcal{H} \coloneqq \bigotimes_{i=1}^n \mathcal{H}_i$ be the Hilbert space of a composite quantum system. A *pure* state $|\psi\rangle \in \mathcal{H}$ is called *entangled* if it cannot be written as a product state of the form

$$|\psi\rangle = \bigotimes_{i=1}^n |\psi_i\rangle , \quad \text{with } |\psi_i\rangle \in \mathcal{H}_i. \tag{A.10}$$

Otherwise, $|\psi\rangle$ is said to be *separable*.

In the special case of a *bipartite* pure state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, there exists a canonical representation known as the *Schmidt decomposition*,

Proposition A.1 (Schmidt decomposition). *Given any bipartite state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, there exist two orthonormal bases $\{|\alpha_j\rangle\}_{j=1}^{d_A}$ and $\{|\beta_k\rangle\}_{k=1}^{d_B}$ for $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively, such that*

$$|\psi\rangle = \sum_{i=1}^{\min(d_A,d_B)} \sqrt{p_i}\, |\alpha_i\rangle \otimes |\beta_i\rangle \tag{A.11}$$

with $p_i \geq 0$ for all i and $\sum_i p_i = 1$.

If a bipartite pure state has only one nonzero *Schmidt coefficient*, it is a *product state* of the form $|\psi\rangle = |\alpha\rangle \otimes |\beta\rangle$. Conversely, if more than one Schmidt coefficient is nonzero, the state is entangled.

In the case of two qubits, a state is said to be *maximally entangled* if both Schmidt coefficients are equal, i.e., $p_1 = p_2 = 1/\sqrt{2}$. Choosing a shared orthonormal basis $\{|0\rangle, |1\rangle\}$ for both subsystems, the canonical maximally entangled state is given by

$$\left|\phi^+\right\rangle := \frac{1}{\sqrt{2}} \left(|00\rangle + |11\rangle\right). \tag{A.12}$$

In the multipartite setting, the notion of entanglement becomes more intricate. Several inequivalent generalizations of the Schmidt decomposition exist, and entanglement classification becomes significantly more complex. For a detailed analysis in the case of three qubits, see, e.g., [130].

Definition A.2. A *mixed state* $\hat{\rho} \in L(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n)$ is said to be *separable* if it can be written as a convex combination of product states:

$$\hat{\rho} = \sum_i p_i \left(\bigotimes_{j=1}^{n} \hat{\rho} j, i \right), \tag{A.13}$$

where each $\hat{\rho} j, i$ is a density operator on $\mathcal{H}_j$, $p_i \geq 0$, and $\sum_i p_i = 1$.

If no such decomposition exists, the state $\hat{\rho}$ is called *entangled*.

A separable state that is not a product state is often referred to as *classically correlated*, as it may exhibit statistical correlations without quantum entanglement.

While determining whether a pure state is entangled is straightforward, no general algorithm is known for deciding the separability of an arbitrary mixed state. A state can be proven to be separable by explicitly exhibiting a decomposition into product states.

A.3 Purification and Naimark dilation

A.3.1 Purification of a mixed state

Given a quantum system S described by a mixed state $\hat{\rho}_S$, it is always possible to represent this state as the reduced state of a pure state defined on a larger Hilbert space. This process is known as *purification* [131], and it plays a central role in quantum information theory, where pure state methods often simplify the analysis of mixed states and quantum channels.

Definition A.3. Let $\hat{\rho}_S \in L(\mathcal{H}_S)$ be a density operator. A *purification* of $\hat{\rho}_S$ is a pure state $|\psi\rangle_{RS} \in \mathcal{H}_R \otimes \mathcal{H}_S$, where R is an auxiliary system (the *reference system*), such that

$$\hat{\rho}_S = \mathrm{Tr}_R \left(|\psi\rangle_{RS} \langle\psi|_{RS} \right). \tag{A.14}$$

Theorem A.1 (Purification theorem). *Let $\hat{\rho}_S \in L(\mathcal{H}_S)$ admit the spectral decomposition*

$$\hat{\rho}_S = \sum_x P(x) |x\rangle_S \langle x|_S .$$

Then a purification of $\hat{\rho}_S$ is given by

$$|\psi\rangle_{RS} \coloneqq \sum_x \sqrt{P(x)}\, |x\rangle_R \otimes |x\rangle_S \,,$$

where $\{|x\rangle_R\}_x$ is any orthonormal basis for the reference system R.

Remark A.1. Purifications are not unique: if $|\psi\rangle_{RS}$ and $|\phi\rangle_{RS}$ are two purifications of $\hat{\rho}_S$, then there exists a unitary $\hat{U}_R \in L(\mathcal{H}_R)$ on the reference system such that

$$|\phi\rangle_{RS} = (\hat{U}_R \otimes \mathbb{1}_S)\, |\psi\rangle_{RS} \,.$$

Physically, the purification framework allows us to interpret the mixedness of $\hat{\rho}_S$ as arising from entanglement with an inaccessible reference system R. In this sense, $\hat{\rho}_S$ is the marginal of a globally pure entangled state, and the statistical uncertainty associated with $\hat{\rho}_S$ reflects our lack of access to the reference.

A.3.2 Naimark's dilation

The purification idea extends naturally to quantum measurements. Just as every mixed state can be seen as the marginal of a pure state on a larger system, every generalized measurement (POVM) can be modeled as a projective measurement on a larger Hilbert space. This is formalized by the *Naimark dilation theorem* [51, 132–135].

Theorem A.2 (Naimark dilation for the CHSH scenario). *Any quantum correlation in the CHSH scenario can be realized using projective measurements on an enlarged Hilbert space. That is, for every $\boldsymbol{P} \in \mathcal{Q}$, there exist:*

1. *a quantum state $\hat{\rho}_{AB}$ on $\mathcal{H}_A \otimes \mathcal{H}_B$,*
2. *auxiliary (ancillary) Hilbert spaces $\mathcal{H}_{A'}$ and $\mathcal{H}_{B'}$,*
3. *projective binary measurements $\{\hat{\Pi}^{AA'}_{a|x}\} \subset L(\mathcal{H}_A \otimes \mathcal{H}_{A'})$ and $\{\hat{\Pi}^{BB'}_{b|y}\} \subset L(\mathcal{H}_B \otimes \mathcal{H}_{B'})$,*

such that

$$P(a,b|x,y) = \mathrm{Tr}\left[(\hat{\rho}_{AB} \otimes \hat{\omega}_{A'B'})\, (\hat{\Pi}^{AA'}_{a|x} \otimes \hat{\Pi}^{BB'}_{b|y})\right], \tag{A.15}$$

where $\hat{\omega}_{A'B'}$ is an ancillary state, and the projective measurements act jointly on system and ancilla.

This result follows from the general Naimark dilation theorem, which states that any POVM can be implemented as a projective measurement on a larger system that includes an ancilla. In the CHSH scenario, this implies that general quantum correlations can always be reproduced by projective measurements on an extended (purified) state, without loss of generality.

A.4 The no-communication theorem

Although quantum entanglement gives rise to striking nonlocal correlations, it does not allow for faster-than-light communication [89, 93]. This fundamental limitation, formalized by the no-communication theorem, played a key role in clarifying the tension between quantum mechanics and relativity.

Theorem A.3 (No-communication theorem). *Let two spatially separated observers, Alice and Bob, share a bipartite quantum state $\hat{\rho}_{AB}$ on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. Suppose Alice performs a measurement on her subsystem described by a POVM $\{\hat{M}_a\}$, where each $\hat{M}_a \succeq 0$ and $\sum_a \hat{M}_a = \mathbb{1}_A$. Then, the statistics observed by Bob, described by the reduced state*

$$\hat{\rho}_B = \mathrm{Tr}_A[\hat{\rho}_{AB}],$$

remain unchanged regardless of which measurement Alice performs or whether she performs one at all. In particular, no local measurement by Alice can influence the marginal statistics observed by Bob.

Proof. Let $\hat{\rho}_{AB}$ be the shared bipartite state on $\mathcal{H}_A \otimes \mathcal{H}_B$, and let $\{\hat{M}_a\}$ be a POVM that Alice may choose to perform.

After Alice performs her measurement and obtains outcome a, the (unnormalized) post-measurement state of the global system becomes

$$\hat{\rho}_{AB}^{(a)} = (\hat{M}_a \otimes \mathbb{1}_B)\hat{\rho}_{AB}(\hat{M}_a \otimes \mathbb{1}_B)^\dagger.$$

The corresponding (normalized) conditional state on Bob's side is then

$$\hat{\rho}_B^{(a)} = \frac{\mathrm{Tr}_A[\hat{\rho}_{AB}^{(a)}]}{\mathrm{Tr}\left[\hat{\rho}_{AB}^{(a)}\right]}.$$

However, Bob does not have access to Alice's outcome a, so the relevant state for Bob is the ensemble average over all possible outcomes:

$$\hat{\tilde{\rho}}_B = \sum_a \mathrm{Tr}_A\left[(\hat{M}_a \otimes \mathbb{1}_B)\hat{\rho}_{AB}(\hat{M}_a \otimes \mathbb{1}_B)^\dagger\right].$$

Since each $\hat{M}_a$ is a positive operator and the identity acts trivially on Bob's system, we can simplify:

$$\hat{\tilde{\rho}}_B = \mathrm{Tr}_A\left[\sum_a (\hat{M}_a \otimes \mathbb{1}_B)\hat{\rho}_{AB}(\hat{M}_a \otimes \mathbb{1}_B)^\dagger\right].$$

If Alice's measurement is projective (or more generally, if the $\hat{M}_a$ are Hermitian), then we can write:

$$\hat{\tilde{\rho}}_B = \mathrm{Tr}_A\left[\sum_a (\hat{M}_a^2 \otimes \mathbb{1}_B)\hat{\rho}_{AB}\right] = \mathrm{Tr}_A\left[\left(\sum_a \hat{M}_a^2 \otimes \mathbb{1}_B\right)\hat{\rho}_{AB}\right].$$

In the general case, we write:

$$\hat{\tilde{\rho}}_B = \mathrm{Tr}_A\left[\sum_a (\hat{M}_a \otimes \mathbb{1}_B)\hat{\rho}_{AB}(\hat{M}_a \otimes \mathbb{1}_B)\right].$$

Regardless, since $\sum_a \hat{M}_a = \mathbb{1}_A$, the completeness of the POVM implies:

$$\hat{\tilde{\rho}}_B = \mathrm{Tr}_A \left[(\mathbb{1}_A \otimes \mathbb{1}_B)\hat{\rho}_{AB}\right] = \mathrm{Tr}_A[\hat{\rho}_{AB}] = \hat{\rho}_B.$$

Therefore, the reduced state on Bob's side remains unchanged, irrespective of whether Alice performs a measurement, and regardless of which measurement she performs. Hence, no information can be transmitted from Alice to Bob solely through quantum measurements on a shared entangled state. □

Appendix B

Convex geometry

Convex geometry plays a foundational role in many areas of science and engineering, and is particularly central in quantum information theory, where it underpins the study of state spaces, correlations, and optimization [136].

This appendix provides a self-contained summary of the main concepts and results from convex geometry that are used throughout the thesis. The aim is to give the reader the necessary mathematical background to follow the geometric and duality-based analysis of quantum correlations presented in the main chapters.

All results are stated in finite-dimensional real vector spaces, and proofs are included when illuminating or necessary for the main text. Standard references include [64, 126, 137].

B.1 Affine sets

Definition B.1. A set $\mathcal{A} \subseteq \mathbb{R}^n$ is *affine* if for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{A}$ and $\lambda \in \mathbb{R}$, we have that

$$\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2 \in \mathcal{A}. \tag{B.1}$$

We refer to a point of the form $\sum_i \lambda_i \boldsymbol{x}_i$, with $\lambda_i \in \mathbb{R}$ and $\sum_i \lambda_i = 1$ as an *affine combination* of the the points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$.

If $\mathcal{A}$ is an affine set and $\boldsymbol{x}_0 \in \mathcal{A}$, then the set

$$V := \mathcal{A} - \boldsymbol{x}_0 = \{\boldsymbol{x} - \boldsymbol{x}_0 \colon \boldsymbol{x} \in \mathcal{A}\} \tag{B.2}$$

is a subspace, i.e., closed under sums and scalar multiplication. To see this, suppose $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$ and $\alpha, \beta \in \mathbb{R}$. Then we have $\boldsymbol{v}_1 + \boldsymbol{x}_0 \in \mathcal{A}$ and $\boldsymbol{v}_2 + \boldsymbol{x}_0 \in \mathcal{A}$, and so

$$\alpha \boldsymbol{v}_1 + \beta \boldsymbol{v}_2 + \boldsymbol{x}_0 = \alpha(\boldsymbol{v}_1 + \boldsymbol{x}_0) + \beta(\boldsymbol{v}_2 + \boldsymbol{x}_0) + (1 - \alpha - \beta)\boldsymbol{x}_0 \in \mathcal{A},$$

since $\mathcal{A}$ is affine, and $\alpha + \beta + (1 - \alpha - \beta) = 1$. We conclude that $\alpha \boldsymbol{v}_1 + \beta \boldsymbol{v}_2 \in V$, since $\alpha \boldsymbol{v}_1 + \beta \boldsymbol{v}_2 + \boldsymbol{x}_0 \in \mathcal{A}$.

Thus, the affine set $\mathcal{A}$ can be expressed as

$$\mathcal{A} = V + \boldsymbol{x}_0 = \{\boldsymbol{v} + \boldsymbol{x}_0 \colon \boldsymbol{v} \in V\},$$

i.e., as a subspace plus an *offset*. The subspace V associated with the affine set $\mathcal{A}$ does not depend on the choice of $\boldsymbol{x}_0$, so $\boldsymbol{x}_0$ can be chosen as any point in $\mathcal{A}$.

Definition B.2. We define the *dimension of an affine set* $\mathcal{A}$ as the dimension of the subspace $V = \mathcal{A} - \boldsymbol{x}_0$, where $\boldsymbol{x}_0$ is any element of $\mathcal{A}$.

Definition B.3. The set of all affine combinations of points in some set $C \subseteq \mathbb{R}^n$ is called the *affine hull* of C, and denoted $\operatorname{aff} C$:

$$\operatorname{aff} C \coloneqq \left\{ \sum_i \lambda_i \boldsymbol{x}_i \,:\, \boldsymbol{x}_i \in C \quad \forall i, \quad \sum_i \lambda_i = 1 \right\}. \tag{B.3}$$

The affine hull is the smallest affine set that contains C: if $\mathcal{A}$ is any affine set with $C \subseteq \mathcal{A}$, then $\operatorname{aff} C \subseteq \mathcal{A}$.

B.2 Convex sets

B.2.1 Basics

Definition B.4. A set $\mathcal{K} \subseteq \mathbb{R}^n$ is *convex* if for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{K}$ and $\lambda \in [0, 1]$, we have that

$$\lambda \boldsymbol{x}_1 + (1 - \lambda)\boldsymbol{x}_2 \in \mathcal{K}. \tag{B.4}$$

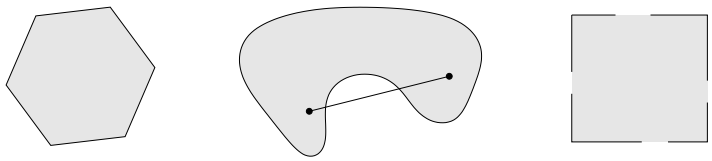

Figure B.1: Some simple convex and nonconvex sets.

We call a point of the form $\sum_i \lambda_i \boldsymbol{x}_i$, where $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$ a *convex combination* of the points $x_1, \ldots, x_k$. As with affine sets, it can be shown that a set is convex if and only if it contains every convex combination of its points. A convex combination of points can be thought of as a mixture or weighted average of the points, with λ_j the fraction of $\boldsymbol{x}_j$ in the mixture.

Definition B.5. The *convex hull* of a set C, denoted $\operatorname{conv} C$, is the set of all convex combinations of points in C:

$$\operatorname{conv} C \coloneqq \left\{ \sum_i \lambda_i \boldsymbol{x}_i \,:\, \boldsymbol{x}_i \in C, \lambda_i \geq 0, \forall i, \sum_i \lambda_i = 1 \right\}. \tag{B.5}$$

As the name suggests, the convex hull $\operatorname{conv} C$ is always convex. It is the smallest convex set that contains C: if $\mathcal{K}$ is any convex set that contains C, then $\operatorname{conv} C \subseteq \mathcal{K}$.

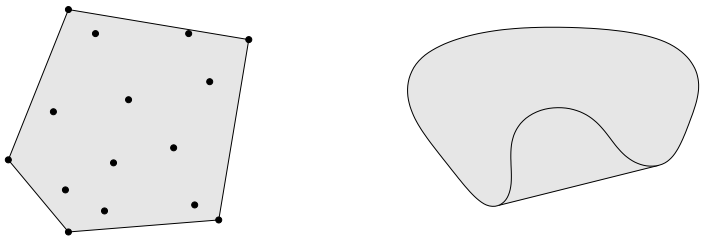

Figure B.2: The convex hulls of two sets in $\mathbb{R}^2$. On the left, the convex hull of a set of fifteen points forms a pentagon. On the right, the convex hull of a kidney-shaped set.

B.2.2 Separating and supporting hyperplanes

Theorem B.1 (Hyperplane Separation Theorem). *Let $\mathcal{K} \subset \mathbb{R}^n$ be a nonempty convex set, and let $\boldsymbol{x} \notin \mathcal{K}$. Then there exists a nonzero vector $\boldsymbol{f} \in \mathbb{R}^n$ and a scalar $\alpha \in \mathbb{R}$ such that*

$$\boldsymbol{f} \cdot \boldsymbol{y} \leq \alpha < \boldsymbol{f} \cdot \boldsymbol{x} \quad \textit{for all } \boldsymbol{y} \in \mathcal{K}. \tag{B.6}$$

That is, there exists a hyperplane that separates the point $\boldsymbol{x}$ strictly from the convex set $\mathcal{K}$.

Definition B.6. Let $\mathcal{K} \subset \mathbb{R}^n$ be a convex set, and let

$$H \coloneqq \{\boldsymbol{x} \in \mathbb{R}^n \colon \boldsymbol{f} \cdot \boldsymbol{x} = 1\}$$

be a hyperplane with normal vector $\boldsymbol{f} \in \mathbb{R}^n$. We say that H is a *supporting hyperplane* of $\mathcal{K}$ if:

- $H \cap \mathcal{K} \neq \emptyset$, and
- $\mathcal{K} \subseteq H_s$, where $H_s \coloneqq \{\boldsymbol{x} \in \mathbb{R}^n \colon \boldsymbol{f} \cdot \boldsymbol{x} \leq 1\}$ is the corresponding closed halfspace.

Remark B.1. Let $\boldsymbol{f} \in \mathbb{R}^n$ be a linear functional. A hyperplane with normal vector $\boldsymbol{f}$ supports $\mathcal{K}$ at a point $\boldsymbol{x} \in \mathcal{K}$ if

$$\boldsymbol{f} \cdot \boldsymbol{x} = \sup_{\boldsymbol{y} \in \mathcal{K}} \boldsymbol{f} \cdot \boldsymbol{y}.$$

This is equivalent to the previous definition by setting $1 = \sup_{\boldsymbol{y} \in \mathcal{K}} \boldsymbol{f} \cdot \boldsymbol{y}$, so that the hyperplane

$$H \coloneqq \{\boldsymbol{y} \in \mathbb{R}^n \colon \boldsymbol{f} \cdot \boldsymbol{y} = 1\}$$

intersects $\mathcal{K}$ at $\boldsymbol{x}$ and contains $\mathcal{K}$ entirely in the halfspace $\{\boldsymbol{y} \in \mathbb{R}^n \colon \boldsymbol{f} \cdot \boldsymbol{y} \leq 1\}$. Thus, this variational formulation provides an equivalent characterization of supporting hyperplanes in terms of linear optimization over convex sets.

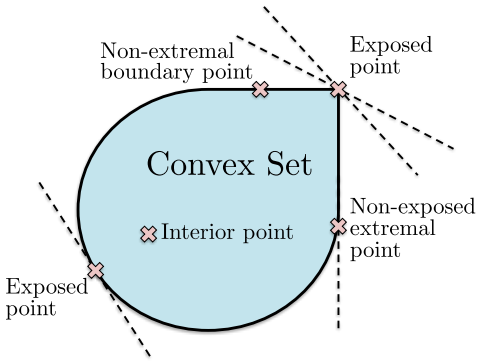

Figure B.3: Different types of points of a compact convex set. From [46].

B.2.3 Characterization of extremal points

In this thesis, we are mostly concerned with *convex bodies*.

Definition B.7. A *convex body* is a bounded closed convex sets with non-empty interior.

Let $\mathcal{K} \subset \mathbb{R}^n$ be a convex body. Then,

Definition B.8. A point $\boldsymbol{u} \in \mathcal{K}$ is called a *boundary point* if it belongs to the topological boundary of $\mathcal{K}$; that is, it does not lie in the interior of the set.

We denote the set of all boundary points of $\mathcal{K}$ by $\mathcal{K}_{\text{bnd}}$.

Definition B.9. A boundary point is said to be *extremal* (or *pure*) if it cannot be written as a non-trivial convex combination of other points in $\mathcal{K}$.

Note that pure points always lie in the boundary of the convex set, but the boundary usually contains non-pure points as well. All non-extremal points are called *mixed*. The set of all extremal points of $\mathcal{K}$ is denoted $\mathcal{K}_{\text{ext}}$. A stronger condition is that of *exposedness*:

Definition B.10. A point $\boldsymbol{u} \in \mathcal{K}$ is *exposed* if there exists a vector $\boldsymbol{v} \in \mathbb{R}^n$ such that $\boldsymbol{u}$ uniquely maximizes the inner product $\boldsymbol{v} \cdot \boldsymbol{w}$ over all $\boldsymbol{w} \in \mathcal{K}$. That is, the set

$$F_{\boldsymbol{v}} := \left\{ \boldsymbol{w} \in \mathcal{K} \colon \boldsymbol{v} \cdot \boldsymbol{w} = \max_{\boldsymbol{z} \in \mathcal{K}} \boldsymbol{v} \cdot \boldsymbol{z} \right\} \tag{B.7}$$

is reduced to the singleton $\{\boldsymbol{u}\}$.

In this case, the hyperplane defined by $\boldsymbol{v} \cdot \boldsymbol{w} = \max_{\boldsymbol{z} \in \mathcal{K}} \boldsymbol{v} \cdot \boldsymbol{z}$ is a supporting hyperplane of $\mathcal{K}$, and the set $F_{\boldsymbol{v}}$ is referred to as the *exposed face of* $\mathcal{K}$ *associated to* $\boldsymbol{v}$. We denote by $\mathcal{K}_{\text{exp}}$ the set of exposed points of $\mathcal{K}$.

By construction, we always have the chain of inclusions:

$$\mathcal{K}_{\text{exp}} \subseteq \mathcal{K}_{\text{ext}} \subseteq \mathcal{K}_{\text{bnd}} \subseteq \mathcal{K}, \tag{B.8}$$

and in general, each inclusion is strict. This hierarchy is illustrated in fig. B.3.

An important structural result in convex geometry is provided by the *Krein-Milman theorem* [64]:

Theorem B.2 (Krein-Milman theorem). *Any convex body $\mathcal{K}$ is equal to the convex hull of its extremal points:*

$$\mathcal{K} = \operatorname{conv}(\mathcal{K}_{\text{ext}}). \tag{B.9}$$

Remark B.2. Theorem B.2 implies that optimization of a linear functional over $\mathcal{K}$ always attains its maximum at an extremal point.

Proof. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a linear functional, $\bar{\boldsymbol{x}} \in \mathcal{K}$ be a maximizer of f over $\mathcal{K}$. We want to show that at least one maximizer is an extremal point.

By the Krein-Milman theorem, the maximizer $\bar{\boldsymbol{x}}$ can be written as a convex combination of extremal points. Thus, there exists a finite collection $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_k\} \subset \mathcal{K}_{\text{ext}}$ and weights $p_i \geq 0$, $\sum_{i=1}^k p_i = 1$, such that

$$\bar{\boldsymbol{x}} = \sum_{i=1}^k p_i \boldsymbol{x}_i.$$

Since f is linear,

$$f(\bar{\boldsymbol{x}}) = f\left(\sum_{i=1}^k p_i \boldsymbol{x}_i\right) = \sum_{i=1}^k p_i f(\boldsymbol{x}_i).$$

But $f(\bar{\boldsymbol{x}})$ is the maximum of f over $\mathcal{K}$, and all $\boldsymbol{x}_i \in \mathcal{K}$, so:

$$f(\boldsymbol{x}_i) \leq f(\bar{\boldsymbol{x}}) \quad \text{for all } i.$$

Putting this into the convex sum:

$$f(\bar{\boldsymbol{x}}) = \sum_{i=1}^k p_i f(\boldsymbol{x}_i) \leq \sum_{i=1}^k p_i f(\bar{\boldsymbol{x}}) = f(\bar{\boldsymbol{x}}).$$

But, a convex combination of numbers less than or equal to $f(\bar{\boldsymbol{x}})$ cannot equal $f(\bar{\boldsymbol{x}})$ unless every term in the sum with positive weight equals $f(\bar{\boldsymbol{x}})$. Hence:

$$f(\boldsymbol{x}_i) = f(\bar{\boldsymbol{x}}) \quad \text{for all } i \text{ such that } p_i > 0.$$

Therefore, each $\boldsymbol{x}_i$ (with $p_i > 0$) is an extremal point and a maximizer of f, since $f(\boldsymbol{x}_i) = f(\bar{\boldsymbol{x}})$. Thus, at least one extremal point of $\mathcal{K}$ attains the maximum of f over $\mathcal{K}$. □

Note that the maximizer need not be unique, nor does it need to be an exposed point. This leads to the distinction between extremal and exposed points: although every exposed point is extremal, the converse is not true in general. Still, a classical result due to Straszewicz [64] shows that in finite-dimensional spaces, the set of exposed points is dense in the set of extremal points. That is, extremal but non-exposed points exist, but they are in a certain sense exceptional.

Although we need all extreme points to describe the interior of a convex body as its convex combinations, a single point always can be written as a convex combination of $d+1$ extremal points, where d is the dimension of the convex set.

Theorem B.3 (Carathéodory). *Any point of a convex body $\mathcal{K} \subset \mathbb{R}^d$ of affine dimension d can be written as a convex combination of $d+1$ extreme points.*

Finally,

Definition B.11. A subset $F \subset \mathcal{K}$ is called a *face* of $\mathcal{K}$, denoted $F \triangleleft \mathcal{K}$, if

$$\forall \boldsymbol{y}, \boldsymbol{z} \in \mathcal{K},\ \lambda \in (0,1),\ \lambda \boldsymbol{y} + (1-\lambda)\boldsymbol{z} \in F \implies \boldsymbol{y}, \boldsymbol{z} \in F.$$

The *face dimension* d_F of a face F is the affine dimension of F.

Every face is itself a convex set. Intuitively, a face is a part of the boundary of $\mathcal{K}$ such that if any line segment lies on this part of the boundary, its endpoints must also be on this part of the boundary. Extremal points correspond to the faces of dimension zero.

B.3 Cones

Cones arise naturally in convex optimization, where many constraint sets—such as the set of positive semidefinite matrices or the nonnegative orthant—are modeled as convex cones. Here, we provide just the basic definitions.

Definition B.12. A set $\mathcal{C} \subseteq \mathbb{R}^n$ is called a *cone*, or *nonnegative homogeneous*, if for every $\boldsymbol{x} \in \mathcal{C}$ and $\lambda \geq 0$, we have $\lambda \boldsymbol{x} \in \mathcal{C}$.

Definition B.13. A set $\mathcal{C}$ is a *convex cone* if it is convex and a cone, which means that for any $\boldsymbol{x}_1,\ \boldsymbol{x}_2 \in \mathcal{C}$ and $\lambda_1, \lambda_2 \geq 0$, we have

$$\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 \in \mathcal{C}. \tag{B.10}$$

Points of this form can be described geometrically as forming the two-dimensional pie slice with apex 0 and edges passing through $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$.

A point of the form $\sum_i \lambda_i \boldsymbol{x}_i$ with $\lambda_i \geq 0$ is called a *conic combination* (or a nonnegative linear combination) of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$. If $\boldsymbol{x}_i$ are in a convex cone $\mathcal{C}$, then every conic combination of $\boldsymbol{x}_i$ is in $\mathcal{C}$. Conversely, a set $\mathcal{C}$ is a convex cone if and only if it contains all conic combinations of its elements.

B.4 Convex duality

Convex duality provides a powerful perspective on the geometry of the quantum set. It plays a central role in this thesis, where we use it to study supporting hyperplanes and orthogonal faces of the quantum set.

Definition B.14. Given a convex set $\mathcal{K} \subset \mathbb{R}^n$, we define its *dual set* as:

$$\mathcal{K}^* \coloneqq \{\boldsymbol{f} \in \mathbb{R}^n \colon \forall \boldsymbol{x} \in \mathcal{K}, \boldsymbol{f} \cdot \boldsymbol{x} \leq 1\}. \tag{B.11}$$

This set captures the linear functionals (vectors $\boldsymbol{f}$) that are bounded by 1 on $\mathcal{K}$. The dual set is always convex and reflects the linear constraints satisfied by elements of $\mathcal{K}$.

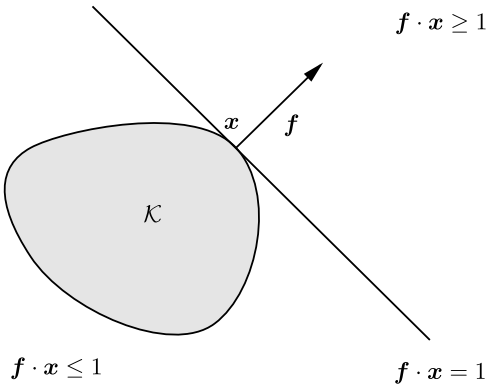

Figure B.4: Illustration of a supporting hyperplane to the convex set $\mathcal{K}$ at the boundary point $\boldsymbol{x}$.

Remark B.3. The relationship between supporting hyperplanes and the dual set is fundamental: if H is a supporting hyperplane for $\mathcal{K}$ with normal vector $\boldsymbol{f}$, and we normalize $\boldsymbol{f}$ such that $\sup_{\boldsymbol{y}\in\mathcal{K}} \boldsymbol{f}\cdot\boldsymbol{y} = 1$, then this $\boldsymbol{f}$ is an element of $\mathcal{K}^*$. Conversely, any $\boldsymbol{f} \in \mathcal{K}^*$ with $\sup_{\boldsymbol{x}\in\mathcal{K}} \boldsymbol{f}\cdot\boldsymbol{x} = 1$ defines a supporting hyperplane. Thus, the boundary of $\mathcal{K}^*$ is defined by the normal vectors of the supporting hyperplanes of $\mathcal{K}$.

Definition B.15. To any face $F \triangleleft \mathcal{K}$, we associate an *orthogonal face* in the dual set $\mathcal{K}^*$, defined by

$$F^\perp \coloneqq \{\boldsymbol{f} \in \mathcal{K}^* \colon \forall \boldsymbol{x} \in F,\ \boldsymbol{f}\cdot\boldsymbol{x} = 1\}\,. \tag{B.12}$$

The *orthogonal dimension* $d_F^\perp$ is the affine dimension of $F^\perp$, often denoted d_F^*.

This face captures the set of linear functionals that are saturated by every point in F. In the context of quantum correlations, if F represents a set of observed correlations, $F^\perp$ represents all Bell expressions that achieve their maximum possible value for all correlations in F.

Proposition B.1. *The set $F^\perp$ is a face of the convex dual $\mathcal{K}^*$.*

Proof. Let $F \triangleleft \mathcal{K}$. Let $\boldsymbol{g}, \boldsymbol{h} \in \mathcal{K}^*$, $\lambda \in (0,1)$ such that $\lambda\boldsymbol{g} + (1-\lambda)\boldsymbol{h} \in F^\perp$. For all $\boldsymbol{x} \in F$, $(\lambda\boldsymbol{g} + (1-\lambda)\boldsymbol{h})\cdot\boldsymbol{x} = 1$. Since $\boldsymbol{g}, \boldsymbol{h}$ are elements of the convex dual and $\boldsymbol{x} \in \mathcal{K}$, their scalar product with $\boldsymbol{x}$ is upper bounded by 1. Thus $\lambda\boldsymbol{g}\cdot\boldsymbol{x} + (1-\lambda)\boldsymbol{h}\cdot\boldsymbol{x} \leq 1$. The equality implies that $\boldsymbol{g}\cdot\boldsymbol{x} = \boldsymbol{h}\cdot\boldsymbol{x} = 1$ and thus $\boldsymbol{g}, \boldsymbol{h} \in F^\perp$. □

Definition B.16. We call *dimension pair* of F the couple $(d_F, d_F^\perp)$.

Definition B.17. We say that a face $F \triangleleft \mathcal{K}$ is *exposed* if there exists $\boldsymbol{f} \in \mathcal{K}^*$ such that

$$\boldsymbol{f}\cdot\boldsymbol{x} = 1 \implies \boldsymbol{x} \in F \quad \forall \boldsymbol{x} \in \mathcal{K}. \tag{B.13}$$

An exposed face is one that can be defined as the set of points in $\mathcal{K}$ that maximize a specific linear functional. In this case, one can show that $d_F = d_{F^\perp}^\perp$. This relationship, often referred to as the *face-polar duality*, provides a powerful link between the geometry of a face and the

geometry of its dual counterpart. The dimension pair $(d_F, d_F^\perp)$ thus provides a useful invariant that characterizes the geometry of exposed faces.

We conclude with one more definition:

Definition B.18. Let $\mathcal{K}, \mathcal{K}' \subset \mathbb{R}^n$ be two convex sets. We say that they are *isomorphic* if there exists a linear isomorphism

$$U \colon \mathbb{R}^n \to \mathbb{R}^n$$

such that

$$U(\mathcal{K}) = \mathcal{K}'. \tag{B.14}$$

This notion preserves all the relevant convex-geometric properties: if $F \triangleleft \mathcal{K}$, then $U(F) \triangleleft \mathcal{K}'$ is a face with the same dimension pair. Moreover, $\mathcal{K}^*$ and $\mathcal{K}'^*$ are isomorphic dual sets, specifically, $(U(\mathcal{K}))^* = (U^T)^{-1}(\mathcal{K}^*)$, where U^T is the transpose of U. This implies that the dual structure is also preserved under isomorphism.

B.5 Polytopes

Polytopes play a central role in convex geometry and combinatorics. In the context of Bell nonlocality, local and no-signaling sets are described by convex polytopes, and understanding their structure is crucial for analyzing classical and general correlations. For a detailed treatment of the theory of polytopes, we refer the reader to [138].

Definition B.19. A *polytope*[1] is a subset of $\mathbb{R}^n$ that can be written as the convex hull of a finite set of points.

For polytopes, the hierarchy introduced in (B.8) collapses: all extremal points are exposed and correspond to the *vertices* of the polytope.

A polytope is said to be in *V-representation* if it is specified by the list of its vertices. By definition, every polytope admits such a representation. Geometric intuition suggests that convex polytopes can also be described in terms of their bounding hyperplanes or *facets*. This alternative description is known as the *H-representation*, where the polytope is expressed as the intersection of finitely many closed half-spaces.

The equivalence of these two representations is guaranteed by the following fundamental result:

Theorem B.4 (Weyl-Minkowski). *Every convex polytope $P \subseteq \mathbb{R}^n$ admits both a V-representation and an H-representation. Moreover, any bounded intersection of finitely many half-spaces in $\mathbb{R}^n$ is a polytope.*

The V- and H-representations are dual in nature and suited to different types of analysis. A natural computational question then arises: given one representation, how can we obtain the other? This is known respectively as the *vertex enumeration problem* (from H to V) and the *facet enumeration problem* (from V to H).

We conclude this section with a geometric stability result about polytopes under linear maps:

[1] Since we restrict our attention to convex sets, we use the terms polytope and convex polytope interchangeably.

Theorem B.5 (Projection of a polytope). *Let $P \subset \mathbb{R}^n$ be a polytope, and let $\pi \colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear projection. Then $\pi(P)$ is a polytope in $\mathbb{R}^m$.*

This fact is highly relevant in this thesis, since we analyze the local set as the projection of a higher-dimensional convex body.

Appendix C

Elements of convex optimization

This appendix summarizes key concepts and tools from convex optimization essential for the thesis. It covers problem formulations, duality theory, and semidefinite programming [126, 139], providing the mathematical foundation for analyzing quantum correlations and related optimization tasks [140, 141].

C.1 Convex optimization

Definition C.1. A *convex program* is an optimization problem of the form:

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & f_0(\boldsymbol{x}) \\ \text{s. t.} \quad & f_i(\boldsymbol{x}) \leq 0, \quad i = 1, \ldots, m, \\ & h_i(\boldsymbol{x}) = 0, \quad i = 1, \ldots, p, \end{aligned} \tag{C.1}$$

where $f_0, f_1, \ldots, f_m \colon \mathbb{R}^n \to \mathbb{R}$ are convex functions and $h_1, \ldots, h_p \colon \mathbb{R}^n \to \mathbb{R}$ are affine functions.

A point $\boldsymbol{x} \in \mathcal{D}$, where $\mathcal{D}$ is the intersection of the domains of all functions involved, is called *feasible* if it satisfies the constraints. If no such point exists, the program is *infeasible*, and by convention one sets the optimal value $p^* = +\infty$.

To analyze such problems, one typically introduces the *Lagrangian*:

$$L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \coloneqq f_0(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i f_i(\boldsymbol{x}) + \sum_{i=1}^{p} \nu_i h_i(\boldsymbol{x}), \tag{C.2}$$

where $\boldsymbol{\lambda} \in \mathbb{R}_+^m$ and $\boldsymbol{\nu} \in \mathbb{R}^p$ are the *Lagrange multipliers*.

Definition C.2. Given a convex optimization problem with Lagrangian $L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu})$, the *Lagrange dual function* is defined as

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \coloneqq \inf_{\boldsymbol{x}\in\mathcal{D}} L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}),$$

where $\mathcal{D}$ is the domain of the primal variables.

The dual function g is always concave, even if the primal problem is not convex, because it is the pointwise infimum of a family of affine functions in $(\boldsymbol{\lambda}, \boldsymbol{\nu})$. Moreover, it provides a lower bound on the optimal value p^* of the primal problem:

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq p^* \quad \text{for all } \boldsymbol{\lambda} \succeq 0.$$

Definition C.3. The *Lagrange dual program* associated with the primal problem (C.1) is given by:

$$d^* \coloneqq \sup_{\boldsymbol{\lambda} \in \mathbb{R}^m_+, \boldsymbol{\nu} \in \mathbb{R}^p} g(\boldsymbol{\lambda}, \boldsymbol{\nu}). \tag{C.3}$$

This is a convex optimization problem, since it involves the maximization of a concave function.

It always holds that $d^* \leq p^*$ (*weak duality*). When equality holds, i.e., $d^* = p^*$, the problem exhibits *strong duality*. Otherwise, the quantity $p^* - d^*$ is referred to as the *duality gap*.

If both the primal and dual problems are feasible, then for any feasible primal point $\boldsymbol{x}_f$ and any feasible dual point $\boldsymbol{\xi}_f \coloneqq (\boldsymbol{\lambda}, \boldsymbol{\nu})_f$, the following chain of inequalities holds:

$$g(\boldsymbol{\xi}_f) \leq d^* \leq p^* \leq f_0(\boldsymbol{x}_f). \tag{C.4}$$

If strong duality holds, then the bounds coincide up to numerical precision.

C.1.1 Linear programs

Definition C.4. A *linear program (LP)* is a convex program in which all functions f_0, f_i are affine. In standard form, it is written as: [1]

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathbb{R}^n} \quad & \boldsymbol{c}^T \boldsymbol{x} \\ \text{s. t.} \quad & \boldsymbol{x} \succeq 0, \\ & A\boldsymbol{x} = \boldsymbol{b}, \end{aligned} \tag{C.5}$$

where $\boldsymbol{c} \in \mathbb{R}^n$, $A \in \mathbb{R}^{p,n}$, and $\boldsymbol{b} \in \mathbb{R}^p$.

An important example is given by optimization over a probability vector: the constraints $\boldsymbol{x} \succeq 0$ ensure non-negativity of the probabilities, and $A\boldsymbol{x} = \boldsymbol{b}$ may include a normalization condition.

Let us derive the dual of the standard form LP. The Lagrangian reads:

$$\begin{aligned} L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) &\coloneqq \boldsymbol{c}^T \boldsymbol{x} - \boldsymbol{\lambda}^T \boldsymbol{x} + \boldsymbol{\nu}^T (A\boldsymbol{x} - \boldsymbol{b}) \\ &= -\boldsymbol{\nu}^T \boldsymbol{b} + (\boldsymbol{c} - \boldsymbol{\lambda} + A^T \boldsymbol{\nu})^T \boldsymbol{x}. \end{aligned} \tag{C.6}$$

The dual function is finite only if the coefficient of $\boldsymbol{x}$ vanishes:

$$\boldsymbol{c} - \boldsymbol{\lambda} + A^T \boldsymbol{\nu} = 0 \quad \Rightarrow \quad \boldsymbol{\lambda} = A^T \boldsymbol{\nu} + \boldsymbol{c}.$$

Since $\boldsymbol{\lambda} \succeq 0$, this gives the dual constraint $A^T \boldsymbol{\nu} + \boldsymbol{c} \succeq 0$. Under this condition, the infimum of the Lagrangian is simply $-\boldsymbol{b}^T \boldsymbol{\nu}$, leading to the dual problem

$$\begin{aligned} \max_{\boldsymbol{\nu} \in \mathbb{R}^p} \quad & -\boldsymbol{b}^T \boldsymbol{\nu} \\ \text{s. t.} \quad & A^T \boldsymbol{\nu} + \boldsymbol{c} \succeq 0. \end{aligned} \tag{C.7}$$

[1] Throughout this thesis, we use the symbol $\succeq$ to denote different types of orderings, depending on context: for real vectors, $\boldsymbol{x} \succeq \boldsymbol{y}$ means $x_i \geq y_i$ for all components i; for Hermitian matrices or operators, $A \succeq B$ means that $A - B$ is positive semidefinite.

C.1.2 Semidefinite programs

Definition C.5. A *semidefinite program (SDP)* is a convex optimization problem of the form:

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s. t.} \quad & \sum_{i=1}^{n} x_i F_i \preceq 0, \\ & A\boldsymbol{x} = \boldsymbol{b}, \end{aligned} \tag{C.8}$$

where $F_i \in \mathbb{R}^{k\times k}$ are symmetric matrices.

Note that, for $k = 1$, the problem reduces to a linear program.

To derive the dual of a semidefinite program, we introduce a matrix-valued Lagrange multiplier $Z \succeq 0$ for the matrix inequality constraint, and a vector of multipliers $\boldsymbol{\nu}$ for the equality constraint. The Lagrangian becomes:

$$\begin{aligned} L(\boldsymbol{x}, Z, \boldsymbol{\nu}) &\coloneqq \boldsymbol{c}^T\boldsymbol{x} + \operatorname{Tr}\left(Z\sum_{i=1}^{n} x_i F_i\right) + \boldsymbol{\nu}^T(A\boldsymbol{x} - \boldsymbol{b}) \\ &= -\boldsymbol{\nu}^T\boldsymbol{b} + \sum_{i=1}^{n}\left(c_i + \operatorname{Tr}(ZF_i) + (A^T\boldsymbol{\nu})_i\right)x_i. \end{aligned} \tag{C.9}$$

The dual function is finite only if the coefficient of each x_i vanishes. This yields the stationarity condition:

$$c_i + \operatorname{Tr}(ZF_i) + (A^T\boldsymbol{\nu})i = 0 \quad \text{for all } i.$$

Under this condition, the infimum of the Lagrangian is simply $-\boldsymbol{b}^T\boldsymbol{\nu}$, leading to the dual problem

$$\begin{aligned} \max_{\boldsymbol{\nu},Z} \quad & -\boldsymbol{b}^T\boldsymbol{\nu} \\ \text{s. t.} \quad & c_i + \operatorname{Tr}(ZF_i) + (A^T\boldsymbol{\nu})_i = 0, \quad \forall i, \\ & Z \succeq 0. \end{aligned} \tag{C.10}$$

Semidefinite programs (SDPs) can be efficiently implemented using optimization solvers and modeling frameworks such as MOSEK [124], YALMIP [142] (with MATLAB), and CVXPY [143] (with Python). For problems involving noncommutative polynomials, specialized tools like `ncpol2sdpa` [123] generate SDP relaxations compatible with MOSEK.

References

[1] G. Auletta and G. Parisi, *Foundations and Interpretation of Quantum Mechanics: In the Light of a Critical-historical Analysis of the Problems and of a Synthesis of the Results*, World Scientific, 2001.

[2] H. Nikolić, Quantum mechanics: myths and facts, in: *Foundations of Physics* 37.11 (Sept. 2007), pp. 1563–1611, DOI: 10.1007/s10701-007-9176-y.

[3] F. Laloë, *Do We Really Understand Quantum Mechanics?*, 2nd ed., Cambridge University Press, 2019.

[4] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete?, in: *Physical Review* 47.10 (May 1935), pp. 777–780, DOI: 10.1103/physrev.47.777.

[5] D. Bohm, *Quantum Mechanics*, Dover Publications, 1951.

[6] D. Bohm, A suggested interpretation of the quantum theory in terms of "hidden" variables. I, in: *Physical Review* 85 (2 Jan. 1952), pp. 166–179, DOI: 10.1103/PhysRev.85.166.

[7] T. Norsen, *Foundations of Quantum Mechanics: An Exploration of the Physical Meaning of Quantum Theory*, Springer International Publishing, 2017.

[8] G. Auletta, M. Fortunato, and G. Parisi, *Quantum Mechanics*, Cambridge University Press, 2009.

[9] J. S. Bell, On the Einstein-Podolsky-Rosen paradox, in: *Physics* 1 (3 Nov. 1964), pp. 195–200, DOI: 10.1103/PhysicsPhysiqueFizika.1.195.

[10] J. F. Clauser and A. Shimony, Bell's theorem. Experimental tests and implications, in: *Reports on Progress in Physics* 41.12 (Dec. 1978), pp. 1881–1927, DOI: 10.1088/0034-4885/41/12/002.

[11] A. Aspect, P. Grangier, and G. Roger, Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: a new violation of Bell's inequalities, in: *Physical Review Letters* 49 (2 July 1982), pp. 91–94, DOI: 10.1103/PhysRevLett.49.91.

[12] A. Aspect, J. Dalibard, and G. Roger, Experimental test of Bell's inequalities using time-varying analyzers, in: *Physical Review Letters* 49 (25 Dec. 1982), pp. 1804–1807, DOI: 10.1103/PhysRevLett.49.1804.

[13] M. A. Rowe et al., Experimental violation of a Bell's inequality with efficient detection, in: *Nature* 409.6822 (Feb. 2001), pp. 791–794, DOI: 10.1038/35057215.

[14] B. Hensen et al., Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres, in: *Nature* 526.7575 (Oct. 2015), pp. 682–686, DOI: 10.1038/nature15759.

[15] L. K. Shalm et al., Strong loophole-free test of local realism, in: *Physical Review Letters* 115 (25 Dec. 2015), p. 250402, DOI: 10.1103/PhysRevLett.115.250402.

[16] T. Norsen, Local causality and completeness: Bell vs. Jarrett, in: *Foundations of Physics* 39.3 (Feb. 2009), pp. 273–294, DOI: 10.1007/s10701-009-9281-1.

[17] S. Goldstein et al., Bell's theorem, in: *Scholarpedia* 6.10 (2011), p. 8378, DOI: 10.4249/scholarpedia.8378.

[18] T. Maudlin, *Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics*, Wiley-Blackwell, 2011.

[19] N. Brunner et al., Bell nonlocality, in: *Reviews of Modern Physics* 86 (2 June 2014), pp. 839–840, DOI: 10.1103/RevModPhys.86.839.

[20] V. Scarani, *Bell Nonlocality*, Oxford University Press, 2019.

[21] B. S. Cirel'son, Quantum generalizations of Bell's inequality, in: *Letters in Mathematical Physics* 4.2 (Mar. 1980), pp. 93–100, DOI: 10.1007/bf00417500.

[22] L. Hardy, *Quantum theory from five reasonable axioms*, 2001, arXiv: quant-ph/0101012 (quant-ph).

[23] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Informational derivation of quantum theory, in: *Physical Review A* 84 (1 July 2011), p. 012311, DOI: 10.1103/PhysRevA.84.012311.

[24] L. Masanes and M. P. Müller, A derivation of quantum theory from physical requirements, in: *New Journal of Physics* 13.6 (June 2011), p. 063001, DOI: 10.1088/1367-2630/13/6/063001.

[25] S. Popescu, Nonlocality beyond quantum mechanics, in: *Nature Physics* 10.4 (Apr. 2014), pp. 264–270, DOI: 10.1038/nphys2916.

[26] S. Aaronson, *Is quantum mechanics an island in theoryspace?*, 2004, arXiv: quant-ph/0401062 (quant-ph).

[27] M. Müller, Probabilistic theories and reconstructions of quantum theory, in: *SciPost Physics Lecture Notes* (Mar. 2021), DOI: 10.21468/scipostphyslectnotes.28.

[28] M. Pawłowski et al., Information causality as a physical principle, en, in: *Nature* 461.7267 (Oct. 2009), pp. 1101–1104, DOI: 10.1038/nature08400.

[29] M. Navascués and H. Wunderlich, A glance beyond the quantum model, in: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 466.2115 (Nov. 2009), pp. 881–890, DOI: 10.1098/rspa.2009.0453.

[30] T. Fritz et al., Local orthogonality as a multipartite principle for quantum correlations, in: *Nature Communications* 4.1 (Aug. 2013), DOI: 10.1038/ncomms3263.

[31] V. Scarani, The device-independent outlook on quantum physics, in: *Acta Physica Slovaca* 62.4 (2012), pp. 347–409, DOI: 10.2478/v10155-012-0003-4.

[32] S. Pironio, V. Scarani, and T. Vidick, Focus on device independent quantum information, in: *New Journal of Physics* 18.10 (Oct. 2016), p. 100202, DOI: 10.1088/1367-2630/18/10/100202.

[33] A. K. Ekert, Quantum cryptography based on Bell's theorem, in: *Physical Review Letters* 67 (6 Aug. 1991), pp. 661–663, DOI: 10.1103/PhysRevLett.67.661.

[34] A. Acín, N. Gisin, and L. Masanes, From Bell's theorem to secure quantum key distribution, in: *Physical Review Letters* 97 (12 Sept. 2006), p. 120405, DOI: 10.1103/PhysRevLett.97.120405.

[35] D. P. Nadlinger et al., Experimental quantum key distribution certified by Bell's theorem, in: *Nature* 607.7920 (July 2022), pp. 682–686, DOI: 10.1038/s41586-022-04941-5.

[36] S. Pironio et al., Random numbers certified by Bell's theorem, in: *Nature* 464.7291 (Apr. 2010), pp. 1021–1024, DOI: 10.1038/nature09008.

[37] I. Šupić and J. Bowles, Self-testing of quantum systems: a review, in: *Quantum* 4 (Sept. 2020), p. 337, DOI: 10.22331/q-2020-09-30-337.

[38] R. Jozsa and N. Linden, On the role of entanglement in quantum-computational speed-up, in: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 459.2036 (Aug. 2003), pp. 2011–2032, DOI: 10.1098/rspa.2002.1097.

[39] J. Anders and D. E. Browne, Computational power of correlations, in: *Physical Review Letters* 102 (5 Feb. 2009), p. 050502, DOI: 10.1103/PhysRevLett.102.050502.

[40] A. Tavakoli et al., Bell nonlocality in networks, in: *Reports on Progress in Physics* 85.5 (Mar. 2022), p. 056001, DOI: 10.1088/1361-6633/ac41bb.

[41] H. J. Kimble, The quantum internet, in: *Nature* 453.7198 (June 2008), pp. 1023–1030, DOI: 10.1038/nature07127.

[42] S. Wehner, D. Elkouss, and R. Hanson, Quantum internet: A vision for the road ahead, in: *Science* 362.6412 (Oct. 2018), DOI: 10.1126/science.aam9288.

[43] P. P. Rohde, *The Quantum Internet: The Second Quantum Revolution*, Cambridge University Press, Sept. 2021.

[44] L. J. Landau, Empirical two-point correlation functions, in: *Foundations of Physics* 18.4 (Apr. 1988), pp. 449–460, DOI: 10.1007/bf00732549.

[45] L. Masanes, *Necessary and sufficient condition for quantum-generated correlations*, 2003, arXiv: quant-ph/0309137 (quant-ph).

[46] K. T. Goh et al., Geometry of the set of quantum correlations, in: *Physical Review A* 97 (2 Feb. 2018), p. 022104, DOI: 10.1103/PhysRevA.97.022104.

[47] K.-S. Chen et al., Quantum correlations on the no-signaling boundary: self-testing and more, in: *Quantum* 7 (July 2023), p. 1054, DOI: 10.22331/q-2023-07-11-1054.

[48] T. P. Le et al., Quantum correlations in the minimal scenario, in: *Quantum* 7 (Mar. 2023), p. 947, DOI: 10.22331/q-2023-03-16-947.

[49] M. Navascués, S. Pironio, and A. Acín, Bounding the set of quantum correlations, in: *Physical Review Letters* 98.1 (Jan. 2007), p. 010401, DOI: 10.1103/PhysRevLett.98.010401.

[50] M. Navascués, S. Pironio, and A. Acín, A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations, in: *New Journal of Physics* 10.7 (July 2008), p. 073013, DOI: 10.1088/1367-2630/10/7/073013.

[51] V. Barizien and J.-D. Bancal, Quantum statistics in the minimal Bell scenario, in: *Nature Physics* (Mar. 2025), DOI: 10.1038/s41567-025-02782-3.

[52] V. Barizien and J.-D. Bancal, Extremal Tsirelson inequalities, in: *Physical Review Letters* 133 (1 July 2024), p. 010201, DOI: 10.1103/PhysRevLett.133.010201.

[53] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, in: *Foundations of Physics* 24.3 (Mar. 1994), pp. 379–385, DOI: 10.1007/bf02058098.

[54] J. S. Bell, "La nouvelle cuisine", in: *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, 2010, pp. 232–248.

[55] J. S. Bell, The theory of local beables, in: *Epistemological Letters* 9 (1975).

[56] A. Fine, Hidden variables, joint probability, and the Bell inequalities, in: *Physical Review Letters* 48.5 (Feb. 1982), pp. 291–295, DOI: 10.1103/PhysRevLett.48.291.

[57] B. Tsirelson, Some results and problems on quantum Bell-type inequalities, in: *Hadronic Journal. Supplement* 8.4 (1993), pp. 329–345, DOI: 10.1016/0370-1573(93)90050-8.

[58] M. Navascués et al., A Physical Approach to Tsirelson's Problem, in: *Foundations of Physics* 42.8 (Mar. 2012), pp. 985–995, DOI: 10.1007/s10701-012-9641-0.

[59] A. Coladangelo and J. Stark, *Unconditional separation of finite and infinite-dimensional quantum correlations*, 2018, arXiv: 1804.05116 (quant-ph).

[60] W. Slofstra, Tsirelson's problem and an embedding theorem for groups arising from non-local games, in: *Journal of the American Mathematical Society* 33.1 (Sept. 2019), pp. 1–56, DOI: 10.1090/jams/929.

[61] Z. Ji et al., MIP* = RE, in: *Communication ACM* 64.11 (Oct. 2021), pp. 131–138, DOI: 10.1145/3485628.

[62] I. Pitowsky, The range of quantum probability, in: *Journal of Mathematical Physics* 27.6 (June 1986), pp. 1556–1565, DOI: 10.1063/1.527066.

[63] S. Pironio, Lifting Bell inequalities, in: *Journal of Mathematical Physics* 46.6 (June 2005), p. 062112, DOI: 10.1063/1.1928727.

[64] B. Simon, *Convexity: An Analytic Viewpoint*, Cambridge Tracts in Mathematics, Cambridge University Press, 2011.

[65] J. F. Clauser et al., Proposed experiment to test local hidden-variable theories, in: *Physical Review Letters* 23 (15 Oct. 1969), pp. 880–884, DOI: 10.1103/PhysRevLett.23.880.

[66] N. Gisin, Bell's inequality holds for all non-product states, in: *Physics Letters A* 154.5 (1991), pp. 201–202, DOI: 10.1016/0375-9601(91)90805-I.

[67] N. Linden et al., Quantum nonlocality and beyond: Limits from nonlocal computation, in: *Physical Review Letters* 99.18 (Oct. 2007), DOI: 10.1103/physrevlett.99.180502.

[68] M. L. Almeida et al., Guess your neighbor's input: a multipartite nonlocal game with no quantum advantage, in: *Physical Review Letters* 104 (23 June 2010), p. 230404, DOI: 10.1103/PhysRevLett.104.230404.

[69] S. Wehner, Tsirelson bounds for generalized Clauser-Horne-Shimony-Holt inequalities, in: *Physical Review A* 73 (2 Feb. 2006), p. 022110, DOI: 10.1103/PhysRevA.73.022110.

[70] A. C. Doherty et al., "The quantum moment problem and bounds on entangled multi-prover games", in: *Proceedings of the 2008 IEEE 23rd Annual Conference on Computational Complexity*, CCC '08, USA: IEEE Computer Society, 2008, pp. 199–210, DOI: 10.1109/CCC.2008.26.

[71] P. A. Parrilo, Semidefinite programming relaxations for semialgebraic problems, in: *Mathematical Programming* 96.2 (May 2003), pp. 293–320, DOI: 10.1007/s10107-003-0387-5.

[72] A. Rai et al., Geometry of the quantum set on no-signaling faces, in: *Physical Review A* 99 (3 Mar. 2019), p. 032106, DOI: 10.1103/PhysRevA.99.032106.

[73] A. Montanaro and R. de Wolf, in: *Theory of Computing* 1.1 (2016), pp. 1–81, DOI: 10.4086/toc.gs.2016.007.

[74] J.-D. Bancal et al., Device-independent witnesses of genuine multipartite entanglement, in: *Physical Review Letters* 106 (25 June 2011), p. 250404, DOI: 10.1103/PhysRevLett.106.250404.

[75] T. Moroder et al., Device-independent entanglement quantification and related applications, in: *Physical Review Letters* 111 (3 July 2013), p. 030501, DOI: 10.1103/PhysRevLett.111.030501.

[76] D. Mayers and A. Yao, Self-testing quantum apparatus, in: 4.4 (July 2004), pp. 273–286.

[77] F. Baccari et al., Device-independent certification of genuinely entangled subspaces, in: *Physical Review Letters* 125 (26 Dec. 2020), p. 260507, DOI: 10.1103/PhysRevLett.125.260507.

[78] J.-D. Bancal, N. Sangouard, and P. Sekatski, Noise-resistant device-independent certification of Bell state measurements, in: *Physical Review Letters* 121 (25 Dec. 2018), p. 250506, DOI: 10.1103/PhysRevLett.121.250506.

[79] P. Sekatski et al., Certifying the building blocks of quantum computers from Bell's theorem, in: *Physical Review Letters* 121.18 (Nov. 2018), p. 180505, DOI: 10.1103/PhysRevLett.121.180505.

[80] R. Colbeck, "Quantum and Relativistic Protocols for Secure Multi-Party Computation", PhD thesis, University of Cambridge, 2009, arXiv: 0911.3814 (quant-ph).

[81] A. Acín et al., Device-independent security of quantum cryptography against collective attacks, in: *Physical Review Letters* 98 (23 June 2007), p. 230501, DOI: 10.1103/PhysRevLett.98.230501.

[82] E.-J. Kuo and L.-Y. Hsu, *Self-testing quantum error correcting codes: analyzing computational hardness*, 2024, arXiv: 2409.01987 (quant-ph).

[83] Y. Wang, X. Wu, and V. Scarani, All the self-testings of the singlet for two binary measurements, in: *New Journal of Physics* 18.2 (Feb. 2016), p. 025021, DOI: 10.1088/1367-2630/18/2/025021.

[84] A. Coladangelo, K. T. Goh, and V. Scarani, All pure bipartite entangled states can be self-tested, in: *Nature Communications* 8.1 (May 2017), p. 15485, DOI: 10.1038/ncomms15485.

[85] M. Balanzó-Juandó et al., *All pure multipartite entangled states of qubits can be self-tested up to complex conjugation*, 2024, arXiv: 2412.13266 (quant-ph).

[86] S. J. Summers and R. Werner, Maximal violation of Bell's inequalities is generic in quantum field theory, in: *Communications in Mathematical Physics* 110.2 (June 1987), pp. 247–259, DOI: 10.1007/bf01207366.

[87] S. Popescu and D. Rohrlich, Which states violate Bell's inequality maximally?, in: *Physics Letters A* 169.6 (1992), pp. 411–414, DOI: 10.1016/0375-9601(92)90819-8.

[88] S. L. Braunstein, A. Mann, and M. Revzen, Maximal violation of Bell inequalities for mixed states, in: *Physical Review Letters* 68 (22 June 1992), pp. 3259–3261, DOI: 10.1103/PhysRevLett.68.3259.

[89] G. C. Ghirardi, A. Rimini, and T. Weber, A general argument against superluminal transmission through the quantum mechanical measurement process, in: *Lettere al Nuovo Cimento* 27.10 (Mar. 1980), pp. 293–298, DOI: 10.1007/bf02817189.

[90] R. Tumulka, *Foundations of Quantum Mechanics*, Lecture Notes in Physics, Springer International Publishing, 2022, DOI: 10.1007/978-3-031-09548-1.

[91] A. Shimony, *The Search for a Naturalistic World View*, Cambridge University Press, 1993, DOI: 10.1017/cbo9780511621147.

[92] L. E. Ballentine and J. P. Jarrett, Bell's theorem: does quantum mechanics contradict relativity?, in: *American Journal of Physics* 55.8 (Aug. 1987), pp. 696–701, DOI: 10.1119/1.15059.

[93] P. H. Eberhard and R. R. Ross, Quantum field theory cannot provide faster-than-light communication, in: *Foundations of Physics Letters* 2.2 (Mar. 1989), pp. 127–149, DOI: 10.1007/bf00696109.

[94] A. Peres and D. R. Terno, Quantum information and relativity theory, in: *Reviews of Modern Physics* 76 (1 Jan. 2004), pp. 93–123, DOI: 10.1103/RevModPhys.76.93.

[95] L. Vervoort, Bell's theorem: two neglected solutions, in: *Foundations of Physics* 43.6 (Apr. 2013), pp. 769–791, DOI: 10.1007/s10701-013-9715-7.

[96] R. D. Gill, Statistics, causality and Bell's theorem, in: *Statistical Science* 29.4 (Nov. 2014), DOI: 10.1214/14-sts490.

[97] Č. Brukner, Quantum causality, in: *Nature Physics* 10.4 (Apr. 2014), pp. 259–263, DOI: 10.1038/nphys2930.

[98] M. Żukowski and Č. Brukner, Quantum non-locality—it ain't necessarily so... In: *Journal of Physics A: Mathematical and Theoretical* 47.42 (Oct. 2014), p. 424009, DOI: 10.1088/1751-8113/47/42/424009.

[99] A. Khrennikov, Get rid of nonlocality from quantum physics, in: *Entropy* 21.8 (2019), DOI: 10.3390/e21080806.

[100] J. R. Hance and S. Hossenfelder, Bell's theorem allows local theories of quantum mechanics, in: *Nature Physics* 18.12 (Oct. 2022), pp. 1382–1382, DOI: 10.1038/s41567-022-01831-5.

[101] S. Hossenfelder, *Quantum confusions, cleared up (or so I hope)*, 2024, arXiv: 2309.12299 (quant-ph).

[102] W. Myrvold, M. Genovese, and A. Shimony, "Bell's Theorem", in: *The Stanford Encyclopedia of Philosophy*, ed. by E. N. Zalta and U. Nodelman, Spring 2024, Metaphysics Research Lab, Stanford University, 2024.

[103] G. M. D'Ariano, G. Chiribella, and P. Perinotti, *Quantum Theory from First Principles: An Informational Approach*, Cambridge University Press, 2017.

[104] J. Barrett et al., The computational landscape of general physical theories, in: *npj Quantum Information* 5.1 (May 2019), p. 41, DOI: 10.1038/s41534-019-0156-9.

[105] M. Weilenmann and R. Colbeck, Self-testing of physical theories, or, Is quantum theory optimal with respect to some information-processing task?, in: *Physical Review Letters* 125.6 (Aug. 2020), p. 060406, DOI: 10.1103/PhysRevLett.125.060406.

[106] M. Plávala, General probabilistic theories: An introduction, in: *Physics Reports* 1033 (2023), pp. 1–64, DOI: 10.1016/j.physrep.2023.09.001.

[107] J. Barrett, Information processing in generalized probabilistic theories, in: *Physical Review A* 75 (3 Mar. 2007), p. 032304, DOI: 10.1103/PhysRevA.75.032304.

[108] W. K. Wootters and W. H. Zurek, A single quantum cannot be cloned, in: *Nature* 299.5886 (Oct. 1982), pp. 802–803, DOI: 10.1038/299802a0.

[109] N. Gisin, Quantum cloning without signaling, in: *Physics Letters A* 242.1 (1998), pp. 1–3, DOI: 10.1016/S0375-9601(98)00170-4.

[110] H. Barnum et al., Generalized no-broadcasting theorem, in: *Physical Review Letters* 99 (24 Dec. 2007), p. 240501, DOI: 10.1103/PhysRevLett.99.240501.

[111] W. van Dam, *Implausible Consequences of Superstrong Nonlocality*, 2005, arXiv: quant-ph/0501159 (quant-ph).

[112] T. Fritz, Polyhedral duality in Bell scenarios with two binary observables, in: *Journal of Mathematical Physics* 53.7 (July 2012), p. 072202, DOI: 10.1063/1.4734586.

[113] B. S. Tsirel'son, Quantum analogues of the Bell inequalities. The case of two spatially separated domains, in: *Journal of Soviet Mathematics* 36.4 (Feb. 1987), pp. 557–570, DOI: 10.1007/bf01663472.

[114] L. Wooltorton, P. Brown, and R. Colbeck, Device-independent quantum key distribution with arbitrarily small nonlocality, in: *Physical Review Letters* 132 (21 May 2024), p. 210802, DOI: 10.1103/PhysRevLett.132.210802.

[115] V. Barizien, P. Sekatski, and J.-D. Bancal, Custom Bell inequalities from formal sums of squares, in: *Quantum* 8 (May 2024), p. 1333, DOI: 10.22331/q-2024-05-02-1333.

[116] S. Wagner et al., Device-independent characterization of quantum instruments, in: *Quantum* 4 (Mar. 2020), p. 243, DOI: 10.22331/q-2020-03-19-243.

[117] S. Pironio, M. Navascués, and A. Acín, Convergent relaxations of polynomial optimization problems with noncommuting variables, in: *SIAM Journal on Optimization* 20.5 (2010), pp. 2157–2180, DOI: 10.1137/090760155.

[118] C. Bamps and S. Pironio, Sum-of-squares decompositions for a family of Clauser-Horne-Shimony-Holt-like inequalities and their application to self-testing, in: *Physical Review A* 91 (5 May 2015), p. 052111, DOI: 10.1103/PhysRevA.91.052111.

[119] M. Navascués et al., Almost quantum correlations, in: *Nature Communications* 6.1 (Feb. 2015), DOI: 10.1038/ncomms7288.

[120] V. Barizien, "The set of quantum correlations: Bell inequalities, self-testing and convex geometry", PhD thesis, 2024.

[121] L. Masanes, *Extremal quantum correlations for N parties with two dichotomic observables per site*, 2005, arXiv: quant-ph/0512100 (quant-ph).

[122] A. Mikos-Nuszkiewicz and J. Kaniewski, Extremal points of the quantum set in the Clauser-Horne-Shimony-Holt scenario: conjectured analytical solution, in: *Physical Review A* 108 (1 July 2023), p. 012212, DOI: 10.1103/PhysRevA.108.012212.

[123] P. Wittek, Algorithm 950: ncpol2sdpa—sparse semidefinite programming relaxations for polynomial optimization problems of noncommuting variables, in: *ACM Transactions on Mathematical Software* 41.3 (June 2015), pp. 1–12, DOI: 10.1145/2699464.

[124] MOSEK ApS, *MOSEK Optimization Toolbox for Python*, Version 11.0.13, MOSEK ApS, Copenhagen, Denmark, 2024.

[125] J. W. Helton and M. C. de Oliveira, *NCAlgebra: A Noncommutative Algebra Package for Mathematica*, https://ncalgebra.github.io, Accessed: April 2025, 2023.

[126] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[127] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.

[128] E. Schrödinger, Die gegenwärtige Situation in der Quantenmechanik, in: *Naturwissenschaften* 23.49 (Dec. 1935), pp. 823–828, DOI: 10.1007/BF01491914.

[129] R. Horodecki et al., Quantum entanglement, in: *Reviews of Modern Physics* 81.2 (June 2009), pp. 865–942, DOI: 10.1103/RevModPhys.81.865.

[130] A. Acín et al., Three-qubit pure-state canonical forms, in: *Journal of Physics A: Mathematical and General* 34.35 (Aug. 2001), p. 6725, DOI: 10.1088/0305-4470/34/35/301.

[131] M. M. Wilde, *Quantum Information Theory*, 2nd ed., Cambridge University Press, 2017.

[132] M. A. Naimark, Spectral functions of a symmetric operator, in: *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya* 4.3 (1940), pp. 277–318.

[133] A. Peres, *Quantum Theory: Concepts and Methods*, Springer Netherlands, 2002.

[134] J. Preskill, *Lecture Notes for Physics 229: Quantum Computation*, 2004.

[135] J. Watrous, *The Theory of Quantum Information*, Cambridge University Press, 2018.

[136] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement*, Cambridge University Press, 2007.

[137] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[138] G. M. Ziegler, *Lectures on Polytopes*, Springer New York, 1995.

[139] L. Vandenberghe and S. Boyd, Semidefinite programming, in: *SIAM Reviews* 38.1 (Mar. 1996), pp. 49–95, DOI: 10.1137/1038003.

[140] P. Skrzypczyk and D. Cavalcanti, *Semidefinite Programming in Quantum Information Science*, 2053-2563, IOP Publishing, 2023, DOI: 10.1088/978-0-7503-3343-6.

[141] A. Tavakoli et al., Semidefinite programming relaxations for quantum correlations, in: *Reviews of Modern Physics* 96 (4 Dec. 2024), p. 045006, DOI: 10.1103/RevModPhys.96.045006.

[142] J. Löfberg, "YALMIP: A Toolbox for Modeling and Optimization in MATLAB", in: *In Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[143] S. Diamond and S. Boyd, CVXPY: A Python-embedded modeling language for convex optimization, in: *Journal of Machine Learning Research* 17.83 (2016), pp. 1–5.

[144] A. Acín, S. Massar, and S. Pironio, Randomness versus nonlocality and entanglement, in: *Physical Review Letters* 108.10 (Mar. 2012), p. 100402, DOI: 10.1103/PhysRevLett.108.100402.

[145] G. Adesso, M. Cianciaruso, and T. R. Bromley, *An introduction to quantum discord and non-classical correlations beyond entanglement*, 2016, arXiv: 1611.01959 (quant-ph).

[146] L. E. Ballentine, *Quantum Mechanics: A Modern Development*, World Scientific, 2015.

[147] J. Barrett et al., Nonlocal correlations as an information-theoretic resource, in: *Physical Review A* 71 (2 Feb. 2005), p. 022101, DOI: 10.1103/PhysRevA.71.022101.

[148] J. Barrett and S. Pironio, Popescu-Rohrlich correlations as a unit of nonlocality, in: *Physical Review Letters* 95 (14 Sept. 2005), p. 140401, DOI: 10.1103/PhysRevLett.95.140401.

[149] C. Branciard, Detection loophole in Bell experiments: how postselection modifies the requirements to observe nonlocality, in: *Physical Review A* 83 (3 Mar. 2011), p. 032123, DOI: 10.1103/PhysRevA.83.032123.

[150] G. Brassard et al., Limit on nonlocality in any world in which communication complexity is not trivial, in: *Physical Review Letters* 96 (25 June 2006), p. 250401, DOI: 10.1103/PhysRevLett.96.250401.

[151] P. Cieśliński et al., *How likely are you to observe non-locality with imperfect detection efficiency and random measurement settings?*, 2025, arXiv: 2503.21519 (quant-ph).

[152] M. Froissart, Constructive generalization of Bell's inequalities, in: *Il Nuovo Cimento B* 64.2 (Aug. 1981), pp. 241–251, DOI: 10.1007/bf02903286.

[153] M. Genovese, Research on hidden variable theories: A review of recent progresses, in: *Physics Reports* 413.6 (July 2005), pp. 319–396, DOI: 10.1016/j.physrep.2005.03.003.

[154] N. Gisin, Weinberg's non-linear quantum mechanics and supraluminal communications, in: *Physics Letters A* 143.1 (1990), pp. 1–2, DOI: 10.1016/0375-9601(90)90786-N.

[155] N. Gisin, *Can relativity be considered complete? From Newtonian nonlocality to quantum nonlocality and beyond*, 2005, arXiv: quant-ph/0512168 (quant-ph).

[156] N. Gisin, Quantum nonlocality: how does nature do it?, in: *Science* 326.5958 (2009), pp. 1357–1358, DOI: 10.1126/science.1182103.

[157] N. Gisin, *Quantum Chance: Nonlocality, Teleportation and Other Quantum Marvels*, Springer International Publishing, 2014, DOI: 10.1007/978-3-319-05473-5.

[158] R. Haag, *Local Quantum Physics: Fields, Particles, Algebras*, Springer Berlin Heidelberg, 1996.

[159] J. B. Lasserre, *Moments, Positive Polynomials and Their Applications*, Imperial College Press, 2009.

[160] J. Oppenheim and S. Wehner, The uncertainty principle determines the nonlocality of quantum mechanics, in: *Science* 330.6007 (2010), pp. 1072–1074, DOI: 10.1126/science.1192065.

[161] A. Peres, Unperformed experiments have no results, in: *American Journal of Physics* 46.7 (July 1978), pp. 745–747, DOI: 10.1119/1.11393.

[162] C. Pfister and S. Wehner, An information-theoretic principle implies that any discrete physical theory is classical, in: *Nature Communications* 4.1 (May 2013), DOI: 10.1038/ncomms2821.

[163] D. Rosset, J.-D. Bancal, and N. Gisin, Classifying 50 years of Bell inequalities, in: *Journal of Physics A: Mathematical and Theoretical* 47.42 (Oct. 2014), p. 424022, DOI: 10.1088/1751-8113/47/42/424022.

[164] T. Vidick and S. Wehner, More nonlocality with less entanglement, in: *Physical Review A* 83 (5 May 2011), p. 052310, DOI: 10.1103/PhysRevA.83.052310.

[165] T. Vidick and S. Wehner, *Introduction to Quantum Cryptography*, Cambridge University Press, 2023.

[166] S. Weinberg, *Dreams of a Final Theory: The Scientist's Search for the Ultimate Laws of Nature*, New York: Vintage Books, 1994.

[167] S. Weinberg, *The Quantum Theory of Fields*, Cambridge University Press, 1995.

[168] A. Zingarofalo, *Orthogonal faces in the CHSH scenario. Notebooks for the Master's Thesis*, July 2025, DOI: 10.5281/zenodo.15824241.